Network Lerner Index:
Demand and Distortions across Industries

Lingxuan Wu*

December 17, 2022

Abstract

What determines demand elasticities and distortions across industries, acknowledging that firms' demand and pricings are interdependent in input-output networks? I propose a modeling approach of *competition in markups*, which micro-founds demand elasticities and markups via "network Lerner indexes" based on the network structure of the economy in general equilibrium. At the equilibrium, each firm's markup depends on four industry-level network statistics—cycle, substitution in production, substitution in consumption, and impact on consumer price, in addition to its market share. Empirical analysis of Compustat firms from 1997 to 2019 confirms the theory predictions. Taking into account the sectoral heterogeneity explains about three times more variations in markups. Under the theory-predicted markups, the loss in total factor productivity (TFP) due to misallocation is about 13% over the sample period, which is four times larger than the implied TFP loss that ignores sectoral heterogeneity.

Keywords: production networks, endogenous markups, sectoral heterogeneity, concentration, distortions, misallocation.

JEL: C72, D21, D24, D43, D5, D57, E1, L16, O4

*Harvard University, lingxuanwu@g.harvard.edu. Website: https://www.lingxuanwu.me. I 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), Elisa Rubbo, Kunal Sangani, Ludwig Straub, Maria Voronina, and participants at various seminars at Harvard and 2022 UZH PhD Macro Workshop for helpful comments. I thank Basile Grassi for sharing his codes.

Electronic copy available at: https://ssrn.com/abstract=4306329

1 Introduction

Lerner (1934) concludes his seminal work on market power with anticipation of more general analysis: "When our 'industry' becomes the whole society, there cannot be too much or too little resources used ... the larger the fraction of the whole society one wishes to examine, the less legitimate is it to use particular analysis. ... What is relevant for general analysis is not the *sum* of individual degrees of monopoly but their *deviations*." Lying behind this is a simple question, to which this paper is devoted: Are some parts of the economy more prone to market power? Much research has been conducted on firms' market power within an industry since Lerner (1934), but it is not yet understood whether and why some industries may "possess" a greater degree of market power than others.

When a firm with marginal cost c maximizes its profits facing a demand curve $D(p)$, it optimally charges a markup $\mu := \frac{p}{c}$ that satisfies the first-order condition

$$\frac{\mu}{\mu - 1} = -\frac{\mathrm{d} \log D(p)}{\mathrm{d} \log p}, \tag{1}$$

which links the inverse of the Lerner index to the demand elasticity. The markup is higher if the demand elasticity is lower. Thus an equivalent question is: Do some parts of the economy face lower demand elasticities than others?

As an analog to firm's demand elasticity, imagine if the prices charged by all oil producers go up, how will the oil industry's total sales change? If oil products are heavily used in industrial production, whereas food is directly sold to consumers, how can we compare their demand elasticities, which a priori need not equal? The difficulty entailed in this analysis is that as upstream firms raise their prices, the resulting changes in their demand stem from their downstream firms' decisions to adjust inputs, which in turn depend on how many products these downstream firms want to produce and sell themselves. The interdependent nature of demand and pricing decisions calls for a model of imperfect competition among all firms in a production network in general equilibrium, where the demand elasticities are equilibrium outcomes instead of given primitives.

I develop a new modeling approach of *competition in markups* in general equilibrium to address these questions, since classic Bertrand and Cournot approaches are not applicable to generic networks.[1] My approach micro-founds the demand elasticities of industries

[1] These approaches with simultaneous moves are ill-defined when vertical relations are involved. While a model with sequential price setting works in simple vertical chains, it become either intractable when the

Electronic copy available at: https://ssrn.com/abstract=4306329

and firms, based on the production technologies and consumer preference in an economy. The key idea is that each firm competes by setting its markup, which is like setting a price. When upstream firms consider their pricings, they assume that their downstream firms will hold profit margins fixed and adjust their demand accordingly.[2] In canonical applications, absent any vertical relations, competition in markups boils down to the familiar Bertrand competition. I discuss its properties in detail in Section 2.2.

This modeling approach makes it possible to characterize demand elasticities in an economy with arbitrary input-output linkages. In a nested CES network, my pricing formula generalizes Atkeson and Burstein (2008)'s Bertrand-competition formula for a horizontal economy to reflect each industry's "essentialness." The markup μ_{f_i} charged by firm f_i in industry i obeys

$$\frac{\mu_{f_i}}{\mu_{f_i} - 1} = \left(1 - s_{f_i}\right)\rho + s_{f_i}\eta_i. \tag{2}$$

Its markup depends on the demand elasticity it faces, which is an average of within-industry elasticity of substitution ρ and *variable* industry-level demand elasticity η_i, weighted by the firm's market share s_{f_i}.[3]

I show that industry-level demand elasticity η_i satisfies

$$\eta_i = \text{cycle}_i + \left(\theta_p - 1\right)\text{SP}_i + \left(\theta_c - 1\right)\text{SC}_i + \left(\epsilon - 1\right)\text{CP}_i + 1, \tag{3}$$

in which $\text{cycle}_i, \text{SP}_i, \text{SC}_i, \text{CP}_i$ are four network statistics related to industry i's network position, θ_p is the elasticity of substitution of a CES production technology, θ_c is the elasticity of substitution of a CES final consumption function, and ϵ is the price elasticity of consumer demand. The model makes the following predictions: (i) Industries that are in stronger *cycles* have higher demand elasticities, as a higher markup feeds into their own production costs and prices, dragging down their own demand. (ii) Demand elasticity η_i increases in the elasticities of substitution θ_p, θ_c, modulated by the strength of *substitution in production and consumption* induced by each industry, SP_i, SC_i. Intuitively, the more substitutable goods are, the more elastic their demand. (iii) η_i also increases in the price elasticity of consumer demand ϵ, modulated by industry i's weight in *consumer price*, CP_i. This is because consumer demand contracts when goods get more expensive, which scales

network gets more complicated, or infeasible when cycles are involved in which case there is no natural or finite sequence of moves to apply. Section 2.2 provides a detailed discussion.

[2] Hall and Hitch (1939); Scherer and Ross (1990); von Ungern-Sternberg (1999) argue on factual grounds that business managers form their pricing strategies in percentage markups instead of prices or quantities.

[3] When $\rho > \eta_i$, the larger firms within each industry face lower elasticities and charge larger markups.

Electronic copy available at: https://ssrn.com/abstract=4306329

down the total demand (including intermediate demand) for each good.

Formulas (2, 3) jointly predict that each firm's markup is a function of five sufficient statistics: the firm's within-industry market share s_{f_i} and its interactions with the four industry-level network statistics underlying η_i. In this sense, the Lerner index becomes "networked." The prediction holds at the equilibrium under any primitives, and all variables involved are observed equilibrium outcomes, facilitating my empirical test.

In Section 4.2, I derive novel theory implications on concentration and distortions, under variable industry-level demand elasticity η_i. Within each industry, the more productive firms charge higher markups, as in standard models of imperfect within-industry competition. Perhaps surprisingly, across industries or across time, holding fixed the firms' productivity advantages, less elastic industries see *lower* concentration and *higher* markups. The prediction cautions against naively associating concentration positively with market power, echoing Syverson (2019); Berry, Gaynor and Scott Morton (2019).

To test the pricing formulas (2, 3), I measure η_i based on network statistics constructed from the US Bureau of Economic Analysis (BEA) input-output table from 1997 to 2019. To measure markups, I take two approaches for robustness: a production-function approach following De Loecker, Eeckhout and Unger (2020); De Ridder, Grassi and Morzenti (2022), and a cost-share approach from Bond et al. (2021); Edmond, Midrigan and Xu (2022). Both approaches identify the slope of log markup to log market share in each industry.[4] My results show that all four network statistics are strongly predictive of the markup-to-market-share relations across industries. Taking the heterogeneity across industries into account raises the predictive power of these regressions by about three times in the production-function approach and more in the cost-share approach, attesting to its quantitative importance. Further, their signs and relative magnitudes reflect the elasticities in production technologies and consumer preference. My calibration delivers that $\theta_p, \theta_c < 1$ (i.e., goods are complementary) and $\epsilon > 1$ (i.e., consumer demand contracts when prices go up), as we may expect intuitively.

As an application, I tentatively quantify the contribution of sectoral heterogeneity in the loss of total factor productivity (TFP) due to misallocation. For this exercise, I use the theory-predicted markups instead of the estimated markups from production function estimation. Misallocation leads to an average TFP loss of about 13% over the sample

[4]It is important to note that the markups that are estimated from the production-function approach using revenue data from Compustat are biased in levels, but have valid log variations across firms. (Bond et al., 2021; De Ridder, Grassi and Morzenti, 2022) My empirics only use the variation of markups.

Electronic copy available at: https://ssrn.com/abstract=4306329

period. This number is four times larger than the implied loss if one ignores sectoral heterogeneity. Comparing to the literature, Baqaee and Farhi (2020) show that the total gain in TFP from eliminating markups is about 10% to 25%, while Edmond, Midrigan and Xu (2022) suggest instead a number of just around 2% to 6%. Both numbers include the TFP loss due to aggregate markup, in addition to misallocation, which I focus on. The major difference between these two papers is that Baqaee and Farhi (2020) use estimated markups that are more dispersed than the size-dependent markups simulated in Edmond, Midrigan and Xu (2022) under homogeneous industry-level demand elasticities.[5] Though the specifications vary across these two papers and mine and the numbers should be interpreted with caution, the takeaway from my exercise is that taking into account the sectoral heterogeneity explains more variations in markups than using only market shares, and enlarges the TFP loss due to misallocation, which helps fill in the gap between Baqaee and Farhi (2020) and Edmond, Midrigan and Xu (2022).

1.1 Literature Review

This paper is most closely related to Atkeson and Burstein (2008); Grassi (2017); Baqaee and Farhi (2020); Edmond, Midrigan and Xu (2022). I generalize the Bertrand approach in Atkeson and Burstein (2008) from a horizontal economy to an input-output network to understand variable industry-level demand elasticities and show its empirical support. Grassi (2017) and Baqaee and Farhi (2020) have studied how firm-level shocks affect aggregate outputs due to their network positions. Grassi (2017) models variable markups, assuming that firms compete without internalizing the impacts of their pricing on downstream firms' demand.[6] That model does not predict variable industry-level demand elasticities. Baqaee and Farhi (2020) assume exogenous markups. This paper builds on the analytical tools they develop and pushes one step further by asking: if different industries matter differently in production networks, do they have different degrees of market power?[7] This paper shares the theme of demand-related distortions and misallocation

[5] In Edmond, Midrigan and Xu (2022), homogeneous industry-level demand elasticities are a result of their assumption of a continuum of industries, similar to Atkeson and Burstein (2008).

[6] Dhyne, Kikkawa and Magerman (2022) build a similar model to study market power in firm-to-firm trade, instead of a generic input-output network.

[7] The premise of this question is that individual firms matter for the aggregate economy. (Gabaix, 2011) My model reduces to monopolistic competition when there is a continuum of homogeneous firms within each industry, and reduces to Atkeson and Burstein (2008) when there is a continuum of industries producing consumption goods. In reality, firms are highly heterogeneous in size and productivity, and industries are of a small number and linked in networks.

Electronic copy available at: https://ssrn.com/abstract=4306329

studied in Edmond, Midrigan and Xu (2022), but differs in that I model sectoral heterogeneity and show that it is important for explaining markups and understanding welfare cost due to misallocation.[8]

The work on production networks dates back to Leontief (1936); Long and Plosser (1983). Recently, Acemoglu et al. (2012) and Baqaee and Farhi (2019) study the propagation of microeconomic shocks in efficient networks. Atalay (2017) estimates the elasticity of substitution among intermediate inputs in production, from the relative movements of prices and quantities of different goods over the business cycle. My empirical exercise of predicting firm markups yields an estimate of the elasticity of substitution too. As Atalay specifies a more nuanced production function than I do, the numbers are not directly comparable, but both indicate that goods are highly complementary in production.

Jones (2011, 2013) as well as Liu (2019); Baqaee and Farhi (2020); Bigio and La'O (2020) study efficiency implications of *exogenous* wedges in networks. In contrast, I endogenize markups. Acemoglu and Tahbaz-Salehi (2020) embed bilateral bargaining in production networks, which gives rise to heterogeneous markups, under the assumptions that each firm has at most one customer and at most one customized supplier in any industry and that the production function is Leontief (for their main analysis) or uses inputs as gross complements (as an extension). Their equilibrium is characterized by an implicit fixed-point system.[9] Bizzarri (2022) develops a model of competition in supply and demand functions in production networks. To avoid indeterminacy of best responses, he introduces stochastic parameters to demand and production functions. As a generic result, the distribution of these stochastic parameters matters for the equilibrium. The equilibrium is characterized by a set of partial differential equations in general, and by an implicit fixed-point system in the special case of competition in linear schedules. Unfortunately, linear schedules are not supported at equilibrium under familiar functional forms of production technology, such as translog and CES. While I abstract away from buyer market power which Acemoglu and Tahbaz-Salehi (2020) and Bizzarri (2022) analyze theoretically, my model explicitly characterizes the demand elasticities and markups in terms of industries' network positions that are observed, with no restrictions on the network structure or production function. It is advantageous not only for its tractability in a truly generic network, but also allows us to see the intuition behind and test it empirically.

[8] The welfare cost due to misallocation in general is discussed in Restuccia and Rogerson (2008); Hsieh and Klenow (2009) before Baqaee and Farhi (2020); Edmond, Midrigan and Xu (2022).

[9] Alviarez et al. (2022); Voronina (2022) study two-sided market power with bargaining too, but only in production chains and under other simplifying assumptions.

Electronic copy available at: https://ssrn.com/abstract=4306329

Chaney (2018); Oberfield (2018) and Acemoglu and Azar (2020), among others, study the endogenous formation of networks. My model focuses on endogenous markups given the network structure, which is an important second stage in a larger model: the profits firms can make affect their ex-ante decisions to form linkages.

Variable markups can be driven by other mechanisms: see Melitz and Ottaviano (2008) and references therein for entry; Peters (2020) for endogenous firm productivity growth; and Afrouzi, Drenik and Kim (2021) for endogenous customer acquisition of retail firms. My model of endogenous demand elasticities in networks is complementary to these mechanisms, in that given the number of firms operating and their productivities, the markups they charge depend on the demand elasticities they face. It is well known that variable markups can be generated by flexible demand functional forms, such as Kimball (Kimball, 1995; Klenow and Willis, 2016) and translog (Feenstra, 2003). This paper micro-founds demand elasticities, instead of postulating them as primitives.

A host of recent papers document declining labor share (Karabarbounis and Neiman, 2014; Kehrig and Vincent, 2021), rising concentration (Gutiérrez and Philippon, 2017; Autor et al., 2020), and rising estimated markups (De Loecker, Eeckhout and Unger, 2020). One point of contention is which measure of concentration one should look at, and whether the popular production-function approach of estimating markups without quantity data is valid.[10] This paper contributes a model that explains markups not only within each industry but also across industries, and derives novel theory implications for concentration across industries.[11] The production-function approach used to estimate markups is originated by Hall (1986, 1988). Hall uses it to show the heterogeneity of markups across industries, with which this paper shares a similar theme but with a theoretical approach.

Last, this paper connects to various models of imperfect competition in partial and general equilibrium, which I will discuss in greater detail in Section 2.2. The conduct assumption of competing in markups focuses on seller market power, and thus naturally

[10] See Rossi-Hansberg, Sarte and Trachter (2021); Benkard, Yurukoglu and Zhang (2021) for the former, and Bond et al. (2021); De Loecker (2021); De Ridder, Grassi and Morzenti (2022) for the latter. Basu (2019) discusses different methods to estimate markups.

[11] 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. It was later discredited as researchers realized that the variables involved are all endogenous and that it is not possible to interpret these associations without a model. (Schmalensee, 1989; Bresnahan, 1989) This paper provides such a model and distinguishes between relations that can be tested based on observed industries' network positions (Propositions 1 and 2) and ones that hinge on unobserved primitives (Proposition 3).

Electronic copy available at: https://ssrn.com/abstract=4306329

relates to classic Cournot and Bertrand approaches in partial equilibrium analysis. These two approaches are also extended to general equilibrium by Gabszewicz and Vial (1972); FitzRoy (1974) and Marschak and Selten (1974); Nikaido (1975); Benassy (1988), respectively.[12] To the best of my knowledge, there has not been much success in accommodating generic network structures with these two approaches. The general equilibrium concept with competition in markups I introduce has much in common with the Bertrand approach, but works for networks. Importantly, it is an explicit model of a non-cooperative game firms play, in contrast to the subjective approach (Negishi, 1961, 1972), which can be too "subjective" instead of predictive.[13] (Hart, 1985)

2 General Equilibrium with Competition in Markups

2.1 Equilibrium Definition

There exist a finite number of firms with arbitrary input-output linkages, and a representative consumer that consumes and supplies labor. I use the wage as the numeraire.

The consumer consumes a constant-returns bundle of all goods in this economy,

$$C = C\left(\{c_i\}_{i\in\mathcal{N}}\right), \tag{4}$$

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

$$\sum_{i\in\mathcal{N}} p_i c_i = L + T, \tag{5}$$

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

Each firm i uses a constant-returns production technology, combining intermediate inputs and labor,

$$y_i = z_i F_i\left(\left\{x_{ij}\right\}_{j\in\mathcal{N}}, l_i\right), \tag{6}$$

in which x_{ij} is the quantity of good j as input used by firm i. Correspondingly, there exists

[12] Azar and Vives (2021) present a recent development of the Cournot approach to study common ownership in general equilibrium. Ederer and Pellegrino (2022) study the welfare implication of rising common ownership. Neither analyzes input-output networks.

[13] Partial equilibrium models with conjectural variations that nest the Cournot and Bertrand approaches trace back to Bowley (1924). Salinger (1989) studies vertical relations with conjectural variations.

Electronic copy available at: https://ssrn.com/abstract=4306329

a constant-returns cost function so that the marginal cost of firm i is $\frac{1}{z_i} mc_i\left(\left\{p_j\right\}_{j\in\mathcal{N}}\right)$.

A general equilibrium with *competition in markups*, built on a standard general equilibrium, gives rise to a set of markups $\boldsymbol{\mu} = \{\mu_i\}_{i\in\mathcal{N}}$ firms optimally choose, and a set of prices and quantities determined in the corresponding general equilibrium given $\boldsymbol{\mu}$.

Definition 1. (Fixed-markup general equilibrium) Given markups $\boldsymbol{\mu} = \{\mu_i\}_{i\in\mathcal{N}}$, a fixed-markup general equilibrium is a system of prices and quantities $\left\{p_i, x_{ij}, y_i, l_i, c_i, L\right\}_{i,j\in\mathcal{N}}$ such that

(i) Each firm i chooses inputs $\left\{x_{ij}, l_j\right\}_{j\in\mathcal{N}}$ to minimize its cost, produces $y_i = z_i F_i\left(\left\{x_{ij}\right\}_{j\in\mathcal{N}}, l_i\right)$, and charges markup μ_i;

(ii) The consumer chooses her consumption c_i and labor supply L to maximize $U\left(C\left(\{c_i\}_{i\in\mathcal{N}}\right), L\right)$, subject to her budget constraint;

(iii) The goods and labor markets clear, $c_i + \sum_{j\in\mathcal{N}} x_{ji} = y_i, \forall i$, and $\sum_{i\in\mathcal{N}} l_i = L$.

As a result, firm i earns a (real) profit of

$$\pi_i(\boldsymbol{\mu}) = \left(1 - \frac{1}{\mu_i}\right) R_i(\boldsymbol{\mu}) = \left(1 - \frac{1}{\mu_i}\right) \frac{p_i(\boldsymbol{\mu})\, y_i(\boldsymbol{\mu})}{P(\boldsymbol{\mu})}, \tag{7}$$

in which $R_i(\boldsymbol{\mu}) = \frac{p_i(\boldsymbol{\mu}) y_i(\boldsymbol{\mu})}{P(\boldsymbol{\mu})}$ is the real revenue of firm i, $p_i(\boldsymbol{\mu}), y_i(\boldsymbol{\mu})$ are the equilibrium price and quantity produced by firm i, and $P(\boldsymbol{\mu})$ is the equilibrium price index of the consumption bundle. The reason we need to use the real profit is simply that in general equilibrium, the firm owner should not be interested in monetary profit *per se*, but rather in what that profit can buy. (Gabszewicz and Vial, 1972)[14]

Definition 2. (General equilibrium with competition in markups) A general equilibrium with competition in markups is a pure-strategy Nash equilibrium among firms each with action μ_i to maximize its profit $\pi_i(\boldsymbol{\mu})$ defined by (7), which is determined in a general equilibrium with fixed markups $\boldsymbol{\mu}$.

Throughout this paper, we study the equilibrium under competition in markups, which characterizes imperfect competition among any number of firms in an economy, including the limiting cases where there is only one firm or an infinite amount of firms. It should be immediately clear that the choice of numeraire in general equilibrium does not

[14] See Hart (1985); Azar and Vives (2021) for a discussion of ownership and firm objective, which is not the focus of this paper.

Electronic copy available at: https://ssrn.com/abstract=4306329

affect the equilibrium markups or quantities, since the payoff function (7) is invariant to the numeraire.

In this paper, we will call $1 - \frac{\partial \log R_i}{\partial \log \mu_i}$ the *demand elasticity* (with respect to markup). It is a proper generalization of the usual price elasticity of demand: in our model, firms compete in setting markups, which is like a kind of price. Indeed, if a firm's marginal cost does not depend on its own markup, in the sense that will become clear later, the demand elasticity w.r.t. markup is exactly equivalently to the demand elasticity w.r.t. price.

2.2 Discussion of the Equilibrium Concept

The concept of general equilibrium with competition in markups consists of two parts: a conduct assumption of competing in markups, and a payoff function for each firm, which derives from a fixed-markup general equilibrium with no restriction on the network structure of production. The conduct assumption facilitates the tractable payoff function from general equilibrium. The interpretation of the conduct assumption is that, when each firms sets its own profit margin, it assumes other firms will keep their profit margins unchanged. Consequently, in production networks, as an upstream firm raises its markup, it anticipates that downstream firms will pass the increase in their input costs into their prices, thus changing their production and demand for the upstream firm. Each firm sets its markup to maximize its profits, facing its own downward-sloping demand curve.

Admittedly the analysis of imperfect competition, production networks and general equilibrium each is a distinct challenge and has resulted in a voluminous literature. An attempt to accommodate them at once inevitably involves some stylized assumption for tractability. Now we examine the concept of competition in markup from different angles, and show that it is versatile in handling generic production structure in general equilibrium, and is similar to familiar benchmarks in partial equilibrium when comparable.

1. Let us study a chain (the simplest network) with given downstream demand elasticity. Take the textbook example (Tirole, 1988, Section 4.1.1) of double marginalization in a vertical relation illustrated in Figure 1. There is a manufacturer (firm 1) with constant marginal cost c selling to a retailer (firm 2), which in turn sells to the consumer whose demand is $D(p_2)$.

Electronic copy available at: https://ssrn.com/abstract=4306329

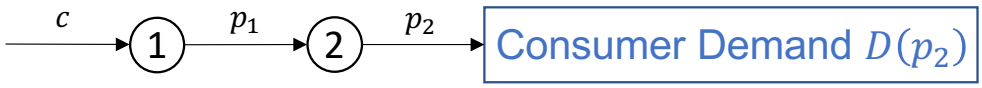


Figure 1: Simplest Vertical Relation

(a) Simultaneous moves in setting prices: If the quantity firm 2 buys from firm 1 is always $D(p_2)$, and if firm 1 assumes that firm 2 will hold p_2 fixed when setting p_1, then firm 1 faces an inelastic demand, and thus wants to charge an infinite price. This exposes the ridiculous result of simultaneous price setting, that has motivated researchers to turn to other assumptions, such as sequential moves.

(b) Sequential moves in setting prices: Consider the commonly studied linear contract, under which the manufacturer sets a price p_1 and the retailer subsequently determines p_2 given its input cost p_1. Both firms maximize their profits. Two firms' markups satisfy

$$\frac{\mu_2}{\mu_2-1}=-\frac{\mathrm{d}\log D}{\mathrm{d}\log p_2},\ \frac{\mu_1}{\mu_1-1}=-\frac{\mathrm{d}\log D}{\mathrm{d}\log p_2}\frac{\mathrm{d}\log p_2}{\mathrm{d}\log p_1},$$

where $-\frac{\mathrm{d}\log D}{\mathrm{d}\log p_2}$ and $\frac{\mathrm{d}\log p_2}{\mathrm{d}\log p_1}=\left(1+\frac{\mathrm{d}^2\log D}{\mathrm{d}\left(\log p_2\right)^2}\left(\mu_2-1\right)\right)^{-1}$ denote the final demand elasticity and cost pass-through of firm 2. For firm 1's demand to be downward sloping, the key is that it anticipates downstream price adjustment $\frac{\mathrm{d}\log p_2}{\mathrm{d}\log p_1}>0$, a behavior that motivates this paper's approach. In generic networks, note that the complexity of solving a sequential-move model with backward induction scales with the size of the network. Further, in networks that involve cycles, there is no natural order or even a finite sequence of moves one can apply.

(c) Simultaneous moves in setting markups: When two firms simultaneously set their markups, understanding that the final price $p_2=\mu_2\mu_1 c$ features double marginalization, they will choose

$$\frac{\mu_2}{\mu_2-1}=\frac{\mu_1}{\mu_1-1}=-\frac{\mathrm{d}\log D}{\mathrm{d}\log p_2}.$$

This is because firm 1 assumes that firm 2 will hold its markup fixed, i.e. $\frac{\mathrm{d}\log p_2}{\mathrm{d}\log p_1}=1$, and thus will face the same demand elasticity as firm 2. While its prediction differs slightly from the linear contract studied above, we can take

Electronic copy available at: https://ssrn.com/abstract=4306329

comfort in the well-behaved demand of firm 1 without resorting to sequential moves.[15] This property carries through to any network structure of production, where solving a sequantial-move model can be intractable or infeasible.[16]

2. The partial equilibrium analysis above takes as its premise some final demand function. However, in general equilibrium the primitives become consumer preferences and endowments, and firms' production technologies. Firms' objective functions have to be characterized in terms of these general equilibrium primitives. Previous researchers have extended both the Cournot and Bertrand approaches to general equilibrium analysis. However, to the best of my knowledge, there has not been much success in accommodating generic production networks with these approaches, possibly due to the difficulty in handling vertical relations with simultaneous choices in quantities or prices. (It is manifested even in partial equilibrium as discussed above.)[17] In contrast, the conduct of competition in markups resolves this difficulty by letting firms anticipate demand adjustments, and can be easily adapted to general equilibrium analysis as shown in Section 2.1. It takes advantage of a standard general equilibrium with fixed markups, whose characterization comes as close as possible to a perfectly competitive general equilibrium.

3. We make a final remark that in many models in macroeconomics and trade that do not involve a vertical relation, competition in markups can be exactly equivalent to competition in prices. Imagine a retail industry with a set $\mathcal{N}$ of firms and a consumer demand function $D_i\left(\left\{p_j\right\}_{j\in\mathcal{N}}\right)$ for each firm i. Under the common assumption that each firm has a constant marginal cost c_i, since each price p_i is simply the product of markup μ_i and marginal cost c_i, competition in markups and competition in prices are equivalent, and each firm's markup satisfies

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

[15] Obviously, assuming sequential moves in markups is equivalent to assuming sequential price setting.

[16] Young (1991); Grant and Quiggin (1994); Irmen (1997); von Ungern-Sternberg (1999) first present analysis of simultaneous moves in markups in simple environments, but none of them attempts to generalize it to more complicated networks or general equilibrium.

[17] While one can define the Cournot or Bertrand approach in restricted setups involving intermediate goods, it becomes implausible for an upstream firm to suppose that if it raises its own output or price, its downstream firms will not change their production plans or the prices accordingly. See Hart (1985) for a more detailed discussion of these issues and a comprehensive overview of seminal papers that extend Cournot and Bertrand approaches to general equilibrium.

Electronic copy available at: https://ssrn.com/abstract=4306329

It is also important to clarify that, while this paper micro-founds the demand elasticity for each firm from a model of imperfect competition among all firms, the *only* thing each firm needs to know for its pricing decision is how its sales will change if it adjusts its price locally. The firm does not have to know what other firms are doing beyond what is already conveyed in the demand elasticity as the sufficient statistic. It is neither more nor less demanding to know the demand elasticity in this model than in any canonical industrial organization model of a particular industry. In reality, firms are certainly not all-knowing game theorists, but they can learn through experiments.[18]

2.3 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-labor tradeoff for simplicity henceforth, as is common in the literature.[19]

The consumer's utility is

$$U(C, L) = \frac{C^{1-\frac{1}{\epsilon-1}}}{1-\frac{1}{\epsilon-1}} - L, \tag{8}$$

with $\epsilon > 1$. This gives rise to the consumer demand under optimal consumption-labor choice, with the consumer price index P,

$$C = P^{1-\epsilon}. \tag{9}$$

The exponent of $1-\epsilon$ may appear unusual, but actually implies a demand elasticity of ϵ for the consumer demand in general equilibrium.[20] If there is only one firm in this economy and it does not use its own output in its production, its markup is $\frac{\epsilon}{\epsilon-1}$ (Example 1).

There are I industries indexed by $i \in \mathcal{I}$. Within each industry i there are F_i firms indexed by $f_i \in \mathcal{F}_i$. In total, there are $N = \sum_i F_i$ firms, and $\mathcal{N} = \cup_i \mathcal{F}_i$ denotes the collection of all firms. Each firm f_i produces y_{f_i} quantity of a distinct good. Goods produced by

[18] See Doraszelski, Lewis and Pakes (2018) for evidence of firms' learning in a new market.

[19] In a more general specification, the results I show can be derived similarly using the local elasticities of substitution and local elasticity of consumer demand.

[20] To see this, recall our more general definition of demand elasticity $1 - \frac{\partial \log R}{\partial \log P}$ based on revenue R. In partial equilibrium, an isoelastic demand function $D(P) = P^{-\epsilon}$ leads to a nominal revenue $R = PD = P^{1-\epsilon}$ and thus a demand elasticity of ϵ. In general equilibrium, since revenue is in real terms, a demand function $D(P) = P^{1-\epsilon}$ leads to a real revenue $R = D = P^{1-\epsilon}$ and thus a demand elasticity of ϵ.

Electronic copy available at: https://ssrn.com/abstract=4306329

firms in the same industry are combined into a CES industry composite i with elasticity ρ,

$$y_i = \left[\sum_{f_i \in \mathcal{F}_i} \left(y_{f_i} \right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}. \tag{10}$$

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

$$C = \left[\sum_{j \in \mathcal{I}} \phi_j c_j^{\frac{\theta_c-1}{\theta_c}} \right]^{\frac{\theta_c}{\theta_c-1}}. \tag{11}$$

Firms in the same industry use the same CES bundle of labor and industry composites in production with elasticity θ_p, but they differ in productivities. Each firm f_i produces, with productivity z_{f_i},

$$y_{f_i} = z_{f_i} \left[b_i l_{f_i}^{\frac{\theta_p-1}{\theta_p}} + \sum_{j \in \mathcal{I}} a_{ij} x_{f_i j}^{\frac{\theta_p-1}{\theta_p}} \right]^{\frac{\theta_p}{\theta_p-1}}, \tag{12}$$

in which $l_{f_i}, x_{f_i j}$ denote firm f_i's usage of labor and industry composite j as inputs. We use q_i to denote the price of the input bundle for industry i, and thus each firm's marginal cost is $\frac{q_i}{z_{f_i}}$. If firm f_i uses m_{f_i} quantity of input bundle, it produces $y_{f_i} = z_{f_i} m_{f_i}$. The total amount of the input bundle used by each industry is thus $m_i = \sum_{f_i \in \mathcal{F}_i} m_{f_i}$.

The market clears for each industry composite $j \in \mathcal{I}$

$$y_j = c_j + \sum_{f_i \in \mathcal{N}} x_{f_i j}, \tag{13}$$

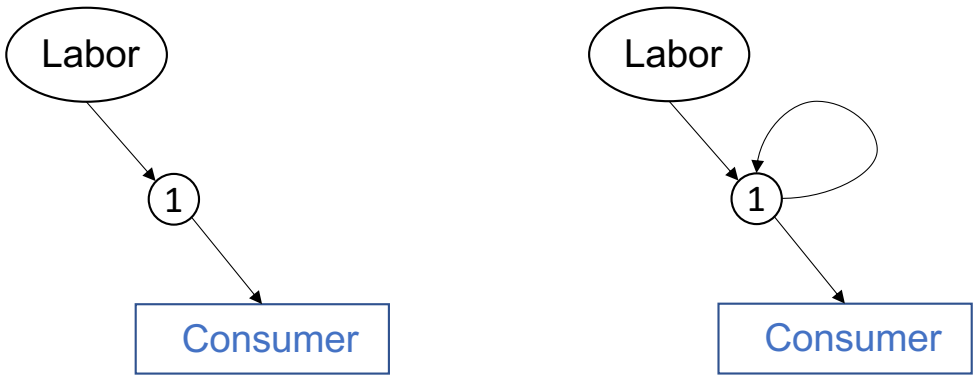
and for labor

$$L = \sum_{f_i \in \mathcal{N}} l_{f_i}. \tag{14}$$

2.4 Examples of One-Industry Economies

Before we introduce the complexity of strategic interaction, it is worth seeing some simplest examples where only one firm exists (Figure 2).

Electronic copy available at: https://ssrn.com/abstract=4306329



(a) Linear economy (b) Roundabout economy

Figure 2: Monopoly Economies

Example 1. (Monopoly linear economy, Figure 2a) In an economy in which there is only one firm that produces one good only from labor and sells it to the consumer, with $y = l$, the markup it charges satisfies

$$\frac{\mu}{\mu - 1} = \epsilon. \tag{15}$$

The markup is lower when the elasticity ϵ is higher.

Example 2. (Monopoly roundabout economy, Figure 2b) In an economy where there is only one firm that produces using labor and its own good according to a Cobb-Douglas production technology, $y = \left(\frac{x}{\alpha}\right)^{\alpha} \left(\frac{l}{1-\alpha}\right)^{1-\alpha}$, the markup it charges satisfies

$$\frac{\mu}{\mu - 1} = \frac{\epsilon - \alpha}{1 - \alpha} + \frac{\alpha}{\mu - \alpha}. \tag{16}$$

The elasticity on the RHS is a composite of ϵ, α, and μ. A higher weight on the intermediate input α and a higher elasticity ϵ both lead to a lower markup μ.[21]

3 Industry-Level Demand Elasticities

In this section, we determine the industry-level demand elasticity. For this purpose, it is intuitive if we assume there is only one firm in each industry, i.e. $F_i = 1$, and study the

[21] To see that, rewrite the expression as $\frac{1}{\mu-1} - \frac{\alpha}{\mu-\alpha} = \frac{\epsilon-1}{1-\alpha}$. Its LHS is decreasing in μ and α, and its RHS is increasing in ϵ and α, and thus the solution $\mu(\epsilon, \alpha)$ is decreasing in both ϵ and α.

Electronic copy available at: https://ssrn.com/abstract=4306329

demand elasticity it faces.[22] When we introduce multiple firms in each industry later, this elasticity will become the industry-level elasticity. With slight abuse of notation, here we use i to denote the monopoly firm in industry i and refer to an industry's optimal markup.

Each monopoly firm in industry i maximizes its profit

$$\pi_i(\boldsymbol{\mu}) = \left(1 - \frac{1}{\mu_i}\right)\frac{p_i y_i}{P} = \left(1 - \frac{1}{\mu_i}\right)\cdot\frac{p_i y_i}{PC}\cdot C = \underbrace{\left(1 - \frac{1}{\mu_i}\right)}_{\text{profit margin}}\cdot\underbrace{\lambda_i}_{\text{Domar weight}}\cdot\underbrace{P^{1-\epsilon}}_{\text{consumer demand}}, \tag{17}$$

in which $\lambda_i := \frac{p_i y_i}{PC}$ is the ratio of industry i's sales to GDP, i.e. its Domar weight.[23] With a given profit margin, the firm prefers larger sales relative to GDP together with a higher value of total consumer demand C. However, when each firm raises its markup, not only will its sales-to-GDP ratio be impacted as their buyers adjust their intermediate demand, but also the consumer demand will contract as the consumption price P gets higher, which scales down the total demand of i.

At the equilibrium, the Lerner condition of firm i is

$$\frac{\mu_i}{\mu_i - 1} = \eta_i,$$

in which the demand elasticity is

$$\eta_i := -\frac{\partial \log\left(p_i y_i / P\right)}{\partial \log \mu_i} + 1 = \underbrace{-\frac{\partial \log \lambda_i}{\partial \log \mu_i}}_{\text{elasticity of Domar weight}} + \underbrace{(\epsilon - 1)\frac{\partial \log P}{\partial \log \mu_i}}_{\text{consumer demand contraction}} + 1. \tag{18}$$

The markup μ_i is higher when the elasticity η_i is lower.

Next we introduce input-output definitions and characterize η_i in terms of industry i's network position in Proposition 1. The derivation builds on the elasticities of consumption price P and Domar weight λ_i to changes in markups μ_i in general equilibrium developed in Baqaee and Farhi (2020).

[22] For each monopoly industry to set a finite markup, one can assume for convenience that the industry-level demand elasticity η_i that we will determine is above 1. However, this restriction is not essential in a realistic environment with multiple firms in each industry, as investigated in Section 4, where within-industry competition will restrain their markups.

[23] When there are intermediate sales, we will have $\sum_i \lambda_i > 1$.

Electronic copy available at: https://ssrn.com/abstract=4306329

3.1 Input-Output Definitions

Include a fictitious firm 0, which exclusively produces the consumption good with a fixed markup $\mu_0 \equiv 1$. Include the labor factor, which does not use any input to produce, as a fictitious firm L too. Then we define input-output matrices of dimension $(1 + N + 1) \times (1 + N + 1)$ following Baqaee and Farhi (2020), where firm 0 is ordered first and firm L last.

Define the *final expenditure share* as

$$b_i := \frac{p_i c_i}{\sum_{j=1}^{N} p_j c_j}, \tag{19}$$

whose denominator is the nominal GDP.

Define the *cost-based input-output matrix* capturing the elasticities of marginal costs to price changes,

$$\tilde{\Omega}_{ij} := \frac{\partial \log mc_i}{\partial \log p_j} = \frac{p_j x_{ij}}{\sum_{k=1}^{N+L} p_k x_{ik} + l_i}, \tag{20}$$

in which the equality uses Shephard's lemma. $\tilde{\Omega}_{ij}$ is thus the cost share of input j for producer i, which characterizes how much directly goes to j for each dollar of inputs used by i. $\tilde{\Omega}_{0j}$ is the consumption expenditure share of good j. The last row of $\tilde{\Omega}$ is zero, corresponding to the labor factor, and its first column is zero too, as no firm uses consumption good 0 in its production.

Define the *cost-based Leontief inverse* as

$$\tilde{\Psi} := \left(I - \tilde{\Omega}\right)^{-1} = \sum_{n=0}^{\infty} \tilde{\Omega}^n. \tag{21}$$

The Leontief inverse captures both direct and indirect linkages among firms. For example, $\tilde{\Omega}^n$ is the indirect use of input through n steps. We define $\tilde{\lambda}_i := \tilde{\Psi}_{0i}$, which is a measure of the importance of good i in final demand.

Define the *revenue-based input-output matrix* as

$$\Omega_{ij} := \frac{p_j x_{ij}}{p_i y_i} = \frac{\tilde{\Omega}_{ij}}{\mu_i}, \tag{22}$$

which is the revenue share of input j of producer i. It characterizes how much directly goes to j for each dollar earned by i. We have $\Omega_{0j} \equiv \tilde{\Omega}_{0j}$ since $\mu_0 \equiv 1$.

Electronic copy available at: https://ssrn.com/abstract=4306329

Define the *revenue-based Leontief inverse* as

$$\Psi := (I - \Omega)^{-1} = \sum_{n=0}^{\infty} \Omega^n. \tag{23}$$

Ψ_{ij} captures how much directly or indirectly goes to j for each dollar of sales by i. Ψ_{0i} is exactly each industry's Domar weight λ_i, as it is the proportion of payment by the consumer that ultimately goes to firm i.

We also introduce the input-output covariance operator

$$\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(k)}, \Psi_{(l)}\right) := \sum_{i=0}^{N+L} \tilde{\Omega}_{ji}\tilde{\Psi}_{ik}\Psi_{il} - \left(\sum_{i=0}^{N+L} \tilde{\Omega}_{ji}\tilde{\Psi}_{ik}\right)\left(\sum_{i=0}^{N+L} \tilde{\Omega}_{ji}\Psi_{il}\right), \tag{24}$$

in which $\tilde{\Omega}^{(j)}$ is the j-th row of $\tilde{\Omega}$, and $\tilde{\Psi}_{(k)}, \Psi_{(l)}$ are the k-th and l-th columns of $\tilde{\Psi}, \Psi$ respectively. This is a covariance, since each row of $\tilde{\Omega}$ that represents a producer has non-negative entries which sum up to one.

3.2 Industry-Level Demand Elasticities

Proposition 1. (Industry-level demand elasticity) *The demand elasticity of industry i,* $\eta_i := -\frac{\partial \log\left(p_i y_i / P\right)}{\partial \log \mu_i} + 1$, *satisfies*

$$\eta_i = \underbrace{cycle_i}_{\text{feedback via cycle}} + \underbrace{\left(\theta_p - 1\right) SP_i}_{\text{substitution in production}} + \underbrace{\left(\theta_c - 1\right) SC_i}_{\text{substitution in consumption}} + \underbrace{\left(\epsilon - 1\right) CP_i}_{\text{consumer demand response}} + 1, \tag{25}$$

with

$$\begin{aligned} cycle_i &= \Psi_{ii} - 1, \\ SP_i &= \sum_{j \in \mathcal{I}} \left[\frac{\Psi_{0j}}{\mu_j \Psi_{0i}} \mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(i)}, \Psi_{(i)}\right)\right], \\ SC_i &= \frac{1}{\lambda_i} \mathrm{Cov}_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(i)}, \Psi_{(i)}\right), \\ CP_i &= \tilde{\lambda}_i. \end{aligned} \tag{26}$$

This proposition suggests that the industry-level demand elasticity η_i additively depends on four network statistics, modulated by elasticity of substitution in production θ_p, elasticity of substitution in consumption θ_c, and consumer demand elasticity ϵ. These

Electronic copy available at: https://ssrn.com/abstract=4306329

four network statistics are generically non-negative, hence η_i increasing in $\theta_p, \theta_c, \epsilon$. The monotonicity in θ_p, θ_c suggests that if goods are more substitutable in use, their producers face higher demand elasticities. The monotonicity in ϵ is intuitive, too, as a lower elasticity in final consumer demand leads to lower demand elasticities across the economy.[24] The four network statistics vary across industries, which we discuss shortly. Industry i has a low demand elasticity η_i 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 $\epsilon > 1$), which we will test empirically.

To derive and discuss each network statistic, we use eq. (18) which shows that η_i additively depends on the elasticity of consumer demand $(\epsilon - 1)\frac{\partial \log P}{\partial \log \mu_i}$ and the elasticity of its Domar weight $-\frac{\partial \log \lambda_i}{\partial \log \mu_i}$. The elasticity of consumer price to each firm i's markup is

$$\frac{\partial \log P}{\partial \log \mu_i} = \tilde{\Psi}_{0i} = \tilde{\lambda}_i. \tag{27}$$

In fact, the elasticity of any price p_j to markup μ_i is captured by $\tilde{\Psi}_{ji}$. The intuition is that when firm i raises its markup μ_i, it directly impacts those firms k who buy from it by $\tilde{\Omega}_{ki}$. Since those firms will keep their profit margins, their price adjustments will have further impacts on firms that buy from them. Via many rounds of iterations, the total response of price p_j to markup μ_i will sum up to $\tilde{\Psi}_{ji}$. When we are interested in the consumer price P, the more important good i is in final demand, the larger price response $\frac{\partial \log P}{\partial \log \mu_i}$ is. This leads to the CP term in eq. (25).

The elasticity of Domar weight $-\frac{\partial \log \lambda_i}{\partial \log \mu_i}$ is

$$-\frac{\partial \log \lambda_i}{\partial \log \mu_i} = \underbrace{(\Psi_{ii} - 1)}_{\text{feedback via cycle}} + \underbrace{\left(\theta_p - 1\right) \sum_{j \in \mathcal{I}} \left[\frac{\Psi_{0j}}{\mu_j \Psi_{0i}} \operatorname{Cov}_{\tilde{\Omega}^{(j)}} \left(\tilde{\Psi}_{(i)}, \Psi_{(i)} \right) \right]}_{\text{substitution in production (SP)}} + \underbrace{(\theta_c - 1) \frac{1}{\lambda_i} \operatorname{Cov}_{\tilde{\Omega}^{(0)}} \left(\tilde{\Psi}_{(i)}, \Psi_{(i)} \right)}_{\text{substitution in consumption (SC)}}. \tag{28}$$

The first term $\Psi_{ii} - 1$ is positive only if industry i is in a cycle of arbitrary length. The intuition is that, when industry i charges a higher markup, it causes the input costs of its downstream to go up and their prices to rise, too. If ultimately industry i buys from its downstream industries, that feeds into its own input cost and price, dragging down its

[24] This property implies that if the consumer demand elasticity is lowered over time as suggested by Sangani (2022); Nord (2022), it will lead to rising markups of not only retailers, but all the firms.

Electronic copy available at: https://ssrn.com/abstract=4306329

sales. This is exactly the additional term featured in Example 2 compared to Example 1.[25]

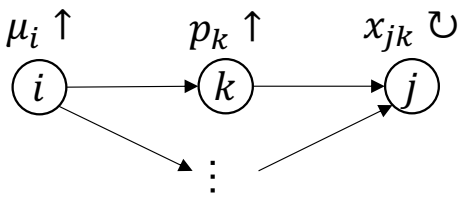


Figure 3: Illustration of Substitution

The remaining terms capture how relative price changes induced by industry i's markup change then cause substitution in other industries' production and final consumption and thus impact industry i's sales. The second term in eq. (28) relating to production elasticity θ_p deals with the substitution in production by each industry j, illustrated in Figure 3: when producer i raises its markup, it increases each price p_k in the economy by $\tilde{\Psi}_{ki}$. If inputs are complementary in production ($\theta_p < 1$), then producer j substitutes its expenditure towards producers k with higher exposure $\tilde{\Psi}_{ki}$ as their prices become relatively higher. If $\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(i)}, \Psi_{(i)}\right) > 0$, it means that the producers that are more exposed to i that j buys from also have high revenue shares that go to i, which is generically the case. As a result, the substitution of producer j leads to higher sales of producer i. The last term which relates 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. All terms in eq. (28) are expressed in terms of observed network positions at the equilibrium.

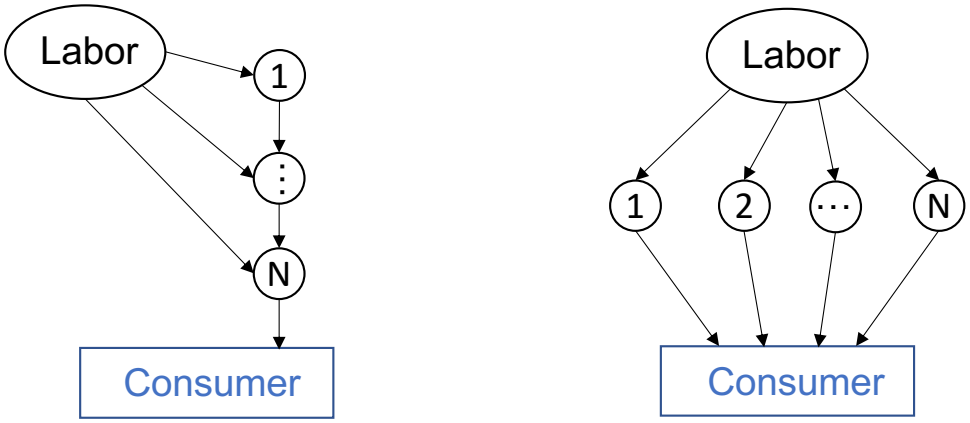
[25] We can revisit Example 2 with our general formula. Setting $\theta_p = \theta_c = 1$, we obtain

$$\eta_i = \Psi_{ii} + (\epsilon - 1)\,\tilde{\lambda}_i,$$

in which $\Psi_{ii} = \left(1 - \frac{\alpha}{\mu}\right)^{-1} = \frac{\mu}{\mu - \alpha}$, $\tilde{\lambda}_i = (1-\alpha)^{-1}$, exactly as the RHS of eq. (16).

Electronic copy available at: https://ssrn.com/abstract=4306329

3.3 Examples of Multi-Industry Economies



(a) Vertical economy (b) Horizontal economy

Figure 4: Oligopoly Economies

Example 3. (Vertical economy, Figure 4a) In this vertical economy, industry i's demand elasticity is

$$\eta_i = \theta_p \left(1 - \tilde{\lambda}_i\right) + \epsilon \tilde{\lambda}_i. \tag{29}$$

In this vertical economy, we have $\tilde{\lambda}_1 < \cdots < \tilde{\lambda}_N$ and thus $\eta_i < \cdots < \eta_N$ if $\theta_p < \epsilon$, since upstream industries have smaller impacts on the consumer price than downstream firms. We see that, if we have a chain with infinitely many industries, each will have a negligible impact on the consumer price and thus face a demand elasticity of θ_p.

Example 4. (Horizontal economy, Figure 4b) In this horizontal economy, industry i's demand elasticity is

$$\eta_i = \theta_c \left(1 - \tilde{\lambda}_i\right) + \epsilon \tilde{\lambda}_i. \tag{30}$$

In this horizontal economy, we obtain $\eta_i = \theta_c$ for any infinitesimal industry ($\tilde{\lambda}_i \to 0$), as in Atkeson and Burstein (2008).

Examples 3 and 4 show two extreme cases where the industry elasticity η_i is governed by θ_p in one case and by θ_c in another case. In general, both elasticities play a role depending on the network position each industry takes. The consumer demand elasticity ϵ is a force that applies universally across all industries.

Electronic copy available at: https://ssrn.com/abstract=4306329

4 Firm-Level Demand Elasticities and Markups

Now we allow for multiple firms in each industry, and see exactly how the elasticity η_i in Proposition 1, where there is a monopoly firm in each industry, becomes a measure of industry-level elasticity η_i here. Note that, no matter how many firms there are in each industry and in the whole economy, we are still applying the same equilibrium concept (competition in markups) to all of them. That is, when firm f_i in industry i decides on its markup μ_{f_i}, it assumes that all other firms in the economy hold their markups fixed, including firms in the same industry and in all other industries.[26]

Each firm $f_i \in \mathcal{F}_i$ maximizes its profit

$$\begin{aligned}\pi_{f_i} &= \left(1-\frac{1}{\mu_{f_i}}\right)\frac{p_{f_i}y_{f_i}}{P} = \left(1-\frac{1}{\mu_{f_i}}\right)\cdot\frac{p_{f_i}y_{f_i}}{p_iy_i}\cdot\frac{p_iy_i}{PC}\cdot C \\ &= \underbrace{\left(1-\frac{1}{\mu_{f_i}}\right)}_{\text{profit margin}}\cdot\underbrace{s_{f_i}}_{\text{within-industry market share}}\cdot\underbrace{\lambda_i}_{\text{industry Domar weight}}\cdot\underbrace{P^{1-\epsilon}}_{\text{consumer demand}}, \end{aligned} \tag{31}$$

in which $s_{f_i} := \frac{p_{f_i}y_{f_i}}{p_iy_i}$ is each firm's within-industry market share, and λ_i is the Domar weight of each industry as defined before. This expression nests eq. (17) in the special case where there is only one firm in each industry, i.e. $s_{f_i} \equiv 1$. For firms within the same industry, they differ in their profit margins and market shares, but share the same industry Domar weight as well as consumer demand.

Within industry. The price of industry bundle i is

$$p_i = \left(\sum_{f_i\in\mathcal{F}_i} p_{f_i}^{1-\rho}\right)^{\frac{1}{1-\rho}}, \tag{32}$$

in which p_{f_i} is firm f_i's price, and thus the market share can be expressed as

$$s_{f_i} = \frac{p_{f_i}^{1-\rho}}{p_i^{1-\rho}}. \tag{33}$$

[26] At a high level, "industry" can be viewed as an emergent concept that groups certain firms together in the nested CES production structure laid out in Section 2.3. It may not be possible, and it is not needed, to define what an industry is, given a generic production structure in Section 2.1.

Electronic copy available at: https://ssrn.com/abstract=4306329

As each firm in the same industry uses the same input bundle albeit differential cost $\frac{q_i}{z_{f_i}}$, the industry-level productivity z_i is defined as the output relative to the total input use. It is thus equal to an *output-quantity-weighted* harmonic mean of firm productivities, as in Edmond, Midrigan and Xu (2015, 2022),[27]

$$z_i := \frac{y_i}{m_i} = \left[\sum_{f_i \in \mathcal{F}_i} \frac{1}{z_{f_i}} \frac{y_{f_i}}{y_i}\right]^{-1}. \tag{34}$$

The industry-level markup is defined as the price p_i relative to the industry-level cost $\frac{q_i}{z_i}$, and is equal to an *input-quantity-weighted* arithmetic mean or a *sales-weighted* harmonic average of firm markups, as in Edmond, Midrigan and Xu (2015, 2022),

$$\mu_i := \frac{p_i}{\frac{q_i}{z_i}} = \sum_{f_i \in \mathcal{F}_i} \mu_{f_i} \frac{l_{f_i}}{l_i} = \left[\sum_{f_i \in \mathcal{F}_i} \frac{1}{\mu_{f_i}} s_{f_i}\right]^{-1}. \tag{35}$$

Across industries. Across industries, since only the industry bundle is used, what matters for allocation and prices in general equilibrium is $\frac{\mu_i}{z_i} = \frac{p_i}{mc_i}$. Thus the industry Domar weights λ_i and the consumer demand $C = P^{1-\epsilon}$ are determined by $\left\{\frac{\mu_i}{z_i}\right\}_{i \in \mathcal{I}}$. In maximizing its profit, firm f_i's markup μ_{f_i} affects λ_i and P only through its effect on $\frac{\mu_i}{z_i}$. All results we derived in Section 3 extend here, if we replace μ_i by $\frac{\mu_i}{z_i}$.

Average markup. The economy-wide average markup is defined as the *sales-weighted* harmonic average of firm markups, or the *sales-weighted* harmonic average of industry markups, too, as in Edmond, Midrigan and Xu (2015, 2022); Baqaee and Farhi (2020),

$$\mu := \left[\sum_{f_i \in \mathcal{F}_i, i \in \mathcal{I}} \frac{1}{\mu_{f_i}} \frac{s_{f_i} \lambda_i}{\sum_{i \in \mathcal{I}} \lambda_i}\right]^{-1} = \left[\sum_{i \in \mathcal{I}} \frac{1}{\mu_i} \frac{\lambda_i}{\sum_{i \in \mathcal{I}} \lambda_i}\right]^{-1}. \tag{36}$$

4.1 Firm-Level Outcomes

Proposition 2. (Firm markup) *For firm f_i in industry i, its Lerner condition is*

$$\frac{\mu_{f_i}}{\mu_{f_i} - 1} = \left(1 - s_{f_i}\right) \rho + s_{f_i} \eta_i, \tag{37}$$

[27] Note that since y_i is a CES bundle of y_{f_i}, the weights $\frac{y_{f_i}}{y_i}$ do not necessarily sum up to one.

Electronic copy available at: https://ssrn.com/abstract=4306329

with industry-level elasticity as, characterized in Proposition 1,

$$\eta_i := cycle_i + \left(\theta_p - 1\right) SP_i + \left(\theta_c - 1\right) SC_i + \left(\epsilon - 1\right) CP_i + 1. \tag{38}$$

This proposition suggests that the demand elasticity each firm faces is an average of the industry-level elasticity η_i and the within-industry elasticity of substitution ρ, weighted by its within-industry market share s_{f_i}. This is because as each firm raises its markup, its market share shrinks (ρ-term) and the industry sale changes (η_i-term). Going forward, we maintain the assumption that $\rho > \eta_i$ and $\left(1 - s_{f_i}\right)\rho + s_{f_i}\eta_i > 1$, in which case larger firms within each industry face smaller demand elasticities. This formula nests the formula under Bertrand competition in Atkeson and Burstein (2008) as a special case of uniform η_i across industries. When η_i is variable, assuming $\theta_p, \theta_c < 1, \epsilon > 1$, an industry that is in a weaker cycle, induces more substitution in production and consumption, and has a smaller impact on the consumer price, will have a lower η_i. This predicted variation across industries is absent in Grassi (2017), whose model assumes that firms compete without internalizing downstream firms' responses. As these network statistics are expressed in observed network positions, we will test this prediction by exploiting the variation of markup-to-market-share relations across industries.

Edmond, Midrigan and Xu (2022) study the welfare cost of markups in environments that are comparable to Atkeson and Burstein (2008), and show that the efficient allocation can be implemented by a nonlinear schedule of size-dependent subsidies. In a generic networked economy, Proposition 2 describes a set of industry-specific size-related distortions. Consequently, it takes a set of industry-specific size-dependent subsidies to restore efficiency. That basically amounts to firm-specific subsidies, suggesting that a networked economy with endogenous markups can be much more challenging for policy making.

4.2 Concentration and Distortions

Before we test Proposition 2, here we take a detour to illustrate how observed concentration is determined by both firm productivities and industry demand elasticities, and consequently how looking at concentration alone can be misleading. While the previous results are equilibrium relations that do not hinge on primitives, the predictions I present here link observed equilibrium outcomes to primitives.

Lemma 1. (Sorting within each industry) *For any two firms f_i, g_i in the same industry i, if f_i*

Electronic copy available at: https://ssrn.com/abstract=4306329

is more productive than g_i, i.e. $z_{f_i} > z_{g_i}$, then at the equilibrium we must have $\mu_{f_i} > \mu_{g_i}, p_{f_i} < p_{g_i}, s_{f_i} > s_{g_i}$. That is, within each industry, a more productive firm charges a higher markup and a lower price, and takes a larger market share.

Lemma 1 suggests that within an industry, the more productive firms take larger market shares and charge higher markups, setting the stage for cross-industry comparison.

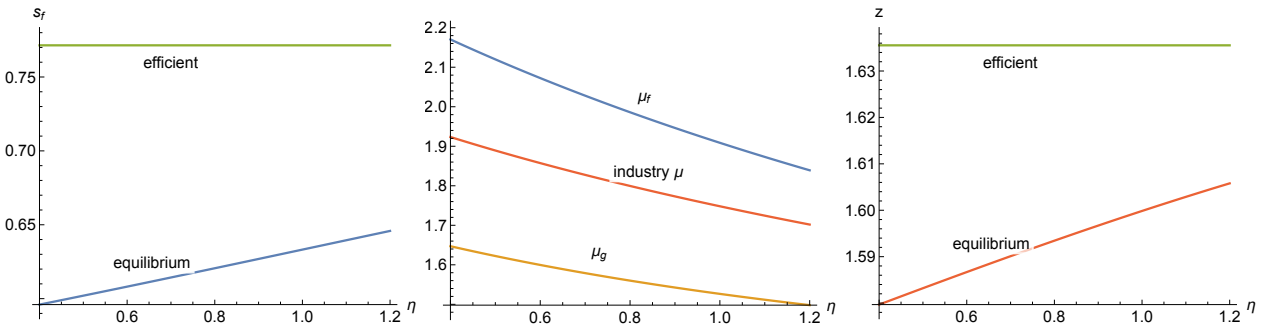
Proposition 3. (Concentration and distortions across industries) *Suppose that two industries i, j each have two firms f_i, g_i and f_j, g_j whose productivities satisfy $z_{f_i} = z_{f_j} > z_{g_i} = z_{g_j}$. For each industry, define $\kappa_i := \frac{\log \mu_{f_i} - \log \mu_{g_i}}{\log s_{f_i} - \log s_{g_i}}$ as the relation of markup to (within-industry) market share. If industry i has a lower industry-level elasticity $\eta_i < \eta_j$ at the equilibrium, we will have:*

1. $\mu_{f_i} > \mu_{g_i}, \mu_{f_j} > \mu_{g_j}, \frac{\mu_{f_i}}{\mu_{g_i}} > \frac{\mu_{f_j}}{\mu_{g_j}}$: in the less elastic industry firms charge higher markups and the more productive firm charges a particularly high markup relative to the other firm;

2. $1/2 < s_{f_i} < s_{f_j}, \kappa_i > \kappa_j$: the less elastic industry is less *concentrated, and sees a steeper relation of markup to market share;*

3. $z_i < z_j, \mu_i > \mu_j$: the less elastic industry is less *productive and* more *distorted.*

Proposition 3 presents a series of comparison across industries each with two firms. Figure 5 plots the market allocation, markups, and the industry's productivity, as a function of the industry-level demand elasticity η.



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

Figure 5: Comparative Statics w.r.t. Industry Demand Elasticity η_i

Notes: Panel (a) shows the equilibrium market share of the more productive firm (firm f), as well as its hypothetical market share assuming that the two firms charge the same markups, as a function of η. Panel (b) plots two firms' markups as well as the industry-level markup at the equilibrium. Panel (c) displays the equilibrium industry-level productivity and the hypothetical industry-level productivity assuming that the two firms charge the same markups.

Electronic copy available at: https://ssrn.com/abstract=4306329

Given Proposition 2, 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* markup 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_i is an output-quantity-weighted harmonic mean of firm productivities, as eq. (34). 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-level markup is a sales-weighted harmonic average of firm markets, as eq. (35). Fixing the market shares, the industry markup is higher if its demand elasticity is lower, since firms will charge higher markups. That continues to be true at equilibrium, even with the more productive firm occupying a relatively smaller market share in a less elastic industry.

At a high level, it is intuitive that firms in less elastic industries will have higher markups, leading to a higher industry markup. In contrast, it can be puzzling at first glance that a less elastic industry has a lower concentration, coexisting with a higher markup and a lower productivity. Our usual intuition seems to be 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 lower distortion. However, concentration is itself an equilibrium outcome—it is a symptom instead of a root cause of imperfect competition. The naive argument 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.[28]

[28]This mechanism is reminiscent of the negative relation between concentration and markup from Melitz

Electronic copy available at: https://ssrn.com/abstract=4306329

5 Empirical Analysis

I present empirical evidence based on Compustat firms in the US input-output network from 1997 to 2019, which confirms the theory prediction of Propositions 1 and 2 that different industries have different slopes of markups to market shares. The empirical test uses two approaches for robustness: a production-function approach and a cost-share approach, both of which exploit the variation of markups across firms.

5.1 Data and Empirical Strategy

I use the BEA input-output table at summary level with 71 industries, which roughly corresponds to 2-digit SIC codes (major groups), from 1997 to 2019 to construct the industry-level input-output matrices and Leontief inverses, $\tilde{\Omega}, \tilde{\Psi}, \Omega, \Psi$. Based on these matrices, I construct the four network statistics that determine the industry-level demand elasticities η_i, following Proposition 1.

I use Compustat data to calculate proxies of firms' markups and within-industry market shares. I follow standard procedures to clean Compustat data. Details are in Appendix C. With these measures, I test Proposition 2, which suggests that industries with lower η_i should see steeper markup-to-market-share relations.

In terms of markups, it is by now well known thanks to Hall (1986, 1988) that markups can be recovered from the first-order condition of cost-minimizing firms for a variable input X as

$$\mu_{f_i,t} = \frac{p_{f_i,t} y_{f_i,t}}{p_{X,t} X_{f_i,t}} \zeta^X_{f_i,t}, \tag{39}$$

in which $p_{X,t} X_{f_i,t}$ is the input cost, $p_{f_i t} y_{f_i,t}$ is the revenue, and $\zeta^X_{f_i,t} := \frac{\partial \log y_{f_i,t}}{\partial \log X_{f_i,t}}$ is the output elasticity of input X. Estimating the output elasticity $\zeta^X_{f_i,t}$ yields the markup. The two approaches I adopt diverge hereafter.

Production-function approach. The difficulty in estimating $\zeta^X_{f_i,t}$ lies in the fact that we rarely have firm-level quantity data. Instead, previous papers have used revenue data. Bond et al. (2021) argue that the level of markups estimated from revenue data is unin-

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-level elasticity η_i instead of the within-industry elasticity of substitution ρ.

Electronic copy available at: https://ssrn.com/abstract=4306329

formative of the true level. De Ridder, Grassi and Morzenti (2022) refine this argument and show that in the presence of firm heterogeneity, the estimated firm-level markup is a product of the true firm-level markup and a constant such that the average markup across firms is indeed not informative as Bond et al. (2021) suggest. However, the log variation across firms does correspond to the true log variation. As should be clear from previous discussions, my theory predictions can be tested based on the variation of markups across firms. Hence, the production-function approach is valid here. Following De Loecker, Eeckhout and Unger (2020); De Ridder, Grassi and Morzenti (2022), I estimate a translog production function with cost of goods sold (cogs) as the variable input, for each 2-digit NAICS industry on 5-year rolling windows.

Cost-share approach. Alternatively, if we assume that the output elasticity $\zeta^X_{f_i,X}$ is common to all firms in the same industry in the same year, the variation in log markups within each industry in each year can be obtained directly from the log of revenue over variable cost $\frac{p_{f_i,t}y_{f_i,t}}{p_{X,t}X_{f_i,t}}$.[29] This is the cost-share approach in Bond et al. (2021); Edmond, Midrigan and Xu (2022). In the panel regression, when we include a set of industry-year fixed effects, log sales over cost of goods sold is a valid proxy for log markup.

Theory predictions. To adapt the theory predictions to our empirical test, acknowledging various limits regarding estimation and model misspecifications which we explain shortly, we present a corollary following from Proposition 2.

Corollary 1. (Log-log markup-to-market-share relation) *Under market shares* $\left\{s_{f_i} = \bar{s}\exp\left(\Delta\log s_{f_i}\right)\right\}_{f_i\in\mathcal{N}}$ *and markups* $\left\{\mu_{f_i} = \bar{\mu}\exp\left(\Delta\log\mu_{f_i}\right)\right\}_{f_i\in\mathcal{N}}$, *the relation of log markup to log market share to the first order in* $\left\{\Delta\log s_{f_i}, \Delta\log\mu_{f_i}\right\}_{f_i\in\mathcal{N}}$ *in each industry is*

$$\Delta\log\mu_{f_i} \approx \frac{\bar{s}\left(\bar{\mu}-1\right)^2}{\bar{\mu}}\left(\rho-\eta_i\right)\Delta\log s_{f_i}. \tag{40}$$

That is, the relation of log markup to log market share is steeper for an industry with a lower η_i.

[29] This assumption basically amounts to assuming a Cobb-Douglas production function. De Ridder, Grassi and Morzenti (2022) discuss its limitations.

Electronic copy available at: https://ssrn.com/abstract=4306329

We run the following panel regression that parallels Corollary 1,

$$\log \mu_{f_i,t} = \beta \log s_{f_i,t} + \log s_{f_i,t} \times \left(\gamma_{\text{cycle}} \text{cycle}_{i,t} + \gamma_{\text{SP}} \text{SP}_{i,t} + \gamma_{\text{SC}} \text{SC}_{i,t} + \gamma_{\text{CP}} \text{CP}_{i,t}\right) + \delta_{i,t} + \delta_{f_i} + e_{f_i,t}. \tag{41}$$

In addition to the industry-year fixed effects $\delta_{i,t}$, we control for firm fixed effects δ_{f_i} to purge persistent firm-level heterogeneity. Standard errors will be clustered at the firm level. Our interest is in the interaction coefficients γ's, which tell us if the markup-to-market-share relation is flatter in some industries than others. In this specification, we use log markup since only the log variation in markups is correctly identified, and we use log market share for better inference since its distribution is closer to a normal distribution than the distribution of market share. Since this is a nonlinear relation, on the RHS I use cross-sectionally demeaned networks statistics. For example, the cycle term is $\text{cycle}_{it} - \overline{\text{cycle}}_t$. Thus the coefficient related to the stand-alone log share term and all interactions can be interpreted as local slopes of an "average" industry in year t.

This regression is comparable to the regression in Edmond, Midrigan and Xu (2022) with two differences: (i) On the LHS, they use a nonlinear increasing function of μ_{f_i} which depends on the level of μ_{f_i}, in contrast to $\log \mu_{f_i}$. (ii) They assume that the coefficient of log market share on the RHS is the same across industries for their analysis.

Proposition 1 and Corollary 1 predict that industries in stronger cycles have higher demand elasticities and flatter markup-to-market share relations, that is,

$$\gamma_{\text{cycle}} < 0. \tag{42}$$

Further, in terms of the relative magnitudes of coefficients compared to γ_{cycle}, we have

$$\frac{\gamma_{\text{SP}}}{\gamma_{\text{cycle}}} = \theta_p - 1, \ \frac{\gamma_{\text{SC}}}{\gamma_{\text{cycle}}} = \theta_c - 1, \ \frac{\gamma_{\text{CP}}}{\gamma_{\text{cycle}}} = \epsilon - 1. \tag{43}$$

If goods are complementary in production and consumption ($\theta_p, \theta_c < 1$), we will see $\gamma_{\text{SP}}, \gamma_{\text{SC}} < \left|\gamma_{\text{cycle}}\right|$.[30]

[30] We do not interpret the relative magnitudes of γ's to β, as it is sensitive to model specifications. Proposition 4 in Appendix B provides an ad hoc example where firms compete à la Cournot in each industry with variable elasticity η_i. In that example, the signs and relative magnitudes of terms underlying η_i are the same as here, but the relative magnitudes of γ's to β is different.

Electronic copy available at: https://ssrn.com/abstract=4306329

Summary statistics. The four network statistics in 2019 are exhibited as a heat map in Figure 6. The order from top to bottom follows BEA industry codes and roughly reflects the upstreamness of each industry. It is clear that the substitution in production (SP) term is much more important for upstream industries than for downstream industries. The opposite is true for the substitution in consumption (SC) term. The cycle term and the impact on consumer price (CP) terms are relatively smaller, but will provide enough variation to identify their coefficients.

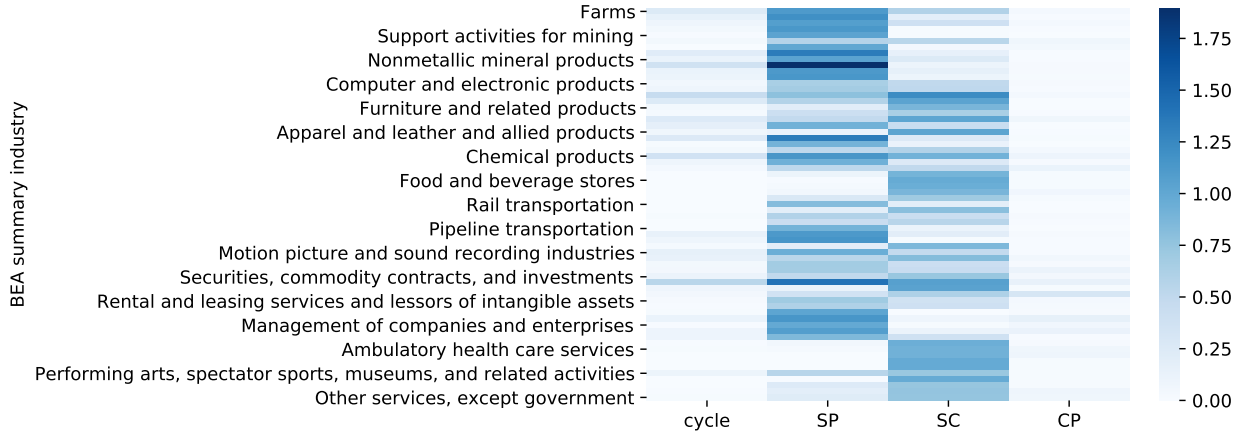


Figure 6: Industry-Level Network Statistics in 2019

Notes: The four network statistics are defined in Proposition 1. Only selected industry names are presented due to space constraints, but all industries are included in the heat map. Some finance-related industries feature prominently according to these network statistics, but financial firms will be excluded from our regressions as is standard in the literature, since their balance sheets are different. Table 4 in Appendix A lists the top industries by each network statistic in 2019.

Table 1 presents summary statistics of firms' log markups estimated from the production-function approach and log sales across all firm-year observations, along with the network statistics which are the same for all firms in the same industry in a given year.

Electronic copy available at: https://ssrn.com/abstract=4306329

Table 1: Summary Statistics

	mean	sd	min	p25	p50	p75	p90	max
log markup (PF)	0.162	0.466	-5.720	0.039	0.255	0.410	0.562	2.090
log sale	5.737	1.657	1.800	4.457	5.765	7.007	7.963	9.543
cycle	0.153	0.141	0.000	0.040	0.108	0.263	0.398	0.512
SP	0.789	0.425	0.000	0.488	0.756	1.127	1.246	2.256
SC	0.584	0.326	0.000	0.322	0.644	0.876	0.970	1.274
CP	0.054	0.053	0.002	0.018	0.034	0.084	0.116	0.315
Observations	55890							

Notes: Sales in millions (deflated in 2012 dollars).

5.2 Markup-to-Market-Share Relations across Industries

Table 2 presents the results of regression (41) under the production-function approach. (Table 5 in Appendix A presents similar results under the cost-share approach.) Across all specifications, we control for firm fixed effects, and industry-year fixed effects and cluster standard errors at the firm level. Columns (1-3) and Columns (4-6) use SIC 2-digit industry classification and BEA summary-level industry classification, respectively.

Columns (1) and (4) confirm the conventional wisdom that within an industry, markups increase in market shares. This can be predicted by models of Atkeson and Burstein (2008); Grassi (2017). Columns (2) and (5) add in our novel network statistics interacted with log market shares. All statistics are highly significant.[31] As predicted by the theory, the cycle interaction carries a negative sign. Further, the ratio of other interaction coefficients relative to the cycle interaction reflect $\theta_p, \theta_c, \epsilon$, which we will use for calibration.

Drawing from the interaction coefficients in Column (2) we can compile a composite measure of industry-level demand elasticity

$$\eta_{it} = \text{cycle}_{it} - 0.48 \times \text{SP}_{it} - 0.78 \times \text{SC}_{it} + 0.52 \times \text{CP}_{it} + 1. \tag{44}$$

Columns (3) and (6) report the results of including η_{it} interacted with log share, and find coefficients that are indistinguishable from the coefficients of the cycle interaction in Columns (2) and (5).

[31] The CP term has a smaller standard error than is displayed here if we do not exclude housing-related firms which have larger impacts on the consumer price, and in the cost-share approach.

Electronic copy available at: https://ssrn.com/abstract=4306329

I do not report the overall R^2 for these regressions, as those numbers are mostly driven by the fixed effects we include. Instead, I show the within R^2 as the measure of predictive power. Comparing Column (2) against (1) suggests that including these interaction terms increases the within R^2 by about three times. (Under the cost-share approach, taking these interactions into account raises the within R^2 by even more; see Table 5.) This attests to the quantitative importance of heterogeneity across industries.

Table 2: Markup-to-Market-Share Relations (Production-Function Approach)

	(1)	(2)	(3)	(4)	(5)	(6)
log s	0.0306***	0.0295***	0.0295***	0.0310***	0.0285***	0.0299***
	(0.00521)	(0.00514)	(0.00505)	(0.00505)	(0.00492)	(0.00483)
log s × cycle		-0.180***			-0.250***	
		(0.0361)			(0.0726)	
log s × SP		0.0863***			0.137***	
		(0.0164)			(0.0299)	
log s × SC		0.141***			0.183***	
		(0.0184)			(0.0334)	
log s × CP		-0.0929*			-0.0993	
		(0.0512)			(0.0640)	
log s × eta			-0.180***			-0.237***
			(0.0214)			(0.0356)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	sic2 × year	sic2 × year	sic2 × year	BEA × year	BEA × year	BEA × year
Observations	55890	55890	55890	55965	55965	55965
Within R^2	.0032	.0083	.0083	.0036	.0087	.0084

Standard errors in parentheses
Standard errors clustered at firm level
* p<0.10, ** p<0.05, *** p<0.01

Calibration. Taking estimates from Column (2) in Table 2 and using expressions (43), I calibrate elasticities $\theta_p, \theta_c, \epsilon$, shown in Table 3. The elasticity θ_p is very tightly estimated and is in line with previous literature that suggests θ_p is much less than 1. Atalay (2017) finds that the elasticity of substitution among different intermediate inputs is much below 1 (some of his estimates indicate a negative elasticity), and the elasticity of substitution between value added and the intermediate input bundle is around 1. My model has only

Electronic copy available at: https://ssrn.com/abstract=4306329

one parameter governing the substitution among value added and different intermediate inputs, which is plausibly between these two ends. While the previous literature suggests that θ_c is below but close to one, my estimate appears lower. The elasticity ϵ, along with the quasi-linear utility specification eq. (8), is a convenient assumption to give rise to an isoelastic demand for the consumption good. While it may not correspond to a parameter in reality, the number appears reasonable. Overall, the model does well in predicting markups across industries and delivers sensible calibrated parameters.

Table 3: Calibration

Elasticity	$\theta_p = 1 + \frac{\gamma_{SP}}{\gamma_{cycle}}$	$\theta_c = 1 + \frac{\gamma_{SC}}{\gamma_{cycle}}$	$\epsilon = 1 + \frac{\gamma_{CP}}{\gamma_{cycle}}$
Estimate	0.52	0.20	1.52
Standard error	(0.03)	(0.11)	(0.29)

Notes: $\theta_p, \theta_c, \epsilon$ are the elasticity of substitution of CES production, the elasticity of substitution of CES final consumption, and the elasticity of consumer demand, respectively, defined in Section 2.3. To estimate them, I draw a million realizations from the joint distribution of estimated coefficients from Column (2) in Table 2 and calculate estimates and standard errors of $\theta_p, \theta_c, \epsilon$ according to expressions (43).

6 TFP Loss from Misallocation

Last, I tentatively quantify the importance of heterogeneity across industries for total factor productivity (TFP) loss due to misallocation. In developing this application, I use the theory-predicted markups from Column (2) of Table 2, leaving out fixed effects.[32] These are the markups that are related to market size and imperfect competition in the theory, instead of from any source. Focusing on this understandable part of markups follows Edmond, Midrigan and Xu (2022). However, taking into account the heterogeneity across industries explains a lot more variations than using only market shares, as suggested in our empirics. Figure 8 in Appendix A presents the histograms of predicted log markup, with or without considering heterogeneity across industries (using Columns (2) and (1) in Table 2 respectively), the former of which displays much larger variations.

I use the formula of change in TFP from eliminating markups ("distance to frontier") developed in Baqaee and Farhi (2020), and decompose misallocation across and within industries. To use their formula, in addition to the N firms across I industries that produce with elasticity θ_p and charge markups μ_i as laid out in Section 2.3, I introduce

[32] Alternatively, one can construct firm markups using the exact formula in Proposition 2, but that is more prone to model misspecifications than simply using empirical predictions.

Electronic copy available at: https://ssrn.com/abstract=4306329

I fictitious "industry packers" that produce industry composites with elasticity ρ and charge fixed markups of one. Still, postulate a packer 0 that produces the consumption good from industry composites with elasticity θ_c and charges a markup of one. I include the labor factor which does not use any input to produce. This gives us a $(N + I + 2)$-firm representation of the nested CES network, with each firm operating a CES technology indexed by $f_i \in \check{\mathcal{N}}$. Based on the industry-level input-output matrices $\tilde{\Omega}, \Omega, \tilde{\Psi}, \Psi$, we can construct the firm-level input-output matrices $\check{\tilde{\Omega}}, \check{\Omega}, \check{\tilde{\Psi}}, \check{\Psi}$ and each firm's Domar weight λ_{f_i}. (See Appendix D for details.) I set $\theta_p = 0.52, \theta_c = 0.20$ following my calibration, and $\rho = 4$ for a conservative estimate.

Lemma 2. (Distance to frontier, Baqaee and Farhi, 2020) *Under markups* $\left\{\mu_{f_i} = \exp\left(\Delta \log \mu_{f_i}\right)\right\}_{f_i \in \check{\mathcal{N}}}$, *the distance to frontier up to the second order in* $\left\{\Delta \log \mu_{f_i}\right\}_{f_i \in \check{\mathcal{N}}}$ *is*

$$\mathcal{L}\left(\left\{\mu_{f_i}\right\}\right) \approx \frac{1}{2} \sum_{f_i \in \check{\mathcal{N}}} \lambda_{f_i} \theta_{f_i} \operatorname{Var}_{\check{\Omega}(f_i)} \left(\sum_{g_j \in \check{\mathcal{N}}} \check{\Psi}_{(g_j)} \Delta \log \mu_{g_j} \right). \tag{45}$$

Definition 3. (Decomposition of TFP loss) The distance to frontier can be decomposed into misallocation within and across industries and loss from aggregate markup,

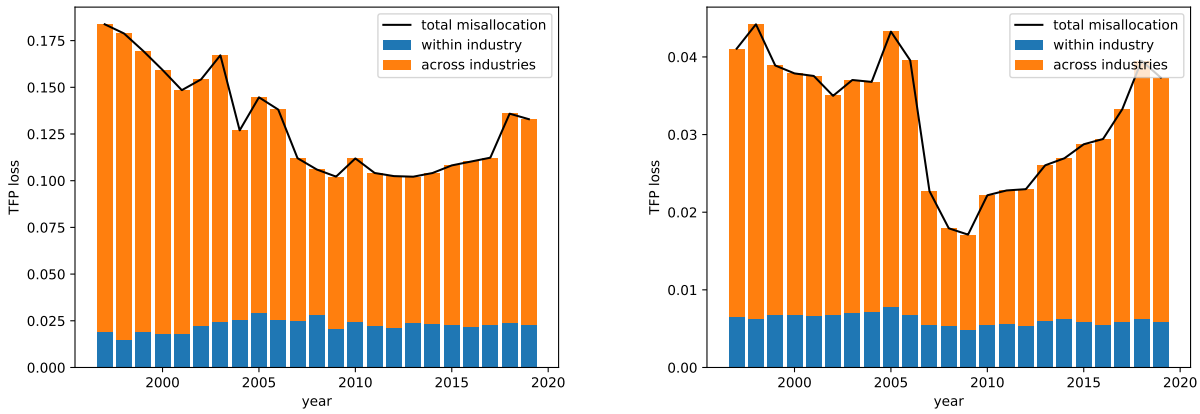
$$\begin{aligned} \mathcal{L}_{\text{within}} &:= \mathcal{L}\left(\left\{\mu_{f_i}\right\}\right) - \mathcal{L}\left(\left\{\mu_{f_i} = \mu_i\right\}\right), \\ \mathcal{L}_{\text{across}} &:= \mathcal{L}\left(\left\{\mu_{f_i} = \mu_i\right\}\right) - \mathcal{L}\left(\left\{\mu_{f_i} = \mu\right\}\right), \\ \mathcal{L}_{\text{aggregate}} &:= \mathcal{L}\left(\left\{\mu_{f_i} = \mu\right\}\right). \end{aligned} \tag{46}$$

The within-industry misallocation term quantifies the gain from equalizing markups within each industry, while the across-industry term is the gain from subsequently equalizing markups across industries. It is evident that both numbers derive from the dispersion of markups, invariant to level shifts. In contrast, the aggregate markup component hinges on the level of markups.[33] Here I focus on quantifying the misallocation terms. Figure 7a shows the TFP loss when taking into account the sectoral heterogeneity in demand elasticities, and Figure 7b shows the implied number that ignores sectoral heterogeneity. In the case with sectoral heterogeneity, the average TFP loss during the sample period is about 13%, which is four times larger than the other case. In both cases, the misallocation

[33] Edmond, Midrigan and Xu (2022) provide an extensive discussion on the sensitivity of this number to the average level of markups.

Electronic copy available at: https://ssrn.com/abstract=4306329

across industries is the major component.



(a) Considering heterogeneity across industries

(b) Using only market shares

Figure 7: TFP Loss due to Misallocation under Theory-Predicted Markups

Notes: Panels (a) and (b) use the predicted log markups from Columns (2) and (1) respectively in Table 2, excluding fixed effects

Comparing to numbers in the literature, Baqaee and Farhi (2020) show that the total gain in TFP from eliminating markups is about 10% to 25%, while Edmond, Midrigan and Xu (2022) suggest instead a number of just around 2% to 6%. Both of their numbers include the TFP loss due to aggregate markup, in addition to misallocation, which I focus on. In my exercise, taking into account the sectoral heterogeneity predicts a TFP loss due to misallocation of about 13%, compared to a number of about 3% if one ignores sectoral heterogeneity. There are two key differences in these papers: (i) Baqaee and Farhi (2020) and this paper build static models based on the true US input-output structure, while Edmond, Midrigan and Xu (2022) consider a simpler economy with one intermediate good in a dynamic setting. Baqaee and Farhi (2020)'s calculation specifies a more complicated nested CES structure than I do. (ii) Perhaps more importantly, Baqaee and Farhi (2020) feed in the estimated markups that have more dispersion than the size-dependent markups simulated in Edmond, Midrigan and Xu (2022) under homogeneous industry-level demand elasticities. In the presence of difference in specifications across these papers and admittedly large uncertainty behind all these estimates (due to model misspecifications and data imputations), these numbers have to be interpreted with caution. Nonetheless, the takeaway from my exercise is that even if one restricts attention to the part of markup

Electronic copy available at: https://ssrn.com/abstract=4306329

variations related to market power, taking into account the sectoral heterogeneity explains a lot more variations and suggests a larger implied TFP loss, which may help fill in the gap between Baqaee and Farhi (2020) and Edmond, Midrigan and Xu (2022).

7 Conclusion

This paper proposes a novel and tractable modeling approach to analyzing demand elasticities and distortions across industries in a generic networked economy. I show that each firm's markup depends on five sufficient statistics—its within-industry market share, along with four industry-level network statistics. I test this prediction using a production-function approach and a cost-share approach, both of which identify the relation of markup to market share in each industry, and confirm that the relation is steeper in some industries than in others as the theory suggests. Taking into account the sectoral heterogeneity explains about three times more variation in markups across firms in Compustat from 1997 to 2019. Further, I present novel theory implications on interpreting concentration which is driven by both firm productivity advantages and industry-level demand elasticities. In an application, I show that under the theory-implied markups, there is an average loss in TFP of 13% due to misallocation over the sample period. This number is four times larger than what one would get when ignoring sectoral heterogeneity.

The Lerner's inverse elasticity rule has inspired decades of theoretical and empirical research in the field of industrial organization, most often of specific industries. The modeling approach developed in this paper is tractable for analyzing the overall organization of production in the economy, where a broad set of questions remain mysterious: How do firms and sectors endogenously respond to shocks and contribute to economic fluctuations? How do we evaluate and design industrial policy and competition policy, acknowledging the importance of sectoral linkages and heterogeneity? What are the bottleneck industries in global trade networks and for economic growth? I hope that this modeling approach can be helpful for answering these questions in the future.

Electronic copy available at: https://ssrn.com/abstract=4306329

Bibliography

Acemoglu, Daron, and Alireza Tahbaz-Salehi. 2020. "Firms, Failures, and Fluctuations: The Macroeconomics of Supply Chain Disruptions." *NBER Working Paper*. 5

Acemoglu, Daron, and Pablo D. Azar. 2020. "Endogenous Production Networks." *Econometrica*, 88(1): 33–82. 6

Acemoglu, Daron, Vasco M. Carvalho, Asuman E. Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica*, 80(5): 1977–2016. 5

Afrouzi, Hassan, Andres Drenik, and Ryan Kim. 2021. "Growing by the Masses - Revisiting the Link between Firm Size and Market Power." *SSRN Electronic Journal*. 6

Alviarez, Vanessa, Michele Fioretti, Ayumu Ken Kikkawa, and Monica Morlacco. 2022. "Two-Sided Market Power in Firm-to-Firm Trade." *Working Paper*, 1–47. 5

Atalay, Enghin. 2017. "How Important Are Sectoral Shocks?" *American Economic Journal: Macroeconomics*, 9(4): 254–280. 5, 31

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review*, 98(5): 1998–2031. 2, 4, 20, 23, 30, 53

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. 6

Azar, José, and Xavier Vives. 2021. "General Equilibrium Oligopoly and Ownership Structure." *Econometrica*, 89(3): 999–1048. 7, 8

Bain, Joe S. 1951. "Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936-1940." *The Quarterly Journal of Economics*, 65(3): 293. 6

Bain, Joe S. 1956. *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries.* Harvard University Press. 6

Baqaee, David Rezza, and Emmanuel Farhi. 2019. "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem." *Econometrica*, 87(4): 1155–1203. 5

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *The Quarterly Journal of Economics*, 135(1): 105–163. 4, 5, 15, 16, 22, 32, 33, 34, 35, 44, 54

Basu, Susanto. 2019. "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence." *Journal of Economic Perspectives*, 33(3): 3–22. 6

Electronic copy available at: https://ssrn.com/abstract=4306329

Benassy, Jean-Pascal. 1988. "The Objective Demand Curve in General Equilibrium with Price Makers." *The Economic Journal*, 98(390): 37. 7

Benkard, C. Lanier, Ali Yurukoglu, and Anthony Lee Zhang. 2021. "Concentration in Product Markets." *NBER Working Paper*, , (February). 6

Berry, Steven, Martin Gaynor, and Fiona M. Scott Morton. 2019. "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization." *Journal of Economic Perspectives*, 33(3): 44–68. 3

Bigio, Saki, and Jennifer La'O. 2020. "Distortions in Production Networks." *The Quarterly Journal of Economics*, 135(4): 2187–2253. 5

Bizzarri, Matteo. 2022. "Supply and demand function competition in general equilibrium: endogenous market power in input-output networks." *Working Paper*. 5

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data." *Journal of Monetary Economics*, 121: 1–14. 3, 6, 26, 27

Bowley, Arthur L. 1924. *The Mathematical Groundwork of Economics.* Oxford University Press. 7

Bresnahan, Timothy F. 1989. "Chapter 17 Empirical studies of industries with market power." In *Handbook of Industrial Organization*. Vol. 2, 1011–1057. 6

Chaney, Thomas. 2018. "The Gravity Equation in International Trade: An Explanation." *Journal of Political Economy*, 126(1): 150–177. 5

De Loecker, Jan. 2021. "Comment on (Un)pleasant ... by Bond et al (2020)." *Journal of Monetary Economics*, 121: 15–18. 6

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. 3, 6, 27

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti. 2022. "The Hitchhiker's Guide to Markup Estimation." *Working Paper*. 3, 6, 27

Dhyne, Emmanuel, Ayumu Ken Kikkawa, and Glenn Magerman. 2022. "Imperfect Competition in Firm-to-Firm Trade." *Journal of the European Economic Association*, 20(5): 1933–1970. 4

Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes. 2018. "Just starting out: Learning and equilibrium in a new market." *American Economic Review*, 108(3): 565–615. 12

Electronic copy available at: https://ssrn.com/abstract=4306329

Ederer, Florian, and Bruno Pellegrino. 2022. "A Tale of Two Networks: Common Ownership and Product Market Rivalry." *NBER Working Paper*. 7

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015. "Competition, Markups, and the Gains from International Trade." *American Economic Review*, 105(10): 3183–3221. 22

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2022. "How Costly Are Markups?" *Journal of Political Economy*, , (Accepted). 3, 4, 5, 22, 23, 27, 28, 32, 33, 34, 35

Feenstra, Robert C. 2003. "A homothetic utility function for monopolistic competition models, without constant price elasticity." *Economics Letters*, 78(1): 79–86. 6

FitzRoy, Felix R. 1974. "Monopolistic equilibrium, non-convexity and inverse demand." *Journal of Economic Theory*, 7(1): 1–16. 7

Gabaix, Xavier. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica*, 79(3): 733–772. 4

Gabszewicz, Jean Jaskold, and Jean-Philippe Vial. 1972. "Oligopoly "A la Cournot" in a General Equilibrium Analysis." *Journal of Economic Theory*, 4(3): 381–400. 7, 8

Grant, Simon, and John Quiggin. 1994. "Nash equilibrium with mark-up-pricing oligopolists." *Economics Letters*, 45(2): 245–251. 11

Grassi, Basile. 2017. "IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important." *Working Paper*. 4, 23, 30

Gutiérrez, Germán, and Thomas Philippon. 2017. "Investmentless Growth: An Empirical Investigation." *Brookings Papers on Economic Activity*, 2017(2): 89–190. 6

Hall, R. L., and C. J. Hitch. 1939. "Price theory and business behaviour." *Oxford Economic Papers*, os-2(1): 12–45. 2

Hall, Robert E. 1986. "Market Structure and Macroeconomic Fluctuations." *Brookings Papers on Economic Activity*, 1986(2): 285–338. 6, 26

Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy*, 96(5): 921–947. 6, 26

Hart, Oliver D. 1985. "Imperfect Competition in General Equilibrium: An Overview of Recent Work." In *Frontiers of Economics*. , ed. Kenneth Arrow and Seppo Honkapohja, 100–149. Basil Blackwell, Oxford. 7, 8, 11

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403–1448. 5

Irmen, Andreas. 1997. "Mark-up pricing and bilateral monopoly." *Economics Letters*, 54(2): 179–184. 11

Electronic copy available at: https://ssrn.com/abstract=4306329

Jones, Charles I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." *American Economic Journal: Macroeconomics*, 3(2): 1–28. 5

Jones, Charles I. 2013. "Misallocation, Economic Growth, and Input-Output Economics." In *Advances in Economics and Econometrics*. , ed. Daron Acemoglu, Manuel Arellano and Eddie Dekel, 419–456. Cambridge:Cambridge University Press. 5

Karabarbounis, Loukas, and Brent Neiman. 2014. "The Global Decline of the Labor Share." *The Quarterly Journal of Economics*, 129(1): 61–103. 6

Kehrig, Matthias, and Nicolas Vincent. 2021. "The Micro-Level Anatomy of the Labor Share Decline." *The Quarterly Journal of Economics*, 136(2): 1031–1087. 6

Kimball, Miles S. 1995. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit and Banking*, 27(4): 1241. 6

Klenow, Peter J., and Jonathan L. Willis. 2016. "Real Rigidities and Nominal Price Changes." *Economica*, 83(331): 443–472. 6

Leontief, Wassily W. 1936. "Quantitative Input and Output Relations in the Economic Systems of the United States." *The Review of Economics and Statistics*, 18(3): 105. 5

Lerner, Abba P. 1934. "The Concept of Monopoly and the Measurement of Monopoly Power." *The Review of Economic Studies*, 1(3): 157–175. 1

Liu, Ernest. 2019. "Industrial Policies in Production Networks." *The Quarterly Journal of Economics*, 134(4): 1883–1948. 5

Long, John B., and Charles I. Plosser. 1983. "Real Business Cycles." *Journal of Political Economy*, 91(1): 39–69. 5

Marschak, Thomas, and Reinhard Selten. 1974. *General Equilibrium with Price-Making Firms.* Springer-Varlag, Berlin. 7

Melitz, Marc J., and Gianmarco I. P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies*, 75(1): 295–316. 6, 25

Negishi, Takashi. 1961. "Monopolistic Competition and General Equilibrium." *The Review of Economic Studies*, 28(3): 196. 7

Negishi, Takashi. 1972. *General Equilibrium Theory and International Trade.* North-Holland, Amsterdam. 7

Nikaido, Hukukane. 1975. *Monopolistic Competition and Effective Demand.* Preceton University Press. 7

Nord, Lukas. 2022. "Shopping, Demand Composition, and Equilibrium Prices." *SSRN Electronic Journal*. 18

Electronic copy available at: https://ssrn.com/abstract=4306329

Oberfield, Ezra. 2018. "A Theory of Input-Output Architecture." *Econometrica*, 86(2): 559–589. 6

Peters, Michael. 2020. "Heterogeneous Markups, Growth, and Endogenous Misallocation." *Econometrica*, 88(5): 2037–2073. 6

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic Dynamics*, 11(4): 707–720. 5

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2021. "Diverging Trends in National and Local Concentration." *NBER Macroeconomics Annual*, 35: 115–150. 6

Salinger, Michael A. 1989. "The Meaning of "Upstream" and "Downstream" and the Implications for Modeling Vertical Mergers." *The Journal of Industrial Economics*, 37(4): 373. 7

Sangani, Kunal. 2022. "Markups Across the Income Distribution: Measurement and Implications." *SSRN Electronic Journal*. 18

Scherer, Frederic M., and David Ross. 1990. *Industrial Market Structure and Economic Performance.* Boston: Houghton Mifflin Company. 2

Schmalensee, Richard. 1989. "Chapter 16 Inter-industry studies of structure and performance." In *Handbook of Industrial Organization*. Vol. 2, 951–1009. 6

Syverson, Chad. 2019. "Macroeconomics and Market Power: Context, Implications, and Open Questions." *Journal of Economic Perspectives*, 33(3): 23–43. 3

Tirole, Jean. 1988. *The Theory of Industrial Organization.* The MIT Press. 9

von Ungern-Sternberg, Thomas. 1999. "Percentage Retail Mark-Ups." *Swiss Journal of Economics and Statistics (SJES)*, IV(1): 539–557. 2, 11

Voronina, Mariia. 2022. "Market Power in Production Networks." *Doctoral dissertation, Harvard University Graduate School of Arts and Sciences*, 74–101. 5

Young, Allan Richard. 1991. "Vertical Structure and Nash Equilibrium: A Note." *The Journal of Industrial Economics*, 39(6): 717. 11

Electronic copy available at: https://ssrn.com/abstract=4306329

A Additional Tables and Figures

Table 4: 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	1.895
	Insurance carriers and related activities	1.411
	Paper products	1.343
	Wood products	1.340
	Forestry, fishing, and related activities	1.197
	Warehousing and storage	1.158
	Chemical products	1.153
	Miscellaneous professional, scientific, and technical services	1.148
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

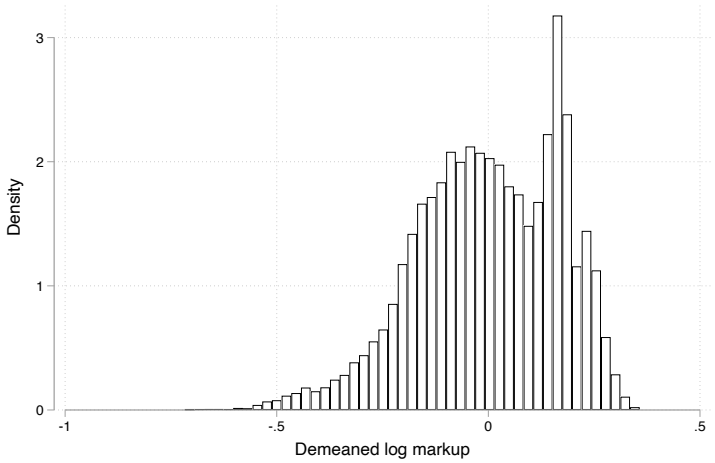
Electronic copy available at: https://ssrn.com/abstract=4306329

Table 5: Markup-to-Market-Share Relations (Cost-Share Approach)

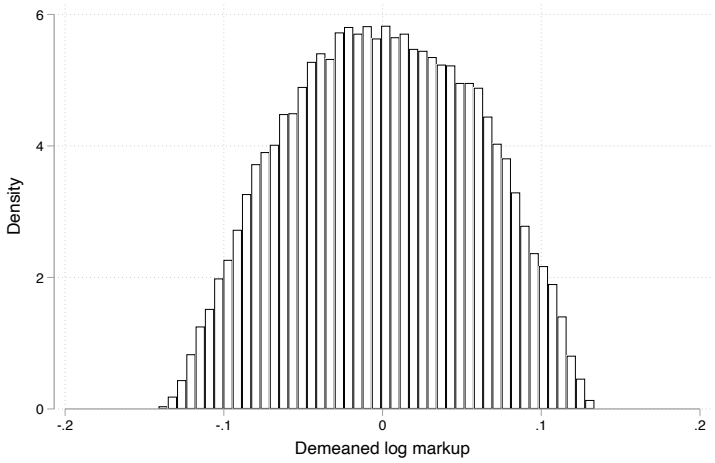
	(1)	(2)	(3)	(4)	(5)	(6)
log s	0.00847*	0.00787*	0.00787*	0.00924**	0.00970**	0.00896*
	(0.00471)	(0.00477)	(0.00466)	(0.00466)	(0.00477)	(0.00460)
log s × cycle		-0.174***			-0.101	
		(0.0396)			(0.0714)	
log s × SP		0.0830***			0.0642**	
		(0.0177)			(0.0298)	
log s × SC		0.109***			0.0860***	
		(0.0195)			(0.0323)	
log s × CP		-0.220***			-0.171**	
		(0.0670)			(0.0717)	
log s × eta			-0.174***			-0.153***
			(0.0324)			(0.0410)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	sic2 × year	sic2 × year	sic2 × year	BEA × year	BEA × year	BEA × year
Observations	66330	66330	66330	66408	66408	66408
Within R^2	.0006	.0073	.0073	.0007	.0045	.0042

Standard errors in parentheses
Standard errors clustered at firm level
* p<0.10, ** p<0.05, *** p<0.01

Electronic copy available at: https://ssrn.com/abstract=4306329



(a) Considering heterogeneity across industries



(b) Using only market shares

Figure 8: Histogram of Predicted Log Markups (around Mean)

Notes: Panel (a) displays the histogram of log markups around the mean using predicted values from Column (2) in Table 2, excluding fixed effects. Panel (b) uses predicted values from Column (1) in the same table, excluding fixed effects too.

Electronic copy available at: https://ssrn.com/abstract=4306329

Online Appendix for

"Network Lerner Index: Demand and Distortions across Industries"

Lingxuan Wu

December 17, 2022

B Proofs and Additional Results

The results stated in Lemmas 3 and 4 are developed by Baqaee and Farhi (2020). Here I reuse them with slight modification. I assume only one factor (labor) in production, and use the wage as numeraire.

Lemma 3. (Elasticity of consumption price, Baqaee and Farhi, 2020) *The elasticity of consumption price P w.r.t. markup μ_i is*

$$\frac{\partial \log P}{\partial \log \mu_i} = \tilde{\Psi}_{0i} = \tilde{\lambda}_i. \tag{47}$$

Proof of Lemma 3. Firms minimize input costs and charge markups. By Shephard's lemma,

$$\mathrm{d} \log p_j = \mathrm{d} \log \mu_j + \sum_{i=1}^{N} \tilde{\Omega}_{ji} \mathrm{d} \log p_i, \tag{48}$$

leading to

$$\mathrm{d} \log p_j = \sum_{j=1}^{N} \tilde{\Psi}_{ji} \mathrm{d} \log \mu_i. \tag{49}$$

□

Lemma 4. (Elasticity of Domar weight, Baqaee and Farhi, 2020) *The elasticity of each firm's Domar weight w.r.t. its own markup is*

$$\frac{\partial \log \lambda_i}{\partial \log \mu_i} = -\underbrace{(\Psi_{ii} - 1)}_{\text{feedback via cycle}} + \underbrace{\left(1 - \theta_p\right) \sum_{j \in \mathcal{I}} \left[\frac{\Psi_{0j}}{\mu_j \Psi_{0i}} \mathrm{Cov}_{\tilde{\Omega}^{(j)}} \left(\tilde{\Psi}_{(i)}, \Psi_{(i)} \right) \right]}_{\text{substitution in production}} + \underbrace{(1 - \theta_c) \frac{1}{\lambda_i} \mathrm{Cov}_{\tilde{\Omega}^{(0)}} \left(\tilde{\Psi}_{(i)}, \Psi_{(i)} \right)}_{\text{substitution in consumption}}, \tag{50}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

$$= - \underbrace{(\Psi_{ii} - 1)}_{\text{feedback via cycle}} + \underbrace{\left(1-\theta_p\right)\left(\tilde{\Psi}_{ii} + \Psi_{ii} - 1 - \tilde{\lambda}_i\right)}_{\text{substitution with uniform elasticity}} - \underbrace{\left(\theta_c - \theta_p\right)\left(\frac{\sum_j \Omega_{0j}\tilde{\Psi}_{ji}\Psi_{ji}}{\lambda_i} - \tilde{\lambda}_i\right)}_{\text{substitution due to differential elasticity}}. \quad (51)$$

Proof of Lemma 4. First we derive the changes in intput-output matrices $\Omega, \tilde{\Omega}$, and then characterize the changes in Leontief inverses $\Psi, \tilde{\Psi}$, which gives rise to eq. (50). Last we show how it simplifies to eq. (51).

$$\mathrm{d}\Omega_{ji} = -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\left(1-\theta_j\right)\left(\mathrm{d}\log p_i - \sum_l \tilde{\Omega}_{jl}\mathrm{d}\log p_l\right) \quad (52)$$

$$= -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\mathrm{d}\log p, I_{(i)}\right), \quad (53)$$

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. Using

$$\mathrm{d}\log p = \sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\mu_k \quad (54)$$

from Lemma 3, we rewrite $\mathrm{d}\Omega_{ji}$ as

$$\mathrm{d}\Omega_{ji} = -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\mu_k, I_{(i)}\right). \quad (55)$$

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

$$\begin{aligned}\mathrm{d}\lambda_n = \mathrm{d}\Psi_{0n} &= -\sum_j \Psi_{0j}\mathrm{d}\log\mu_j \sum_i \Omega_{ji}\Psi_{in} \\ &\quad + \sum_j \frac{\Psi_{0j}}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\mu_k, \sum_i I_{(i)}\Psi_{in}\right) \quad (56) \\ &= -\sum_j \Psi_{0j}\mathrm{d}\log\mu_j\left(\Psi_{jn} - \delta_{jn}\right) \\ &\quad + \sum_j \frac{\Psi_{0j}}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\mu_k, \Psi_{(n)}\right), \quad (57)\end{aligned}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

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

$$\frac{\mathrm{d}\log\lambda_n}{\mathrm{d}\log\mu_n} = -\left(\Psi_{nn} - 1\right) + \frac{1}{\lambda_n}\sum_j \frac{\lambda_j}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right), \tag{58}$$

as claimed in eq. (50).

Since we have $\theta_0 = \theta_c$ for the consumer and $\theta_j = \theta_p, \forall j \neq 0$ for the producers, we make use of the following result to simplify the own elasticity of Domar weight to markup,

$$\sum_j \frac{\lambda_j}{\mu_j}\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) = \lambda_n\left(\tilde{\Psi}_{nn} + \Psi_{nn} - 1 - \tilde{\lambda}_n\right). \tag{59}$$

To see that, we open up the LHS

$$\begin{aligned}
\sum_j \frac{\lambda_j}{\mu_j}\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) &= \sum_j \frac{\lambda_j}{\mu_j}\left[\sum_m \tilde{\Omega}_{jm}\tilde{\Psi}_{mn}\Psi_{mn} - \left(\sum_m \tilde{\Omega}_{jm}\tilde{\Psi}_{mn}\right)\left(\sum_m \tilde{\Omega}_{jm}\Psi_{mn}\right)\right] && (60)\\
&= \sum_{j,m} \Psi_{0j}\Omega_{jm}\tilde{\Psi}_{mn}\Psi_{mn} - \sum_j \Psi_{0j}\left(\sum_m \tilde{\Omega}_{jm}\tilde{\Psi}_{mn}\right)\left(\sum_m \Omega_{jm}\Psi_{mn}\right) && (61)\\
&= \sum_m \left(\Psi_{0m} - \delta_{0m}\right)\tilde{\Psi}_{mn}\Psi_{mn} - \sum_j \Psi_{0j}\left(\tilde{\Psi}_{jn} - \delta_{jn}\right)\left(\Psi_{jn} - \delta_{jn}\right) && (62)\\
&= -\tilde{\Psi}_{nn}\Psi_{nn} + \sum_j \delta_{jn}\Psi_{0j}\left(\tilde{\Psi}_{jn} + \Psi_{jn} - \delta_{jn}\right) && (63)\\
&= -\tilde{\lambda}_n\lambda_n + \lambda_n\left(\tilde{\Psi}_{nn} + \Psi_{nn} - 1\right), && (64)
\end{aligned}$$

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 $\sum_j \Psi_{0j}\tilde{\Psi}_{jn}\Psi_{jn}$.

Therefore, we can simplify the own elasticity of Domar weight to

$$\frac{\mathrm{d}\log\lambda_n}{\mathrm{d}\log\mu_n} = -\left(\Psi_{nn} - 1\right) + \frac{1}{\lambda_n}\left[\left(1-\theta_p\right)\sum_j \frac{\lambda_j}{\mu_j}\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) - \left(\theta_c - \theta_p\right)\lambda_0\,\mathrm{Cov}_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right)\right] \tag{65}$$

$$= -\left(\Psi_{nn} - 1\right) + \left(1-\theta_p\right)\left(\tilde{\Psi}_{nn} + \Psi_{nn} - 1 - \tilde{\lambda}_n\right) - \left(\theta_c - \theta_p\right)\frac{1}{\lambda_n}\mathrm{Cov}_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right). \tag{66}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

The last term can be further simplified as

$$\left(\theta_c - \theta_p\right) \frac{1}{\lambda_n} \mathrm{Cov}_{\tilde{\Omega}^{(0)}} \left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) = \left(\theta_c - \theta_p\right) \frac{1}{\lambda_n} \left[\sum_j \Omega_{0j} \tilde{\Psi}_{jn} \Psi_{jn} - \left(\sum_j \tilde{\Omega}_{0j} \tilde{\Psi}_{jn} \right) \left(\sum_j \Omega_{0j} \Psi_{jn} \right) \right] \tag{67}$$

$$= \left(\theta_c - \theta_p\right) \left[\frac{\sum_j \Omega_{0j} \tilde{\Psi}_{jn} \Psi_{jn}}{\lambda_n} - \tilde{\lambda}_n \right]. \tag{68}$$

□

Proof of Proposition 1. We combine Lemma 3 and 4,

$$\eta_n = \left(\Psi_{nn} - 1\right) - \left(1 - \theta_p\right) \left(\tilde{\Psi}_{nn} + \Psi_{nn} - 1 - \tilde{\lambda}_n\right) + \left(\theta_c - \theta_p\right) \left[\frac{\sum_j \Omega_{0j} \tilde{\Psi}_{jn} \Psi_{jn}}{\lambda_n} - \tilde{\lambda}_n \right] + \left(\epsilon - 1\right) \tilde{\lambda}_n + 1 \tag{69}$$

$$= \theta_p \Psi_{nn} - \left(1 - \theta_p\right) \left(\tilde{\Psi}_{nn} - 1\right) + \left(\theta_c - \theta_p\right) \frac{\sum_j \Omega_{0j} \tilde{\Psi}_{jn} \Psi_{jn}}{\lambda_n} + \left(\epsilon - \theta_c\right) \tilde{\lambda}_n. \tag{70}$$

□

Proof of Proposition 2. The FOC of firm f_i is, based on eq. (31),

$$\frac{\mu_{f_i}}{\mu_{f_i} - 1} = -\frac{\partial \log s_{f_i}}{\partial \log \mu_{f_i}} - \frac{\partial \log \left(p_i y_i / P\right)}{\partial \log \left(\mu_i / z_i\right)} \frac{\partial \log \left(\mu_i / z_i\right)}{\partial \log \left(\mu_{f_i}\right)} + 1 \tag{71}$$

$$= -\frac{\partial \log s_{f_i}}{\partial \log \mu_{f_i}} + \left(\eta_i - 1\right) \frac{\partial \log \left(\mu_i / z_i\right)}{\partial \log \left(\mu_{f_i}\right)} + 1. \tag{72}$$

Since industry sales share s_{f_i} satisfies

$$s_{f_i} = \frac{p_i^{1-\rho}}{p_i^{1-\rho}} = \frac{\left(\mu_{f_i} / z_{f_i}\right)^{1-\rho}}{\left(\mu_i / z_i\right)^{1-\rho}}, \tag{73}$$

we can rewrite the FOC as

$$\frac{\mu_{f_i}}{\mu_{f_i} - 1} = -\left(1 - \rho\right) + \left(1 - \rho\right) \frac{\partial \log \left(\mu_i / z_i\right)}{\partial \log \left(\mu_{f_i}\right)} + \left(\eta_i - 1\right) \frac{\partial \log \left(\mu_i / z_i\right)}{\partial \log \left(\mu_{f_i}\right)} + 1 \tag{74}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

$$= \rho + (\eta_i - \rho) \frac{\partial \log(\mu_i / z_i)}{\partial \log\left(\mu_{f_i}\right)} \tag{75}$$

$$= \rho + (\eta_i - \rho) s_{f_i} \tag{76}$$

$$= \left(1 - s_{f_i}\right) \rho + s_{f_i} \eta_i, \tag{77}$$

in which the second-to-last step uses

$$\frac{\mu_i}{z_i} = \frac{p_i}{mc_i} = \frac{\left(\sum_{f_i \in \mathcal{F}_i} \left(\frac{\mu_{f_i}}{z_{f_i}} mc_i\right)^{1-\rho}\right)^{\frac{1}{1-\rho}}}{mc_i} = \left(\sum_{f_i \in \mathcal{F}_i} \left(\frac{\mu_{f_i}}{z_{f_i}}\right)^{1-\rho}\right)^{\frac{1}{1-\rho}}. \tag{78}$$

□

Proof of Lemma 1. For two firms f_i, f_i' in the same industry i, with $z_{f_i} \geq z_{f_i'}$, take the difference between their Lerner conditions (37) and get

$$\frac{1}{\mu_{f_i} - 1} - \frac{1}{\mu_{f_i'} - 1} = -(\rho - \eta_i)\left(s_{f_i} - s_{f_i'}\right). \tag{79}$$

Using $\mu_{f_i} = \frac{p_{f_i}}{\frac{1}{z_{f_i}} \mathrm{mc}_i}$ and $s_{f_i} = \frac{p_{f_i}^{1-\rho}}{p_i^{1-\rho}}$, we can rewrite the equation in terms of $p_{f_i}, p_{f_i'}$ as

$$\frac{1}{\frac{z_{f_i} p_{f_i}}{\mathrm{mc}_i} - 1} - \frac{1}{\frac{z_{f_i'} p_{f_i'}}{\mathrm{mc}_i} - 1} = -(\rho - \eta_i) \frac{p_{f_i}^{1-\rho} - p_{f_i'}^{1-\rho}}{p_i^{1-\rho}}. \tag{80}$$

The LHS is decreasing in both z_{f_i} and p_{f_i}, and the RHS is increasing in p_{f_i} when $\rho > 1$ and $\rho > \eta_i$. Suppose by contradiction that $p_{f_i} \geq p_{f_i'}$, we would have LHS $< 0 \leq$ RHS. Thus for the equality to hold, it must be the case that $p_{f_i} < p_{f_i'}$. Since s_{f_i} is decreasing in p_{f_i}, we also have $s_{f_i} > s_{f_i'}$. Since $\eta_i < \rho$, from the two firms Lerner condition we conclude that $\mu_{f_i} > \mu_{f_i'}$.□

Proof of Proposition 3. 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 $z_1 = z \geq z_2 = 1$. 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-level elasticity is $e(s) = (1 - s)\rho + s\eta$, and

Electronic copy available at: https://ssrn.com/abstract=4306329

the markup is $\mu(s) = \frac{e(s)}{e(s)-1}$. We show that if the industry-level elasticity 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 larger markup.

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 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

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

Electronic copy available at: https://ssrn.com/abstract=4306329

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

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 $e_1' e_2 < e_1 e_2'$. 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 Proposition 1 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

$$\frac{s_1}{s_2} = \frac{(\mu_1/z)^{1-\rho}}{\mu_2^{1-\rho}}, \tag{81}$$

which gives

$$\mu_1 = z\mu_2\left(\frac{s_2}{s_1}\right)^{\frac{1}{\rho-1}}. \tag{82}$$

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 eq. (82) yields

$$\log\mu_1 - \log\mu_2 = \log z + \frac{1}{\rho-1}\left(\log s_2 - \log s_1\right), \tag{83}$$

and thus

$$\kappa = \frac{\log z}{\log s_1 - \log s_2} - \frac{1}{\rho-1}. \tag{84}$$

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

Electronic copy available at: https://ssrn.com/abstract=4306329

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 (34),

$$\frac{1}{z_{ind}} = \frac{1}{z}\frac{y_1}{y} + \frac{y_2}{y} = \frac{\frac{1}{z}\frac{y_1}{y_2} + 1}{\left[1 + \left(\frac{y_1}{y_2}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}}. \tag{85}$$

Claim that the RHS is decreasing in y_1/y_2 and thus the industry z_{ind} is increasing in y_1/y_2. 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

$$\frac{y_1}{y_2} = \frac{p_1^{-\rho}}{p_2^{-\rho}} = \left(\frac{s_1}{s_2}\right)^{\frac{\rho}{\rho-1}} = \left(\frac{s_1}{1-s_1}\right)^{\frac{\rho}{\rho-1}}, \tag{86}$$

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

$$\frac{y_1}{y_2} = \frac{p_1^{-\rho}}{p_2^{-\rho}} = \left(\frac{\mu_1/z}{\mu_2}\right)^{-\rho} = z^{\rho}\left(\frac{\rho(1-s_1)+\eta s_1}{\rho(1-s_2)+\eta s_2}\right)^{-\rho} < z^{\rho} \tag{87}$$

holds as long as $s_1 > s_2$. Define

$$h(x) = \frac{\frac{1}{z}x + 1}{\left(1 + x^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}}, \tag{88}$$

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 $x < z^{\rho}$,

$$h'(x) \propto \frac{1}{z}\left(1 + x^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} - \left(\frac{1}{z}x + 1\right)\left(1 + x^{\frac{\rho-1}{\rho}}\right)^{\frac{1}{\rho-1}} x^{-\frac{1}{\rho}} \tag{89}$$

$$\propto \frac{1}{z}\left(1 + x^{\frac{\rho-1}{\rho}}\right) - \left(\frac{1}{z}x + 1\right)x^{-\frac{1}{\rho}} \tag{90}$$

$$= \frac{1}{z} - x^{-\frac{1}{\rho}}, \tag{91}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

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

5. *Industry markup.* Use the expression of industry markup (35),

$$\frac{1}{\mu_{ind}} = \frac{s_1}{\mu_1} + \frac{s_2}{\mu_2} = \frac{1}{\mu_1}\left(s_1 + \frac{\mu_1}{\mu_2}s_2\right). \tag{92}$$

Use

$$\frac{s_1}{s_2} = \frac{(\mu_1/z)^{1-\rho}}{\mu_2^{1-\rho}}, \tag{93}$$

which leads to

$$\frac{\mu_1}{\mu_2} = z\left(\frac{s_1}{s_2}\right)^{\frac{1}{1-\rho}}, \tag{94}$$

$$\frac{1}{\mu_{ind}} = \frac{1}{\mu_1}\left(s_1 + zs_1^{-\frac{1}{\rho-1}}s_2^{\frac{\rho}{\rho-1}}\right) = \frac{s_1^{\frac{\rho}{\rho-1}} + zs_2^{\frac{\rho}{\rho-1}}}{\mu_1 s_1^{\frac{1}{\rho-1}}} = \frac{s_1^{\frac{\rho}{\rho-1}} + zs_2^{\frac{\rho}{\rho-1}}}{z\mu_2 s_2^{\frac{1}{\rho-1}}}. \tag{95}$$

As we have shown that both μ_2 and s_2 increase in η, the denominator of the last term is increasing in η. If its numerator is decreasing in s_2, we will have μ_{ind} increases in η. To see that, notice the numerator can be rewritten as

$$g(s_2) = (1-s_2)^{\frac{\rho}{\rho-1}} + zs_2^{\frac{\rho}{\rho-1}}, \tag{96}$$

and thus

$$g'(s_2) = \frac{\rho}{\rho-1}\left(zs_2^{\frac{1}{\rho-1}} - s_1^{\frac{1}{\rho-1}}\right), \tag{97}$$

which is positive as long as $s_1/s_2 < z^{\rho-1}$. Since we have $\mu_1 > \mu_2$, it follows from

$$\frac{s_1}{s_2} = \frac{(\mu_1/z)^{1-\rho}}{\mu_2^{1-\rho}} < z^{\rho-1}. \tag{98}$$

□

Proof of Lemma 1. Take first-order approximation of (37),

$$-\frac{\bar{\mu}}{(\bar{\mu}-1)^2}\Delta\log\mu_{f_i} = -(\rho-\eta_i)\,\bar{s}\Delta\log_{s_{f_i}} \tag{99}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

$$\Delta \log \mu_{f_i} = \frac{\bar{s}\left(\bar{\mu}-1\right)^2}{\bar{\mu}}\left(\rho-\eta_i\right)\Delta \log s_{f_i}. \tag{100}$$

□

Proposition 4. (Log-log markup-to-market share relation in a hybrid model) *In an ad hoc model where firms $f_i \in \mathcal{F}_i$ within an industry compete à la Cournot but with variable industry-level demand elasticity η_i, firms' markups obey*

$$\frac{\mu_{f_i}}{\mu_{f_i}-1} = \left(\frac{s_{f_i}}{\eta_i} + \frac{1-s_{f_i}}{\rho}\right)^{-1}. \tag{101}$$

To the first order in $\left\{\Delta \log s_{f_i}, \Delta \log \mu_{f_i}\right\}_{f_i \in \mathcal{N}'}$, the markup-to-market-share relation in each industry simplifies to

$$\Delta \log \mu_{f_i} \approx \frac{\bar{s}\left(\bar{\mu}-1\right)^2}{\bar{\mu}} \frac{\left(\frac{1}{\eta_i}-\frac{1}{\rho}\right)}{\left(\frac{\bar{s}}{\eta_i}+\frac{1-\bar{s}}{\rho}\right)^2}\Delta \log s_{f_i}, \tag{102}$$

$$\approx \frac{\bar{s}\left(\bar{\mu}-1\right)^2}{\bar{\mu}} \left(\frac{\frac{1}{\bar{\eta}}-\frac{1}{\rho}}{\left(\frac{\bar{s}}{\bar{\eta}}+\frac{1-\bar{s}}{\rho}\right)^2} - \frac{\frac{1+\bar{s}}{\rho}-\frac{\bar{s}}{\bar{\eta}}}{\bar{\eta}^2\left(\frac{\bar{s}}{\bar{\eta}}+\frac{1-\bar{s}}{\rho}\right)^3}\Delta\eta_i \right)\Delta \log s_{f_i}, \tag{103}$$

the last step of which approximates the slope to the first order in $\Delta\eta_i$.

This proposition suggests that the slope of log markup to log market share is higher in industries with lower $\Delta\eta_i$, as long as the "average" firm in the economy has a market share $\bar{s}$ smaller than $\frac{\bar{\eta}}{\rho-\bar{\eta}}$. This condition holds as the average firm is pretty small.

Proof of Proposition 4. We get (4) by simply replacing η with η_i in the well-known result from Atkeson and Burstein (2008, eq. 16). When firms within an industry compete à la Bertrand/Cournot, the demand elasticity they face is an arithmetic/harmonic average of the industry demand elasticity η_i and within-industry substitution ρ.

Take first-order approximation of (4),

$$-\frac{\bar{\mu}}{\left(\bar{\mu}-1\right)^2}\Delta \log \mu_{f_i} = \frac{\left(\frac{1}{\eta_i}-\frac{1}{\rho}\right)}{\left(\frac{\bar{s}}{\eta_i}+\frac{1-\bar{s}}{\rho}\right)^2}\bar{s}\Delta \log s_{f_i} \tag{104}$$

Electronic copy available at: https://ssrn.com/abstract=4306329

$$\Delta \log \mu_{f_i} = \frac{\bar{s}\,(\bar{\mu}-1)^2}{\bar{\mu}} \left(\frac{\frac{1}{\bar{\eta}} - \frac{1}{\rho}}{\left(\frac{\bar{s}}{\bar{\eta}} + \frac{1-\bar{s}}{\rho}\right)^2} - \frac{\frac{1+\bar{s}}{\rho} - \frac{\bar{s}}{\bar{\eta}}}{\bar{\eta}^2 \left(\frac{\bar{s}}{\bar{\eta}} + \frac{1-\bar{s}}{\rho}\right)^3} \Delta\eta_i \right) \Delta \log s_{f_i}. \tag{105}$$

□

C Data Construction

Network statistics. I use the BEA use tables before redefinitions in producer value at summary level (71 industries) from 1997 to 2019.[34] 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. Based on $\Omega, \Psi, \tilde{\Omega}, \tilde{\Psi}$, I construct the determinants of industry-level elasticities for each industry following the theory predictions in Proposition 1.

Firm-level variables. From Compustat, choose the "Fundamentals Annual" dataset. Select data variable *Fiscal Year,* choose date range 1950-01 to 2022-11, set *gvkey* for company codes, exclude industry format *FS* and currency *CAD*, select identifying information *Company Name, Stock Exchange Code* and *Foreign Incorporation Code*, and inquire *sic, naics, ppegt, ppent, at, sale, oibdp, oiadp, emp, cogs, xsga, capx, bkvlps, mkvalt, lt.*

Keep firms with *fic* equal to *USA* and *curcd* equal to *USD.* Exclude firms with *sic* between 4900 - 4999 (utility firms which are heavily regulated) and between 6900 - 6999 (financial firms whose balance sheets are different). Exclude firm-year observations with assets (*at*) less than 10 million, with negative book, market value, sale or employment (*bkvlps, mkvalt, sale, emp*), or with missing year, SIC, NAICS, sale, COGS, assets, or book liabilities (*fyear, sic, naics, sale, cogs, at, lt*). Deflate capital-related items (*capx, ppegt, ppent)* using the implicit price deflators for non-residential investment in 2012 dollars. Deflate other variables related to sales and cost *(bkvlps, cogs, emp, lt, sale, oiadp, oibdp, xsga, mkvalt)*

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

Electronic copy available at: https://ssrn.com/abstract=4306329

using the implicit deflators for GDP. For each year, exclude firms with sales or COGS-to-sales ratio *(cogs/sale)* in the top and bottom 5% for the cost-share approach, and exclude firms with estimated markups in the top and bottom 5% for the production-function approach. When estimating markups using production function for each 2-digit NAICS industry on 5-year rolling windows, I drop panels with less than 500 observations for accuracy, but the result is robust to this.

D TFP Loss

We use the predicted log markups from the empirical analysis, whose variation is valid but level is not. Thus we shift the level of the log markups to target a (sales-weighted harmonic) average markup of 1.2. Since our focus here is on the TFP loss due to misallocation which is only tied to dispersion of markups, this should be innocuous. This gives us the levels of markups μ_{f_i} for each firm. For those BEA industries not in the empirical analysis, we assume that there is a monopoly firm in each of them which charge a markup equal to the average markup of 1.2.

To apply the formula (45), we calculate the firm-level input-output matrices $\tilde{\check{\Omega}}, \check{\Omega}, \tilde{\check{\Psi}}, \check{\Psi}$ from the industry-level input-output matrices $\tilde{\Omega}, \Omega, \tilde{\Psi}, \Psi$. We order firm 0 first. The for each industry, we order the fictitious industry packer before the real production firms. The last firm, after all industries, represents the fictitious labor producer.

We first calculate $\tilde{\check{\Omega}}$ from $\tilde{\Omega}$ row by row. For firm 0, it uses industry composites from industry packers only. These are the only non-zero elements and have the same value as in $\tilde{\Omega}$. The case is similar for each row corresponding to a real firm. As they only use industry composites, the only non-zero elements correspond to sourcing from industry packers and have the same values as in $\tilde{\Omega}$. For each industry packer i, they only source from the real production firms in that industry. Thus the non-zero elements of those rows correspond to the market shares $\left\{s_{f_i}\right\}_{f_i \in \mathcal{F}_i}$. The last row, corresponding to firm L, is zero. This concludes the construction of $\tilde{\check{\Omega}}$. Then $\check{\Omega}$ is given by $\tilde{\check{\Omega}}_{f_i g_i} \mu_{f_i}^{-1}$ element by element where μ_{f_i} equals one for firm 0, industry packers, and firm L. Last, we have $\check{\Psi} = \left(I - \check{\Omega}\right)^{-1}$. Each firm's Domar weight λ_{f_i} is equal to $\check{\Psi}_{0 f_i}$.

When calculating the TFP loss according to Definition 3, we only change the markups of real production firms. We fix the markups of firm 0, L and all industry packers to 1, as they are fictitious and only serve the purpose of simplifying the notation.

Electronic copy available at: https://ssrn.com/abstract=4306329