Product Differentiation and Oligopoly: a Network Approach

Bruno Pellegrino

University of Maryland

August 30, 2022

Abstract

This paper develops a theory of oligopoly and markups in general equilibrium, micro-founded by hedonic demand. I propose a model where oligopolistic firms compete in a network of product market rivalries that emerges endogenously out of the characteristics of the products and services they supply; these oligopolists co-exist with a fringe of atomistic competitors that enter and exit endogenously. The model embeds a novel, highly tractable and scalable demand system (GHL) that can be estimated for the universe of public corporations in the USA using publicly-available data. Using the model, I decompose the level and the trend of markups into: 1) a new measure of firm productivity that accounts for product quality that is comparable across firms in the cross section; 2) a metric of network centrality, which captures the intensity of competition from substitute products, and determines firms' ability to capture surplus in the form of monopoly rents. I estimate that, in 2019, oligopoly lowered total surplus by 11.5% and depressed consumer surplus by 31%. My analysis also suggests that both numbers were significantly lower in the mid-90s (7.9% and 21.5%, respectively), and that M&A activity is increasingly skewed towards firm pairs that compete closely in the product market. These estimates should be interpreted with care due to data limitations.

JEL Codes: D2, D4, D6, E2, L1, O4

Keywords: Competition, Concentration, General Equilibrium, Market Power, Markups, Mergers, Monopoly, Networks, Oligopoly, Text Analysis.

This paper received the following awards: EEA Best Economics Job Market Paper, WFA Best Corporate Finance Paper, UCLA Xavier Drèze Prize for Outstanding Research Paper and Best Presentation Award at the Unicredit Job Market Bootcamp. I gratefully acknowledge financial support from the Price Center for Entrepreneurship at UCLA Anderson. I am indebted to my dissertation advisors Hugo Hopenhayn, Nico Voigtländer, Andy Atkeson, Ed Leamer, Stavros Panageas and Romain Wacziarg for their invaluable advice and guidance, and to five anonymous referees for detailed feedback that significantly improved the paper. I benefited from thoughtful conference discussions by Jeff Campbell, Michael Gofman, Naveen Khanna, Isabelle Mejean and Jane Olmstead-Rumsey. I also received valuable comments from John Asker, David Baqaee, Alberto Bisin, Ariel Burstein, Michela Giorcelli, Thomas Groesbeck, Kurt Mitman, Oleg Itskhoki, Nicola Persico, Sebastian Ottinger, Tommaso Valletti, Maria Lucia Yanguas, and workshop participants at the NBER IO Program Meeting, the NBER Mega-Firms Conference, the WFA, the NBER Summer Institute (EFDIS), and many other conferences and seminars. Finally, I acknowledge outstanding data collection assistance by Michele Schincaglia and Elliot Oh.

1. Introduction

Recent empirical evidence suggests that firms vary enormously in their degree of market power, as measured as the price markup over marginal cost (De Loecker et al., 2020). In the United States, both the level and the dispersion of these markups appears to have increased over the last few decades. This increase has been accompanied with a secular rise in industry concentration (Grullon et al., 2019; Kwon et al., 2021).

Why do some firms charge such large markups, while others price close to marginal cost? What drives the observed changes in the distribution of markups? And what are the implications of these trends for aggregate productivity and consumer welfare? The answers to these questions have significant implications for public policy: some argue that they reflect an oligopolization of U.S. industries. The newly-appointed chairman of the Federal Trade Commission even went as far as suggesting that a revised antitrust paradigm needs to be developed in response to these trends (Khan, 2018).

Standard price theory arguments suggest that the recent upward trends in markups and industry concentration may have a significant impact on efficiency and welfare. However, interpreting these trends presents an imposing methodological challenge. The study of market power has traditionally resided within Empirical Industrial Organization (EIO). Yet, there is a consensus that these trends are macroeconomic in nature. The theoretical and empirical tools developed by the EIO literature, which are typically deployed to tackle competition policy problems, require data that is not available for more than a handful of industries (Syverson, 2019), and are thus unavailable to the growing number of economists that seek to study market power from a macroeconomic angle.

Moreover, studying market power in macroeconomics comes with an even more daunting *conceptual* challenge: when we study oligopoly within an industry, a key input into the analysis is the *market definition* – that is, the economist needs to understand, for each relevant product, what is the set of available substitutes and how strong is the demand elasticity of substitution between each product pair. Market definition is often a deciding factor in antitrust litigations. It is hard enough to define individual markets without incurring in a variety of ideological biases; the complexity of this challenge is compounded several times over when we move from an industry setting to a macro setting. To study market power in macro setting, we face the seemingly insurmountable challenge of objectively defining multiple, potentially-interlinked product markets at the same time. Industry classifications (such as NAICS) are not a solution to this problem, as they are based on similarities in the production process, not on the degree of product substitutability. In other words, they are appropriate for estimating production functions, but they are not suitable to estimate demand elasticities. In addition, the very concepts of industry and sector are arbitrary by construction. In fact, larger companies (those more likely to have market power) have been shown to strategically manipulate their industry classification—a phenomenon that has been dubbed *industry window dressing* (Chen et al., 2016).

In this paper I tackle these challenges head-on by developing a novel theory of oligopoly and markups in general equilibrium. I then take the theory to the data to unpack the sources of markups heterogeneity, and to investigate the changing welfare consequences of oligopoly power in the United States. The model that I propose is populated by a finite set of *granular* firms that behave as oligopolists and by a continuum of atomistic producers that behave competitively and can enter and exit. To model product market competition among the oligopolists, I propose a new, highly tractable and scalable demand system – Generalized Hedonic-Linear (GHL) – which I estimate using a dataset recently developed by Hoberg and Phillips (2016). This dataset provides measures of product similarity for all pairs of publicly-traded corporations in the US. The empirical implementation of the model allows me to disentangle the underlying sources of markups heterogeneity, and study how consumer surplus and the welfare costs of oligopoly evolved as a consequence of industry consolidation during this period.

Using my novel theoretical framework, I document that the increased concentration of US industries over the past twenty years was accompanied by an increase in oligopoly power, measured as: (1) an increase in the

deadweight losses induced by oligopolistic behavior; (2) a decline in the share of total surplus that accrues to consumers.

Economists have long been concerned with market power. Since the 1980s, the EIO literature has been developing a conceptual "toolkit" that researchers and antitrust enforcement practitioners have used to analyze market power within industries (Einav and Levin, 2010). The EIO approach requires the researcher to first understand the structure of product market rivalries in an industry: a firm's ability to price above marginal cost depends crucially on the availability of substitutes. As a consequence, this literature has shown that oligopoly power is inextricably linked to the notion of product differentiation: to measure a firm's market power in an industry with n firms, the economist effectively needs to first estimate n^2 cross-price demand elasticities—one for each pair of rivals. This is because variations in the supply of any product cause the residual demand curve of every competing product to shift, and the shift is larger for products that are closer substitutes.

In industry studies, demand estimation is nowadays based on hedonic models (Berry, Levinsohn and Pakes, 1995). The current resurgence in market power and antitrust research, however, has a distinctive macroeconomic angle. Because we do not observe output volume, prices, or product characteristics for a sufficiently large cross-section of industries, this EIO approach cannot be directly applied in a macroeconomic context.

Despite these challenges, the macroeconomics literature has made significant progress in incorporating market power in general equilibrium models: Baqaee and Farhi (2020, henceforth BF) have recently shown how to approximate the welfare costs of imperfect competition, under minimal assumptions, using the cross-sectional distribution of markups. Their approach is (by design) very general, but it is also agnostic about the origin of markups, which are assumed to be exogenous and observable. Because this approach is silent about how the observed cross-section of markups arises in the first place, a separate theory of markups formation is needed in order to simulate the welfare consequences of a change in market structure.

This study breaks new ground by providing a theory of firm size and profitability that generalizes the Cournot oligopoly model to differentiated products and hedonic demand, and embeds it in a general equilibrium model. The objective of this study is not to measure markups, but rather to understand what drives their distribution. Through the model, I can also quantify the contribution of each individual producer to aggregate welfare, and I can study the general equilibrium effects of events that are relevant to antitrust policy, such as mergers.

To achieve this, my theoretical model dispenses with the notions of industry and sector altogether, building instead on the tradition of hedonic demand (Lancaster, 1966; Rosen, 1974). My model links the cross-price elasticity of demand between all firms in the economy to the fundamental attributes of the products they supply. Each firm's output is modeled as a bundle of characteristics that are individually valued by the representative consumer. The cross-price elasticity of demand between two firms depends on the characteristics embedded in their output. If the product portfolios of two companies contain similar characteristics, the cross-price elasticity of demand between their products is high. The result is a rather different picture of the product market: not a static collection of sectors, but a dynamic network, in which the distance between nodes reflects product similarity and strategic interaction between firms.

The key assumptions of my model are: (1) the representative consumer's preferences are described by a utility function that is quadratic-in-characteristics ; (2) firms compete à la Cournot[1]; (3) the marginal cost function is linear in output. Based on these assumptions, the firms in my model play a linear-quadratic game over a weighted network, a type of potential game that has been extensively studied in the micro theory literature (see Ballester, Calvó-Armengol and Zenou, 2006; Ushchev and Zenou, 2018).

This is the first paper to show how to derive the network Cournot model starting from a hedonic utility specification, to embed the game in a general equilibrium framework and to take the model to the data in a structural way.

[1]I also study the Bertrand case in the Appendix.

To estimate my model, I use a recently-developed data set (Hoberg and Phillips, 2016, henceforth HP) that provides continuous measures of product similarity for every pair of publicly traded firms in the United States. These product-similarity scores are based on a computational-linguistics analysis of product descriptions obtained from SEC filings, and are mapped by my model into an $n \times n$ matrix of cross-price demand elasticities. Moreover, because HP's similarity scores are time-varying (yearly observations since 1997), my model is unique in that the degree of product substitution between individual firms is allowed to change over time.

Crucially, the empirical implementation of my model does not require any proprietary or confidential data, and is computationally tractable. Two datasets are required: Compustat and HP's cosine similarity data, which the authors have made publicly-accessible through an online repository.[2]

I validate the GHL framework in several ways. First, I show that the macro-cluster structure of the network of product rivalries overlaps (almost perfectly) with GIC product industries. Then, I show that, for a sample of firm pairs that have been studied in the IO literature, the cross-price demand elasticities implied by the paper match almost perfectly (without directly targeting) the corresponding microeconometric estimates. Finally, I show that the markups implied by the model correlate highly (both over time and in the cross section) with those estimated by De Loecker, Eeckhout and Unger (2020). In addition, HP provide additional validation in their paper.

I use my model to compute the deadweight loss from oligopoly and to simulate changes in total surplus and consumer surplus for a number of counterfactuals. I find that the welfare costs of oligopoly are sizable. By moving to an allocation in which firms price at marginal cost (that is, in which they behave as if they were atomistic players in a perfectly-competitive market), total surplus rises by 12.6 percentage points; consumer surplus increases by 31%, partly due to total surplus being reallocated from producers to consumers. By computing a separate counterfactual that keeps the aggregate labor supply fixed (markups are equalized, rather than eliminated), I determine that a significant share of the welfare loss from oligopoly—about 7.7 percentage points of the aforementioned 12.6—occurs by way of factor misallocation. In other words, the deadweight loss is driven not only by an underutilization of inputs, but also by a suboptimal mix of goods being produced. I also simulate a counterfactual in which all firms in the economy are owned by a single producer that implements a collusive equilibrium. Under this scenario, total surplus would drop by about one-tenth: with some degree of abstraction, we can think of this estimate as an upper bound to the welfare benefits of antitrust. Also, in this monopolistic/collusive equilibrium consumer surplus would decrease by about 38%, due partly to surplus being reallocated from consumers to producers.

By mapping my model to firm-level data for a period of 26 consecutive years, I investigate the welfare consequences of the rise in concentration and markups between 1996 and 2019. I find that the share of surplus appropriated by companies in the form of oligopoly profits has increased from about 15.5% (in 1996) to 22.7% (in 2019). I perform various robustness checks to ensure (to the best of my ability) that these findings are robust to different assumptions about returns to scale, and that they are unlikely to be driven by the exclusion of foreign and private firms, or by nontradable sectors.

The welfare costs of oligopoly have also increased over this period. In terms of total surplus, the gap between the oligopolistic equilibrium and perfect competition (the deadweight loss) has increased from 7.9% (in 1996) to 12.6% (in 2019). Consumer surplus is adversely affected via two channels: less surplus is produced overall (as a percentage of the surplus that *could* be produced), and less of the diminished surplus is allocated to the consumer in equilibrium. Thus, an important contribution of this paper is to investigate the distributional implications of oligopoly. Overall, my empirical findings suggest that rising concentration and markups did indeed translate into a measurable increase in the welfare loss from oligopoly, and affected how surplus is shared between producers and consumers.

I also use the model to decompose the markups of US public firms into two components. The first is a

[2] See hobergphillips.tuck.dartmouth.edu

novel measure of (*hedonic-adjusted*) *productivity*; the second is a measure of *product market centrality*, which captures the intensity of product market competition coming from producers of substitute products. The first is the primary driver of the cyclical variation in markups, while the latter drives the slow-moving upward trend.

This paper aims to connect the new EIO literature (Einav and Levin, 2010) to two recent and growing branches of macroeconomics that use micro-data.

The first is the literature on networks (Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi, 2012; Carvalho, 2014; Acemoglu, Ozdaglar and Tahbaz-Salehi, 2017; Carvalho and Tahbaz-Salehi, 2019; Baqaee and Farhi, 2020; Carvalho, Nirei, Saito and Tahbaz-Salehi, 2020). I contribute to and expand this literature, which has mostly focused on input-output networks, by considering a different type of network: that of product market rivalries.[3]

The second is the literature on markups and industry concentration (De Loecker, Eeckhout and Unger, 2020; Autor, Dorn, Katz, Patterson and Van Reenen, 2020; Edmond, Midrigan and Xu, 2018; Covarrubias, Gutiérrez and Philippon, 2020; Syverson, 2019). This paper builds on and adds to this body of work by incorporating hedonic demand as well as new data. These features allow me to go beyond markups and concentration, and to create a rich, high-dimensional representation of the competitive environment. In my model, firms differ not only by their productivity, but also by their products' characteristics; as a consequence, each firm has a distinct set of competitors that changes over time, as firms update their product's description in their SEC filings.

The rest of the paper is organized as follows. In Section 2, I present my theoretical model. In Section 3, I present the data used in the empirical part of the paper and show how it is mapped to the model. In Section 4, I validate the GHL demand system introduced in this paper. In Section 5, I present my empirical results. In Section 6, I discuss a number of extensions and robustness checks. In Section 7, I present my conclusions and discuss how my findings can inform the current debate on market power and antitrust policy.

[3]Outside the macro-networks literature, Bloom, Schankerman and Van Reenen (2013) have studied rivalry networks in a seminal empirical study of R&D spillovers.

2. A Network Theory of Oligopoly and Markups

In this section, I present a general equilibrium model in which firms produce differentiated products and compete à la Cournot. For expositional purposes, I start by laying out the basic model that only includes granular oligopolistic firms. After characterizing the equilibrium of this model economy and outlining a series of counterfactuals of interest, I extend the model (in Subsection 2.9) by adding a continuum of perfectly-competitive atomistic firms.

2.1. Generalized Hedonic-Linear (GHL) Demand

There are n firms, indexed by $i \in \{1, 2, ..., n\}$ that produce differentiated products. Following the tradition of hedonic demand in differentiated product markets (Lancaster, 1966; Rosen, 1974), I assume that consumers value each product as a bundle of characteristics. The number of characteristics is $(n + m)$.

There are two types of characteristics. The first m characteristics are common across all goods and are indexed by $k \in \{1, 2, ..., m\}$, while the remaining n characteristics are idiosyncratic (that is, they are product-specific and cannot be imitated by other products) and therefore have the same index i as the corresponding product. The scalar a_{ki} is the number of units of common characteristic j provided by product i. Each product is described by an m-dimensional, strictly-positive column vector $\mathbf{a}_i$, which I assume to be of unit length – formally:

$$\mathbf{a}_i \;=\; \left[\begin{array}{cccc} a_{1i} & a_{2i} & \dots & a_{mi} \end{array}\right]' \tag{2.1}$$

$$\text{such that} \qquad \sum_{k=1}^{m} a_{ki}^2 = 1 \quad \forall\, i \in \{1, 2, ..., n\} \tag{2.2}$$

The vector $\mathbf{a}_i$ therefore provides firm i's coordinates in the space of common characteristics. We can stack all the coordinate vectors $\mathbf{a}_i$ inside a $m \times n$ matrix that we call $\mathbf{A}$:

$$\mathbf{A} \quad \equiv \quad \left[\begin{array}{cccc} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right] \quad \equiv \quad \left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right] \tag{2.3}$$

Let q_i be the number of units produced by firm i and consumed by the representative agent, which we write inside the n-dimensional vector $\mathbf{q}$:

$$\mathbf{q} \;=\; \left[\begin{array}{cccc} q_1 & q_2 & \cdots & q_n \end{array}\right]' \tag{2.4}$$

Definition 1. A vector $\mathbf{q}$ that specifies, for every firm, the number of units produced is called an *allocation*.

I assume that there exists a representative agent. Consistent with the hedonic demand literature, the consumer combines linearly the characteristics of different products, and their preferences are defined in terms of these characteristics. Letting x_j be the total units of common characteristic j, we have:

$$x_k \;=\; \sum_{i=1}^{n} a_{ki} q_i \tag{2.5}$$

Hence, intuitively, the matrix $\mathbf{A}$ transforms units of goods into units of common characteristics:

$$\mathbf{x} \;=\; \mathbf{A}\mathbf{q} \tag{2.6}$$

Letting y_i be the number of units of each characteristics I assume that each unit of good i provides exactly one unit of its corresponding idiosyncratic characteristic:

$$\mathbf{y} \;=\; \mathbf{q} \tag{2.7}$$

The representative agent's preferences are described by a utility function that is quadratic in both common characteristics ($\mathbf{x}$) and idiosyncratic characteristics ($\mathbf{y}$). The agent's preferences also incorporate a linear disutility for the total number of hours of work supplied (H):

$$U(\mathbf{x},\mathbf{y},H) \quad \stackrel{\text{def}}{=} \quad \alpha \cdot \sum_{k=1}^{m}\left(b_k^x x_k \;-\; \frac{1}{2}\,x_k^2\right) + (1-\alpha)\sum_{i=1}^{n}\left(b_i^y y_i \;-\; \frac{1}{2}\,y_i^2\right) - H \tag{2.8}$$

where b_k^x and b_i^q are characteristic-specific preference shifters. In linear algebra notation:

$$U(\mathbf{x},\mathbf{y},H) \quad \stackrel{\text{def}}{=} \quad \alpha\left(\mathbf{x}'\mathbf{b}^x - \frac{1}{2}\cdot\mathbf{x}'\mathbf{x}\right) + (1-\alpha)\left(\mathbf{y}'\mathbf{b}^y - \frac{1}{2}\cdot\mathbf{y}'\mathbf{y}\right) - H \tag{2.9}$$

$\alpha \in [0,1]$ is the utility weight that is assigned to common characteristics. Hence, it governs the degree of *horizontal differentiation* among products. By making leisure the outside good, I close the model and make it general equilibrium. A similar utility specification was used before by Epple (1987) in an econometric theory setting.[4]

The representative consumer buys the goods bundle $\mathbf{q}$ taking $\mathbf{p}$ (the vector of prices) as given. Moreover, I assume that the representative consumer is endowed with the shares of all the companies in the economy. As a consequence, the aggregate profits are paid back to them. Their consumption basket, defined in terms of the unit purchased $\mathbf{q}$, respects the following budget constraint:

$$H + \Pi \;\geq\; \sum_{i=1}^{n} p_i q_i \tag{2.10}$$

To streamline notation, let us define::

$$\mathbf{b} \quad \stackrel{\text{def}}{=} \quad \alpha\,\mathbf{A}'\mathbf{b}^x + (1-\alpha)\,\mathbf{b}^q \tag{2.11}$$

Then, plugging equation (2.6) and (2.11) inside equation (2.9), we obtain the following Lagrangian for the representative consumer:

$$\mathscr{L}(\mathbf{q},H) \quad = \quad \mathbf{q}'\mathbf{b} - \frac{1}{2}\mathbf{q}'\left[\mathbf{I} + \alpha\left(\mathbf{A}'\mathbf{A} - \mathbf{I}\right)\right]\mathbf{q} - H - \lambda\left(\mathbf{q}'\mathbf{p} - H - \Pi\right) \tag{2.12}$$

The choice of labor hours as the numéraire immediately pins down the Lagrange multiplier $\lambda = 1$. Then, the consumer chooses a demand function $\mathbf{q}(\mathbf{p})$ to maximize the following consumer surplus function:

$$S(\mathbf{q}) \quad \stackrel{\text{def}}{=} \quad \mathbf{q}'\left(\mathbf{b} - \mathbf{p}\right) - \frac{1}{2}\mathbf{q}'\left[\mathbf{I} + \alpha\left(\mathbf{A}'\mathbf{A} - \mathbf{I}\right)\right]\mathbf{q} \tag{2.13}$$

Let us now define the concept of *cosine similarity.*

Definition 2. We call the dot product $\mathbf{a}_i'\mathbf{a}_j$ the *cosine similarity* between i and j.

[4] Epple (1987) uses a utility function that is quadratic-in-characteristics to study the identification properties of hedonic demand systems. There are several differences: 1) Epple's model is partial equilibrium, while this is general equilibrium; 2) Epple's setting is purely theoretical, while this paper is taken to the data; 3) consumer choice is discrete in Epple's model, and continuous in my model; 4) in Epple's model sellers act as price-takers, while in the baseline version of my model sellers oct oligopolistically.

FIGURE 1: ILLUSTRATIVE EXAMPLE - TWO FIRMS, TWO CHARACTERISTICS

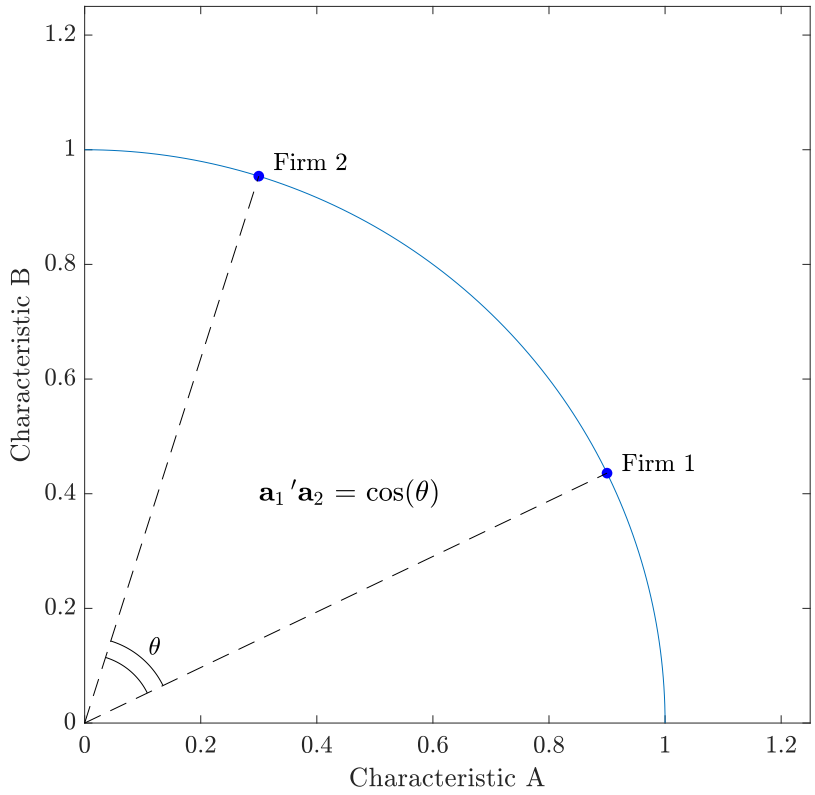


FIGURE NOTES: The following diagram exemplifies the hedonic demand model, for the simple case where there are only two product characteristics (A and B) and only two competitors (1 and 2). Each firm exists as a vector on the unit hypersphere of product characteristics (in this example, we have a circle). The dot product $\mathbf{a}_i'\mathbf{a}_j$ equals the cosine of the angle θ. The tighter the angle, the higher the cosine similarity, and the larger (in absolute value) the inverse cross-price elasticity of demand.

The rationale for this nomenclature is that – geometrically – $\mathbf{a}_i'\mathbf{a}_j$ measures the cosine of the angle between vectors $\mathbf{a}_i$ and $\mathbf{a}_j$ in the space of common characteristics $\mathbb{R}^k$. Hence, the cosine similarity ranges from zero to one. Because, by definition:

$$(\mathbf{A}'\mathbf{A})_{ij} \;=\; \mathbf{a}_i'\mathbf{a}_j \tag{2.14}$$

the matrix $\mathbf{A}'\mathbf{A}$ contains the *cosine similarities* between all firm pairs. A higher cosine similarity implies that two products provide a more overlapping mix of characteristics, and this reflects in patterns of product substitution: if $\mathbf{a}_i'\mathbf{a}_j > \mathbf{a}_i'\mathbf{a}_k$, an increase in the supply of product i leads to a larger decline in the marginal utility of product j than it does on the marginal utility of product k.

Figure 1 helps visualize this setup for the simple case of two firms—1 and 2—competing in the space of two common characteristics A and B. As can be seen in the figure, both firms exist as vectors on the unit circle (with more than three characteristics, it would be a hypersphere instead). The cosine similarity $\mathbf{a}_i'\mathbf{a}_j$ captures the width of the angle θ. An increase in the cosine of the angle θ implies a lower angular distance, and therefore a more overlapping set of common characteristics.

The assumption that $\mathbf{a}_i$ has unit length is a normalization assumption on volumetric units (kilograms, pounds, gallons, etc.). The normalization consists in picking, for each good i, the volume unit so that i is

geometrically represented by a point on the m-dimensional hypersphere.

We can streamline the notation further by defining:

$$\boldsymbol{\Sigma} \overset{\text{def}}{=} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \overset{\text{def}}{=} \alpha \left(\mathbf{A}'\mathbf{A} - \mathbf{I} \right) \tag{2.15}$$

then the demand and inverse demand functions are given by:

$$\text{Aggregate demand}: \quad \mathbf{q} = \left(\mathbf{I} + \boldsymbol{\Sigma} \right)^{-1} \left(\mathbf{b} - \mathbf{p} \right) \tag{2.16}$$

$$\text{Inverse demand}: \quad \mathbf{p} = \mathbf{b} - \left(\mathbf{I} + \boldsymbol{\Sigma} \right) \mathbf{q} \tag{2.17}$$

Notice that the quantity sold by each firm may affect the price of the output sold by every other firm in the economy (unless the matrix $\boldsymbol{\Sigma}$ is null). The derivative $\partial p_i / \partial q_j$ is proportional to $\mathbf{a}_i' \mathbf{a}_j$, the product similarity between i and j. The closer these two firms are in the product characteristics space, the larger is this derivative in absolute value. Because $\mathbf{A}'\mathbf{A}$ is symmetric, we have $\partial q_i / \partial p_j = \partial q_j / \partial p_i$ by construction. In terms of elasticities, we have:

$$\text{Inverse cross}-\text{price demand elasticity}: \quad \frac{\partial \log p_i}{\partial \log q_j} = -\frac{q_j}{p_i} \cdot \sigma_{ij} \qquad \forall \, i \neq j \tag{2.18}$$

$$\text{Cross}-\text{price demand elasticity}: \quad \frac{\partial \log q_i}{\partial \log p_j} = -\frac{p_j}{q_i} \cdot \left(\mathbf{I} + \boldsymbol{\Sigma} \right)^{-1}_{ij} \tag{2.19}$$

It is worth stopping to inspect equation (2.19) more closely. The first thing to notice is that the cross-price demand elasticities depend on the inverse $\left(\mathbf{I} + \boldsymbol{\Sigma} \right)^{-1}$. This implies that, while cosine similarities are positive by construction, it is entirely possible for goods to be complements. This property of the model is discussed at length in Section 6.

Next, let us consider the case $i = j$, where (2.19) simply becomes the own residual demand elasticity. The first major difference between the GHL demand system and CES is that, while in CES the own demand elasticity is equal to a constant, here the own demand elasticity is an equilibrium object (as it depends on $\mathbf{q}$) and will generally differ among firm pairs. This implies that, unlike CES, this demand system produces heterogenous markups. In fact, we can see that two forces drive cross-sectional differences in market power across firms. The more familiar one is the incomplete passthrough from marginal cost to prices: that is, larger firms (high q_i) charge higher markups. The second force, which is instead a feature of hedonic demand models, is asymmetric product differentiation. That is, a firm j that produces a "unique" products, as measured by the term $\left(\mathbf{I} + \boldsymbol{\Sigma} \right)^{-1}_{jj}$, face a less elastic residual demand.

2.2. Supply

I denote by h_i the labor input acquired by every firm. Then the labor market clearing condition is:

$$H = \sum_i h_i \tag{2.20}$$

I assume (without loss of generality) that labor is the numéraire of this economy (the price of one unit of labor is \$1). Therefore h_i is also the total cost incurred by firm i, and it is a function of the output q_i. The marginal cost (also a function of q_i) is c_i:

Special Case: the case where the cost function is quadratic is of particular interest:

$$h_i \;=\; f_i + c_i^0 q_i \;+\; \frac{\delta_i}{2}\, q_i^2 \qquad \text{thus} \qquad c_i \;=\; c_i^0 + \delta_i q_i; \tag{2.21}$$

as it yields closed-form solutions, and it is the one that we take to the data in Section 5.

Firm i maximizes its total profits π_i, defined as follows:

$$\begin{aligned} \pi_i(\mathbf{q}) &\overset{\text{def}}{=} p_i(\mathbf{q}) \cdot q_i - h_i(\mathbf{q}) \\ &= q_i b_i - q_i^2 - \sum_{j \neq i} \sigma_{ij} q_i q_j - h_i \end{aligned}$$

Firms compete à la Cournot: each firm i strategically chooses its output volume q_i by taking as given the output of all other firms. By taking the profit vector as a payoff function and the vector of quantities produced $\mathbf{q}$ as a strategy profile, I have implicitly defined a *network game* (Ballester, Calvó-Armengol and Zenou, 2006, henceforth BCZ). The reason is that the matrix $\mathbf{\Sigma}$ can be conceptualized as the adjacency matrix of a weighted network: in this specific instance, it is the network of product market rivalries that exists among firms, based on the substitutability of their products.

2.3. Equilibrium

Network games belong to a larger class of games known as "potential games" (Monderer and Shapley, 1996): the key feature of potential games is that they can be described by a scalar function $\Phi(\mathbf{q})$, which we call the game's *potential*. The potential function can be thought of, intuitively, as the objective function of the *pseudo-planner* problem that is solved by the Nash equilibrium allocation. The potential function is shown below, together with the aggregate profit function $\Pi(\mathbf{q})$ and the aggregate welfare function $W(\mathbf{q})$:

$$\begin{aligned} \text{Aggregate Profit}: \quad \Pi(\mathbf{q}) &= \mathbf{q}'\mathbf{b} \;-\; \mathbf{q}'(\mathbf{I}+\mathbf{\Sigma})\,\mathbf{q} \;- H(\mathbf{q}) \\ \text{Cournot Potential}: \quad \Phi(\mathbf{q}) &= \mathbf{q}'\mathbf{b} - \frac{1}{2}\cdot\mathbf{q}'(2\mathbf{I}+\mathbf{\Sigma})\,\mathbf{q} - H(\mathbf{q}) \\ \text{Total Surplus}: \quad W(\mathbf{q}) &= \mathbf{q}'\mathbf{b} - \frac{1}{2}\cdot\mathbf{q}'(\mathbf{I}+\mathbf{\Sigma})\,\mathbf{q} - H(\mathbf{q}) \end{aligned} \tag{2.22}$$

The three functions in equation (2.23) are visually similar to each other; they only differ by the scalar weight applied to the quadratic terms. The Cournot potential Φ is somewhat of a hybrid between the aggregate profit Π and the total surplus W: the diagonal entries of the quadratic term are the same as the aggregate profit function, while the off-diagonal terms are the same as the aggregate surplus function. By maximizing the potential $\Phi(\mathbf{q})$, we find the Cournot-Nash equilibrium. I shall assume all these three functions are concave. In the special case where the cost function is quadratic, these three functions are also be quadratic:

$$\begin{aligned} \Pi(\mathbf{q}) &= \mathbf{q}'\left(\mathbf{b}-\mathbf{c}^0\right) \;\; -\frac{1}{2}\cdot\mathbf{q}'(2\mathbf{I}+\mathbf{\Delta}+2\mathbf{\Sigma})\,\mathbf{q} - F \\ \Phi(\mathbf{q}) &= \mathbf{q}'\left(\mathbf{b}-\mathbf{c}^0\right) \;\; -\frac{1}{2}\cdot\mathbf{q}'(2\mathbf{I}+\mathbf{\Delta}+\mathbf{\Sigma})\,\mathbf{q} - F \\ W(\mathbf{q}) &= \mathbf{q}'\left(\mathbf{b}-\mathbf{c}^0\right) \;\; -\frac{1}{2}\cdot\mathbf{q}'(\mathbf{I}+\mathbf{\Delta}+\mathbf{\Sigma})\,\mathbf{q} - F \end{aligned} \tag{2.23}$$

$$\text{where} \qquad \mathbf{\Delta} \overset{\text{def}}{=} \begin{bmatrix} \delta_1 & 0 & \cdots & 0 \\ 0 & \delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_n \end{bmatrix} \qquad \text{and} \qquad F \overset{\text{def}}{=} \sum_{i=1}^{n} f_i \tag{2.24}$$

Because the oligopolists in this model will be actual firms in the data (who produce positive output by definition) we can look directly at the unique internal solution.

Proposition 1. *The Cournot-Nash equilibrium is* $\mathbf{q}^{\Phi}$ *– the maximizer of the potential function* $\Phi(\cdot)$*:*

$$\mathbf{q}^{\Phi} \overset{\text{def}}{=} \arg\max_{\mathbf{q}} \Phi(\mathbf{q}) \tag{2.25}$$

it is the fixed point of the following equation:

$$\mathbf{q}^{\Phi} = (2\mathbf{I} + \mathbf{\Sigma})^{-1} \left[\mathbf{b} - \mathbf{c}\left(\mathbf{q}^{\Phi}\right)\right] \tag{2.26}$$

in particular, if the cost function is quadratic:

$$\mathbf{q}^{\Phi} = (2\mathbf{I} + \mathbf{\Delta} + \mathbf{\Sigma})^{-1} \left(\mathbf{b} - \mathbf{c}^0\right) \tag{2.27}$$

Proof. The derivation of the potential function, as well as the proof that its maximizer $\mathbf{q}^{\Phi}$ is the genuine Nash equilibrium, appear in Appendix A. □

Equation (2.26), which provides a closed-form solution for the case where the cost function is quadratic, allows us to take a closer look at the determinants of equilibrium firm size. The diagonal matrix $\mathbf{\Delta}$, which contains the slopes of the marginal cost functions, captures economies of scale. $\mathbf{\Sigma}$ is the adjacency matrix of the network of product rivalries. $\mathbf{b}$ and $\mathbf{c}^0$ are, respectively, the demand and supply function intercepts. Hence, $(b_i - c_i)$ is simply the marginal surplus of the very first unit produced by firm i; also, b_i can be interpreted as a measure of vertical product differentiation (quality).

BCZ show that another way to interpret equation (2.26) is as a measure of *network centrality* – specifically, that developed by Katz (1953) and Bonacich (1987). The intuition is that firms that are more "isolated" in the network of product similarities face less product market competition and behave more like monopolists. Centrality measures are a recurring feature of the literature on networks in macroeconomics (see Carvalho and Tahbaz-Salehi, 2019). In Appendix C, I discuss in further detail the link between Nash equilibrium and network centrality.

The discrepancy between the potential function and the total-surplus function implies that the network Cournot game delivers an equilibrium allocation that is not socially-optimal. A benevolent social planner can theoretically improve on the market outcome for two reasons. First, they can coordinate output choices across firms; second, they can internalize consumer surplus.

2.4. Generality of the Utility Specification and the role of Idiosyncratic Characteristics

In Appendix B I prove that the utility specification in equation (2.9) is identical, up to series of welfare-invariant normalizations, to the much more general form

$$U(\mathbf{x}, \mathbf{y}, H) \overset{\text{def}}{=} \tau_1 \mathbf{x}'\mathbf{b}^x - \frac{\tau_2}{2} \cdot \mathbf{x}'\mathbf{M}^x\mathbf{x} + \frac{\tau_3}{2}\mathbf{y}'\mathbf{b}^y - \frac{\tau_4}{2}\mathbf{y}'\mathbf{M}^y\mathbf{y} - \tau_5 H \tag{2.28}$$

$$\mathbf{x} = \mathbf{A}^x\mathbf{q} \qquad \text{and} \qquad \mathbf{y} = \mathbf{A}^y\mathbf{q} \tag{2.29}$$

where $\mathbf{M}^x$ is a diagonalizable (but not necessarily diagonal) matrix, $\mathbf{M}^y$ and $\mathbf{A}^y$ are diagonal (but not necessarily identity) matrices, and we do *not* require $\|\mathbf{a}_k^x\| = 1$ for all k. In addition we *do* require – only

when $\alpha < 1$ – that the columns of $\left(\frac{\tau_2}{\alpha}\mathbf{M}^x\right)^{-\frac{1}{2}}\mathbf{A}^x$ have the same euclidian norms as the respective columns of the matrix $\left(\frac{\tau_4}{1-\alpha}\mathbf{M}^y\right)^{-\frac{1}{2}}\mathbf{A}^y$. This assumption simply means that all goods have the same relative loading on common v/s idiosyncratic characteristics – i.e., we only need a single α parameter to capture the weight on common characteristics (as opposed to a different α for each good); while we could relax this assumption in theory, we would still require it to take the model to the data, and it simplifies the exposition significantly.

None of the normalizations required to go from (2.28) to (2.9) involves choosing the labor units, which therefore leaves us with an additional degree of freedom to impose labor as the numeráire good.

2.5. Markups, Productivity and Centrality

In this model, the ability of firms to price above marginal cost depends on two factors. First, firms that are more "productive" (that can generate a lot of surplus surplus with little labor) become larger in the equilibrium: because demand elasticity decreases with size, these firms end up charging higher markups. Second, firms that are more central in the network of product similarities face more intense product market competition, and thus a more elastic residual demand: these firms charge lower markups. In this sub-section, I formalize this intuition by defining some measures of centrality and productivity as well by deriving a few useful identities. To begin with, we formally define a measure of *degree centrality.*

Definition 3. We define d_i –the (output-weighted) *degree centrality* of firm i– as follows:

$$d_i \overset{\text{def}}{=} \sum_{j \neq i} \sigma_{ij} q_j \tag{2.30}$$

We can then express the equilibrium quantity and price-cost margins as a decreasing function of this measure of centrality.

Proposition 2. *The equilibrium quantities and price-cost margins satisfy*:

$$q_i \quad \equiv \quad p_i - c_i \quad \equiv \quad \frac{b_i - d_i - c_i}{2} \tag{2.31}$$

Proof. Appendix G. □

The intuition behind this expression is that b_i is the intercept of the "monopolistic" demand function that firm i would face if it had zero similarity to every other firm, while $(b_i - d_i)$ is the intercept of the residual demand actually faced by i in the Cournot oligopoly game.

If i had zero degree centrality, it would charge a monopolistic price-cost margin equal to $\frac{1}{2}(b_i - c_i)$. As competition from rivals increases (higher d_i) firm i charges lower and lower margins, and its equilibrium size decreases. As d_i approaches $(b_i - c_i)$, firm i eventually hits a choke price and exits endogenously.

One shortcoming of d_i as a measure of centrality is that it is an endogenous object (it depends on the output vector $\mathbf{q}$). The next step is to write the equilibrium markup of firm i in terms of exogenous measures of productivity and centrality. We start by defining the equilibrium markup formally.

Definition 4. We define μ_i –the markup of firm i– as the ratio between the output price p_i and the marginal cost (c_i) – formally:

$$\mu_i \overset{\text{def}}{=} \frac{p_i}{c_i} \tag{2.32}$$

The first step is to define a novel measure of productivity that is comparable in the cross section of firms and that accounts for product quality. This is a non-trivial task. We know that cross-sectional comparisons of physical productivity are typically meaningless (Syverson, 2004) for a variety of reasons: a) the output of

different firms is typically measured in volumetric units that are not comparable; b) even when output units are comparable, the production technology may differ across firms; c) even keeping technology constant, the quality of output may vary. In fact hedonic adjustment is increasingly used in national statistics to construct good price indices that account for changes in quality. These adjustments have been shown to exert a significant effect on measured productivity growth (Moulton et al., 2001).

To construct such an "ideal" productivity measures, we start from the observation that the model admits a clear measure of quality, b_i, which is the representative agent's willingness to pay for the first unit of good i when the supply of every other product is zero. The next step is to exploit the fact that a change of volumetric units (say, from pounds to kilograms) has the effect of scaling up b_i and c_i by exactly the same factor. By taking the ratio between b_i and c_i, we obtain a measure of productivity that adjusts for product quality, is welfare-relevant (see below), and is invariant to changes in volumetric units. We call this ratio "hedonic-adjusted productivity".

Definition 5. We define ω_i –the "hedonic-adjusted productivity" of firm i– as the ratio between the marginal utility of the very first unit produced (b_i) and the marginal cost (MC_i) – formally:

$$\omega_i \stackrel{\text{def}}{=} \frac{b_i}{c_i} \tag{2.33}$$

It is easy to prove that the markup that firm i charges in equilibrium (μ_i) is bounded above by its quality-adjusted productivity.

Lemma 1. *In the Nash-Cournot equilibrium allocation, firm i's equilibrium markup is always less than the "monopolistic" markup $\bar{\mu}_i$, which takes on the following expression:*

$$\mu_i \quad \leq \quad \bar{\mu_i} \quad \stackrel{\text{def}}{=} \quad \frac{b_i + c_i}{2c_i} \quad \equiv \quad \frac{1+\omega_i}{2} \tag{2.34}$$

with equality if and only if firm i has degree centrality equal to zero ($d_i = 0$).

Proof. Appendix (G). □

To write the equilibrium markup in terms of productivity and centrality, let us re-write equation (2.26) in terms of the matrix $\mathbf{\Gamma}$:

$$\mathbf{q}^{\Phi} \;=\; \frac{1}{2} \cdot \mathbf{\Gamma} \left(\mathbf{b} - \mathbf{c} \right) \qquad \text{where} \qquad \mathbf{\Gamma} \;\equiv\; \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{bmatrix} \stackrel{\text{def}}{=} \left(\mathbf{I} + \frac{1}{2}\mathbf{\Sigma} \right)^{-1} \tag{2.35}$$

We can further rewrite equation (2.35) as:

$$q_i^{\Phi} \;=\; \frac{1}{2} \left[\gamma_{ii} + \sum_{j \neq i} \gamma_{ij} \frac{b_j - c_j}{b_i - c_i} \right] (b_i - c_i) \tag{2.36}$$

Appealing to the Nash equilibrium-centrality linkage, we can interpret the term in square brackets as a measure of (inverse) centrality, that captures how "far" firm i is from every other firm j, and weights each rival j by its competitiveness ($b_j - c_j$) relative to i. We can thus formally define the product market centrality of firm i.

Definition 6. We define χ_i, the *product market centrality* of firm i as follows:

$$1 - \chi_i \quad \stackrel{\text{def}}{=} \quad \gamma_{ii} + \sum_{j \neq i} \gamma_{ij} \frac{b_j - c_j}{b_i - c_i} \tag{2.37}$$

It is a measure of centrality because it "summarizes" the entire matrix of cross-price derivatives into an n-dimensional vector. What we have done, intuitively, is to replace $\left(\mathbf{I} + \frac{1}{2}\mathbf{\Sigma}\right)^{-1}$ with a diagonal matrix that has $(1 - \chi_i)$ along the diagonal, to obtain:

$$q_i^{\Phi} \;=\; \frac{1 - \chi_i}{2} \left(b_i - c_i\right) \tag{2.38}$$

Importantly, this measure of centrality only depends on exogenous objects, and it is the one that determines how close to competitive (or monopolistic) is the markup charged by firm i in equilibrium.

Proposition 3. *The equilibrium markup μ_i is equal to a product market centrality (χ_i)-weighted convex combination of 1 (the lowest possible markup), and the monopolistic markup $\bar{\mu}_i$:*

$$\mu_i \;=\; \chi_i + (1 - \chi_i)\, \bar{\mu}_i \tag{2.39}$$

Corollary. *The product market centrality ranges from zero to one: $\chi_i \in [0, 1]$.*

Proof. Appendix G. □

This proposition links the topology of the rivalry network to the residual demand elasticity faced by individual firms: a firm that is highly central ($\chi_i \to 1$) has many competitive rivals that produce products with similar characteristic to its own, and thus behaves similarly to an atomistic firm. Vice-versa, a firm that is highly peripheral ($\chi_i \to 0$) behaves like a monopolist.

2.6. Consumer Surplus, Shapley Value and Surplus Appropriation

The surplus that firms produce is either appropriated by firms in the form of profits, or by consumers in the form of consumer surplus. It is thus natural to ask the following question: how much (consumer) surplus does each firm contribute? And how does oligopoly power affect firm's ability to to appropriate surplus? These might seem like questions that are not well posed: after all, the consumer surplus contributed by firm i depends on how much output every other firm j is producing.

Luckily, the problem of how to attribute surplus to players in a game with non-linear utilities has already been studied in the theory literature, and we know that there is a natural, economically-meaningful metric that accomplishes this objective: the Shapley Value. While the Shapley Value is usually utilized to break down total surplus in coalitional games, there is nothing that prevents us from applying the same concept to consumer surplus in a game of oligopoly.

To break down aggregate consumer surplus using the Shapley value, we start by writing down the expression for the consumer surplus generated by q_i units of good i when all other firms supply $\bar{\mathbf{q}}_{-i}$ (taken as given). We know from basic price theory that this quantity can be computed by integrating the difference between the residual demand and the purchase price p_i[5]:

$$\int_0^{q_i} \left(b_i - q_i' - \sum_{j \neq i} \sigma_{ij} \bar{q}_j - p_i \right) \mathrm{d}q_i' \qquad = \qquad q_i \left(b_i - p_i\right) - \frac{1}{2} q_i^2 - \sum_{j \neq i} \sigma_{ij} q_i \bar{q}_j \tag{2.40}$$

[5]When integrating consumer surplus, we must remember to treat p_i as a constant, since consumers are price-takers.

Naturally, the marginal consumer surplus generated by i depends on the vector of quantities supplied by every other competitor ($\bar{\mathbf{q}}_{-i}$). The Shapley Value in terms of consumer surplus, which we call s_i, is just the average of (2.40), taken over the set of allocations where firms $j \neq i$ produce $\bar{q}_j = q_j$ or $\bar{q}_j = 0$:[6]

$$s_i \quad \stackrel{\text{def}}{=} \quad \frac{1}{2^{(n-1)}} \sum_{\iota_1=0}^{1} \sum_{\iota_2=0}^{1} \cdots \sum_{\iota_n=0}^{1} \left[q_i \left(b_i - p_i \right) - \frac{1}{2} q_i^2 - \sum_{j \neq i} \iota_j \sigma_{ij} q_i q_j \right] \tag{2.41}$$

because we know that each firm j produces q_j in exactly half of the allocations considered, the expression above simplifies to:

$$s_i \;=\; q_i \left(b_i - p_i \right) - \frac{1}{2} \left(q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i q_j \right) \tag{2.42}$$

by plugging p_i (as a function of q_i), this expression further simplifies to:

$$s_i \;\stackrel{\text{def}}{=}\; \frac{1}{2} \left(q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i q_j \right) \tag{2.43}$$

Because it is a Shapley Value, one of s_i's desirable properties is that it always aggregates to total consumer surplus, that is:

$$S \quad = \quad \sum_i s_i \tag{2.44}$$

We shall therefore call s_i, going forward, the consumer surplus generated by firm i. We can similarly define a firm-level *total* surplus function, which specifies for every firm i a certain share w_i of total surplus $W(\mathbf{q})$:

$$w_i \quad \stackrel{\text{def}}{=} \quad \pi_i + s_i \quad = \quad q_i b_i - \frac{1}{2} \left(q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i q_j \right) - h_i \tag{2.45}$$

Next, we define a suitable measure of market share for our network oligopoly model, and derive an equation that links this quantity to firm's ability to appropriate surplus in the form of monopoly profits.

Definition 7. I define $\mathcal{M}_i$, the *weighted market share* of firm i, as follows:

$$\mathcal{M}_i \quad \stackrel{\text{def}}{=} \quad \frac{q_i}{q_i + \sum_j \sigma_{ij} q_j} \tag{2.46}$$

This measure of market share weights the output of each competitor by σ_{ij}, a measure of substitutability between i's product and j's product. Notice that, under homogenous products ($\sigma_{ij} = 1 \; \forall \, i, j$) this is simply the market share of firm i.

Following the literature, we can formally define r_i – the monopoly rents of firm i – as:

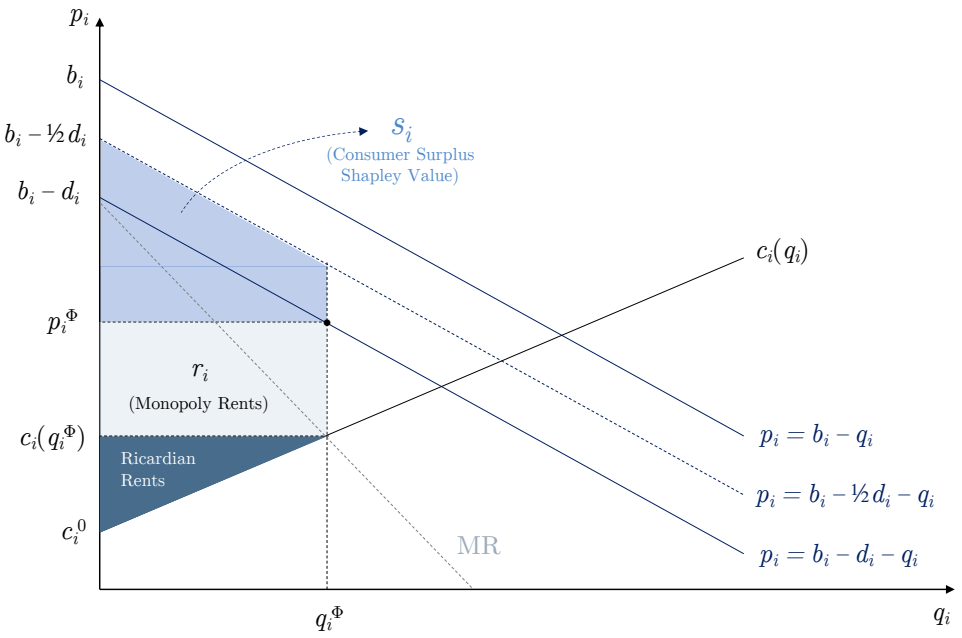
$$r_i \;\stackrel{\text{def}}{=}\; \left(p_i - c_i \right) q_i \tag{2.47}$$

and the Ricardian Rents as the difference between $c_i q_i$ and variable costs[7] (positive if marginal cost is

[6] We could have equivalently taken the average over the continuous set $\bar{q}_j \in [0, q_j]$, which yields the same expression for s_i (due of linearity). I used the discrete average where $\bar{q}_j \in \{0, q_j\}$ because this is how the Shapley Value is usually defined.

[7] Thanks to David Baqaee for providing the correct nomenclature for all these objects.

FIGURE 2: DEMAND, RENTS AND CONSUMER SURPLUS



increasing), so that profits can be decomposed into:

$$\pi_i \quad = \quad \underbrace{r_i}_{\text{Monopoly Rents}} \quad + \quad \underbrace{c_i q_i - \text{TVC}_i}_{\text{Ricardian Rents}} - \underbrace{f_i}_{\text{Fixed Costs}} \tag{2.48}$$

It is possible to show that the ratio of firm profits π_i to consumer surplus (measured using the Shapley Value s_i) is proportional to the weighted market share $\mathcal{M}_i$.

Proposition 4. *In the Cournot-Nash equilibrium allocation, the ratio of monopoly rents to consumer surplus, for firm i, is equal to twice the weighted market share - specifically:*

$$\frac{r_i}{s_i} \quad = \quad 2\,\mathcal{M}_i \tag{2.49}$$

Proof. See Appendix G. □

Proposition 4 reflects the fact that, in my model, there are no clearly-defined industry boundaries. This is also the case in the real world: if we consider antitrust lawsuits for example, a major object of litigation is the market's definition. Consider for example an attempted monopolization case: defendants (alleged monopolies) have an incentive to define the relevant market broadly, while plaintiffs have an incentive to define the relevant market narrowly. In my model, firms exist in a continuous space of product characteristics. Hence, there is no uniquely-defined peer group for each firm. To understand how dominant firm i is, we need to compare its market share vis-à-vis every other firm in the economy, weighting every other firm by σ_{ij}. Notice that it is perfectly possible, in this model, for every firm to have a weighted market share of 100%. This corresponds to the case where the network is completely disconnected ($\sigma_{ij} = 0$ for all i, j) and every firm is a monopolist.

While the weighted market share is an endogenous equilibrium object, we can show that it is entirely pinned down by the procuct market centrality χ_i, of which it is a strictly decreasing function. This implies that also

the ratio r_i/s_i is also a decreasing function of centrality.

Proposition 5. *The equilibrium weighted market share $\mathcal{M}_i$ is equal to the following decreasing function of χ_i:*

$$\mathcal{M}_i \quad = \quad \frac{1-\chi_i}{1+\chi_i} \tag{2.50}$$

This implies that when χ_i is equal to zero $\mathcal{M}_i$ is equal to one, and vice-versa. In sum, χ_i is thus the key statistic of market power in this model: markups, surplus sharing and (weighted) market shares all depend ultimately on it.

2.7. Efficiency and Counterfactuals

A key application of my theoretical model is to study how welfare statistics - such as total surplus - respond to changes in market structure. What that means is that, having made the required assumption that firms compete by a well-defined set of rules (thus far we have assumed Cournot oligopoly), we can then consider counterfactuals in which the same firms play by a different set of rules. In this Subsection, I define four of these counterfactuals: each of these counterfactuals corresponds to the solution of a specific maximization problem.[8]

The first counterfactual that I consider is *perfect competition*: firms act as atomistic producers, and price all units sold at marginal cost.

Definition 8. The *Perfect Competition* allocation $\mathbf{q}^W$ is defined as the maximizer of the aggregate total surplus function $W(\mathbf{q})$:

$$\mathbf{q}^W \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}\geq 0} W(\mathbf{q}) \tag{2.51}$$

for $\mathbf{q}^W > \mathbf{0}$ it satisfies

$$\mathbf{q}^W \quad = \quad (\mathbf{I}+\boldsymbol{\Sigma})^{-1}\left[\mathbf{b}-\mathbf{c}\left(\mathbf{q}^W\right)\right] \tag{2.52}$$

and with quadratic cost function it equals

$$\mathbf{q}^W \quad = \quad (\mathbf{I}+\boldsymbol{\Delta}+\boldsymbol{\Sigma})^{-1}\left(\mathbf{b}-\mathbf{c}^0\right) \tag{2.53}$$

The second counterfactual that I consider is called *Monopoly*: it represents a situation in which one agent (that does not internalize consumer surplus) has ownership and control over all firms and maximizes aggregate profits.

Definition 9. The *Monopoly* allocation is defined as the maximizer of the aggregate profit function $\Pi(\mathbf{q})$:

$$\mathbf{q}^\Pi \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}\geq 0} \Pi(\mathbf{q}) \tag{2.54}$$

for $\mathbf{q}^\Pi > \mathbf{0}$ it satisfies

$$\mathbf{q}^\Pi \quad = \quad \frac{1}{2}(\mathbf{I}+\boldsymbol{\Sigma})^{-1}\left[\mathbf{b}-\mathbf{c}\left(\mathbf{q}^\Pi\right)\right] \tag{2.55}$$

and with quadratic cost function it equals

$$\mathbf{q}^\Pi \quad = \quad (2\mathbf{I}+\boldsymbol{\Delta}+2\boldsymbol{\Sigma})^{-1}\left(\mathbf{b}-\mathbf{c}^0\right) \tag{2.56}$$

[8] The closed-form expressions for the output vector $\mathbf{q}$ which I provide below assume an internal solution. For my empirical analysis, I also compute a numerical solution that is subject to a non-negativity constraint on $\mathbf{q}$ and I verify it is approximately equal to the unconstrained solution (error $< 0.1\%$ for the total surplus function in Perfect Competition). The non-negativity constraint binds for very few firms.

This allocation can be thought of as a stylized representation an economy with no antitrust, where firms have unlimited ability to coordinate their supply choices.

Another interesting counterfactual is one in which resources are allocated efficiently but the labor supply is fixed. That is, the social planner maximizes the aggregate surplus function subject to the constraint of using no more labor than in the observed Cournot equilibrium.

Definition 10. I define the resource-efficient counterfactual $\mathbf{q}^H$ as the solution to the following constrained maximization problem:

$$\mathbf{q}^H \quad \stackrel{\text{def}}{=} \quad \arg\max_{\mathbf{q}\geq 0} W(\mathbf{q}) \quad \text{s.t.} \quad H(\mathbf{q}) = H\left(\mathbf{q}^\Phi\right) \tag{2.57}$$

Setting up the Lagrangian and using $(1-\mu)$ as the Lagrange multiplier, we find that, conditioning on the set of active firms ($\mathbf{q} > \mathbf{0}$) the resource-efficient counterfactual takes the form:

$$\mathbf{q}^H \;=\; (\mathbf{I}+\mathbf{\Sigma})^{-1}(\mathbf{b}-\mu\mathbf{c}) \quad \text{if} \quad \mathbf{q}^H > \mathbf{0} \tag{2.58}$$

where μ solves:

$$H\left(\mathbf{q}^H(\mu)\right) \;=\; H\left(\mathbf{q}^\Phi\right) \tag{2.59}$$

The Lagrange multiplier term μ turns out to be the common markup charged by all firms in the resource-efficient counterfactual.

Proposition 6. *The Resource-efficient counterfactual $\mathbf{q}^H$ equalizes markups across active firms.*

Proof. Let all firms price at a constant markup μ over marginal cost:

$$p_i \;=\; \mu c_i \tag{2.60}$$

expanding the expression for the equilibrium price we have:

$$\mathbf{b} - (\mathbf{I}+\mathbf{\Sigma})\,\mathbf{q} \;=\; \mu\mathbf{c} \tag{2.61}$$

rearranging the equation above we obtain (2.58). □

Because this counterfactual uses the same amount of labor as the observed equilibrium, by comparing welfare in this allocation to the first-best we can effectively break down the deadweight loss into two components – one linked to misallocation, the other linked to labor suppression. We can also interpret this counterfactual as the deadweight loss in an alternative model where the supply of labor is completely inelastic. Notice that when this allocation is not constrained by the labor supply (the Lagrange multiplier $1-\mu$ is zero), the common markup is one (firms price at marginal cost) and the resource-efficient allocation coincides with perfect competition.

2.8. Multi-product Firms, Collusion and M&A

Next, I generalize the model to accommodate multi-product firms and show how to perform counterfactuals where firm boundaries are altered.

Suppose now that i indicates product lines, and firms are denoted by $z = 1, 2, ..., Z$. We thus define an $n \times Z$ *ownership matrix* $\mathbf{O}$, whose (i,z) entry is equal to one if firm z owns product line i. Each firm z maximizes ϖ_z - the sum of the profits from all product lines:

$$\varpi_z \;=\; \sum_{i=1}^{n} o_{iz}\pi_i \tag{2.62}$$

We next derive the equilibrium of the multi-product Cournot model:

Proposition 7. *The multi-product Nash-Cournot equilibrium quantity vector maximizes the following modified potential function:*

$$\Phi(\mathbf{q}) \;=\; \mathbf{q}'\mathbf{b} - \frac{1}{2}\mathbf{q}'\left(2\mathbf{I} + \mathbf{\Sigma} + \mathbf{K} \circ \mathbf{\Sigma}\right)\mathbf{q} - H(\mathbf{q}) \tag{2.63}$$

where $\mathbf{K}$ *is the co-ownership matrix, defined as follows:*

$$\mathbf{K} \;\equiv\; \begin{bmatrix} \kappa_{11} & \kappa_{21} & \cdots & \kappa_{1n} \\ \kappa_{12} & \kappa_{22} & \cdots & \kappa_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \kappa_{n1} & \kappa_{n2} & \cdots & \kappa_{nn} \end{bmatrix} \;\overset{\text{def}}{=}\; \mathbf{O}'\mathbf{O} \tag{2.64}$$

and the operator $(\circ)$ *is the Hadamard (entry-by-entry) product. The solution satisfies*

$$\mathbf{q}^{\Phi} \;=\; \left(2\mathbf{I} + \mathbf{\Sigma} + \mathbf{K} \circ \mathbf{\Sigma}\right)^{-1}\left[\mathbf{b} - \mathbf{c}\left(\mathbf{q}^{\Phi}\right)\right] \tag{2.65}$$

and with quadratic cost function it is equal to:

$$\mathbf{q}^{\Phi} \;=\; \left(2\mathbf{I} + \mathbf{\Delta} + \mathbf{\Sigma} + \mathbf{K} \circ \mathbf{\Sigma}\right)^{-1}\left(\mathbf{b} - \mathbf{c}^{0}\right) \tag{2.66}$$

Proof. Appendix G. □

The symmetric matrix $\mathbf{K}$ has a simple structure and a straightforward economic interpretation: its (i, j) entry is equal to one if product lines i and j are owned by the same firm, and zero otherwise.

The multi-product extension of the model can be used to perform counterfactuals on firm boundaries, including mergers, acquisitions, break-ups and divestures. When it comes to modeling mergers and collusion, the I.O. literature has used multiple approaches. Following Baker and Bresnahan (1985), we can model mergers and collusion interchangeably as *coordinated pricing.* That is, the merger or the collusion does not alter the product range offered by the merging/colluding enterprises; instead, a single agent determines the output of the merging firms to maximize the joint profits. Then, even in a single-product setting ($\kappa_{ij} = 0 \;\forall\; i, j$), we can simulate a merger or a collusion between a subset $\mathcal{J}$ of the set of firms by re-setting to one the diagonal entries of $\mathbf{K}$ that correspond to the elements of $\mathcal{J} \times \mathcal{J}$. It is easily verified that when all firms are merged ($\kappa_{ij} = 1 \;\forall\; i, j$), the potential function $\Phi(\mathbf{q})$ converges to the aggregate profit function $\Pi(\mathbf{q})$, and the equilibrium allocation converges to the Monopoly counterfactual (equation 2.54).

More in general, to simulate a more complex counterfactual, all we need to do is to update the matrix $\mathbf{K}$ to reflect the updated firm boundaries. Consider the following illustrative example.

Example. Firm z_1 produces products 1 and 2, while firm z_2 produces product 3. This implies that $\kappa_{12} = \kappa_{21} = 1$ and $\kappa_{13} = \kappa_{31} = \kappa_{23} = \kappa_{23} = 0$. If firm z_1 sells product line 2 to firm z_2, the matrix $\mathbf{K}$ must to be updated so that $\kappa_{12} = \kappa_{12} = \kappa_{13} = \kappa_{31} = 0$ and $\kappa_{23} = \kappa_{23} = 1$.

An interesting question that we can ask is whether there is a way to identify, ex-ante, mergers that are unlikely to produce significant consumer welfare harm. The FTC-DOJ horizontal merger highlight the pre- and post-merger Herfindahl Indices in the relevant industry. It is possible to formally prove that in this model, where there are no formal industry boundaries, cosine similarities can fulfill a similar role.

Proposition 8. *The equilibrium allocation is unaffected by co-ownership of production lines with low cosine similarity. Formally, consider two identical economies that only differ by their co-ownership matrices* $\mathbf{K}^{(1)}$ *and* $\mathbf{K}^{(2)}$*. Let* $\mathbf{q}^{(1)}$ *and* $\mathbf{q}^{(2)}$ *be their respective Cournot-Nash equilibria. If*

$$\mathbf{a}_i'\mathbf{a}_j \;\approx\; 0 \,\text{for all}\; (i, j) \;\text{such that}\; \kappa_{ij}^{(1)} \neq \kappa_{ij}^{(2)} \tag{2.67}$$

then $\mathbf{q}^{(1)} \approx \mathbf{q}^{(2)}$.

Proof. Follows trivially from the fact that κ_{ij} only appears in the equilibrium allocation (2.65) multiplied by σ_{ij}, which is proportional to the cosine similarity $\mathbf{a}_i'\mathbf{a}_j$. □

2.9. Adding a Continuum of firms with Endogenous Entry and Exit

Next, I show how to expand the model to include a continuum of atomistic firms that behave competitively and can enter and exit endogenously. This extension of the model allows me to accomplish two things: 1) incorporate firms for whom we do not observe product similarity data – that is, foreign and private firms; 2) it allows to incorporate entry and exit in an otherwise static model. The idea is that we can model unobserved companies as atomistic firms.

The key to tractably integrating these atomistic firms in the model is an aggregation result. I describe these atomistic firms through a productivity distribution: the set of active atomistic firms is then characterized by a productivity cut-off value, in the style of Hopenhayn (1992).

Next, I show that these atomistic companies can be aggregated into a representative firm: variations in the size of the representative firm reflect the intensive margin of production as well as the extensive margin (the entry/exit of the atomistic firms). I index this representative firm $i = n + 1$, effectively adding a row and a column to the matrices $\mathbf{A}'\mathbf{A}$ and $\mathbf{\Delta}$ and adding one dimension to the vector $\mathbf{b}$.

Proposition 9. *Assume that there is a continuum of potential entrants that are indexed by a productivity parameter* $\zeta \in \left(\underline{\zeta}, \infty\right)$*, with* $\underline{\zeta} > 0$*, and that produce a homogeneous good using the following quadratic cost function:*

$$h\left(\zeta\right) \;=\; \frac{1}{2\zeta} \cdot q^2\left(\zeta\right) \tag{2.68}$$

Assume also that the firms face cost of entry equal to one unit of labor and that the probability density of type-ζ *potential entrants is given by*

$$f\left(\zeta\right) \;= \frac{\beta - 1}{\zeta^{\beta+1}} \tag{2.69}$$

implying that ζ *follows a Pareto distribution with shape parameter* β *and scale parameter* $\underline{\zeta} \overset{\text{def}}{=} \left[\left(\beta - 1\right)/\beta\right]^{\frac{1}{\beta}}$.[9] *Then, as the parameter* β *converges down to* 1*, the cost function of the corresponding aggregate representative firm is approximated by*

$$h_{n+1} \;=\; \frac{q_{n+1}^2}{2} \tag{2.70}$$

where and h_{n+1} *and* q_{n+1} *are, respectively, the labor input and the output of the representative firm, and the productivity cutoff for entry converges to* $\zeta_{\min} = \frac{1}{q_{n+1}}$.

Proof. See Appendix G. □

Because employment and revenues are proportional to ζ, it follows that, if the assumptions above are respected, both the revenue and employment distribution of firms also approximate a Pareto distribution with shape parameter $\beta = 1$, sometimes called a Zipf Law.

Although this might look like a knife-edge assumption, it is not. It is a well-documented empirical regularity that the size distribution of firms closely approximates a Pareto distribution with shape parameter $\beta = 1$.

[9] While the revenue and employment distribution of US firms approximates a Pareto Distribution with scale parameter equal to one (a Zipf Law), this distribution has the undesirable property that its mean (and therefore q_{n+1} and h_{n+1}) grows unboundedly as $\beta \to 1^+$. This particular choice of the scale parameter ensures that q_{n+1} and h_{n+1} integrate to a finite number as $\beta \to 1^+$.

This stylized fact was confirmed to hold for both the employment *and* the revenue distribution of US firms by Axtell (2001), using Census micro-data.

Next, I show how to add this representative firm to the network oligopoly model. Because the representative firm behaves competitively, its first order condition will differ from that of granular firms $\{1, 2, ..., n\}$. The latter maximize individual profits:

$$\pi_i'(q_i) \;=\; 0 \qquad \text{for} \quad i = 1, 2, ..., n \tag{2.71}$$

The representative firm, on the other hand, prices at marginal cost, and therefore maximizes total surplus:

$$W'(q_i) \;=\; 0 \qquad \text{for} \quad i = n+1 \tag{2.72}$$

We can write the full system of first order conditions in linear algebra notation as:

$$0 \;=\; \begin{bmatrix} \mathbf{b}_{(n)} - \mathbf{c}^0_{(n)} \\ b_{n+1} - c^0_{n+1} \end{bmatrix} + \left(\begin{bmatrix} 2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} + \mathbf{\Sigma} + \mathbf{\Delta} \right) \begin{bmatrix} \mathbf{q}_{(n)} \\ q_{n+1} \end{bmatrix} \tag{2.73}$$

where $c^0_{n+1} = 0$, $\delta_{n+1} = 1$ and the superscript (n) identifies the sub-vector corresponding to the granular firms. A simpler way to rewrite this set of equations is

$$0 \;=\; \mathbf{b} - \mathbf{c} - (\mathbf{I} + \mathbf{G} + \mathbf{\Sigma} + \mathbf{\Delta})\,\mathbf{q} \tag{2.74}$$

where $\mathbf{G}$ is a diagonal matrix that identifies granular firms – that is, whose diagonal elements equal 1 for firms 1 to n and to 0 for firm $n+1$:

$$\mathbf{G} \;=\; \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{2.75}$$

The potential function for the model that includes the representative firm is:

$$\Phi(\mathbf{q}) \;=\; \mathbf{q}'(\mathbf{b} - \mathbf{c}) - \frac{1}{2}\mathbf{q}'(\mathbf{I} + \mathbf{G} + \mathbf{\Sigma} + \mathbf{\Delta})\,\mathbf{q} \tag{2.76}$$

and the equilibrium quantity vector is:

$$\mathbf{q}^{\Phi} \;=\; (\mathbf{I} + \mathbf{G} + \mathbf{\Delta} + \mathbf{\Sigma})^{-1}\left(\mathbf{b} - \mathbf{c}^0\right) \tag{2.77}$$

3. Data, Identification and Validation

In this section, I outline the data used to estimate the model in Section 2 and how it maps to the various objects that make up the empirical model. The mapping and the identification are summarized in Table 1.

3.1. Cost Function

In order to take the model to the data, we need to make a parametric assumption about the cost function. I assume that it follows the quadratic form from equation 2.21. In the baseline model, I further assume that firms face and exogenous (flat) marginal cost ($\delta_i = 0$). In the extended model (section 6.1), I relax this assumption and outline a procedure to identify the slope term $\delta_i = 0$.

3.2. Firm Financials

My data source for firm financials is the Compustat database, which I access via the Wharton Research Data Services (WRDS) platform. From this database, I extract information on firm revenues, Costs of Goods Sold (COGS), Selling General and Administrative (SGA) costs.

I follow (De Loecker, Eeckhout and Unger, 2020, henceforth DEU) in mapping accounting revenues to model revenues, COGS to variable costs, and SG&A to fixed costs (f_i). In addition, in order to make the welfare metrics comparable over time, I deflate every dollar amount by the hourly labor compensation in non-farm business (COMPNFB in FRED), indexed at the level of 2019.

There are various reasons why hourly labor compensation (as opposed to, say, CPI) is the correct deflator. First, goods prices incorporate welfare-relevant information about goods quality. Second (by assumption) labor is the only commodity that is not subject to quality changes and that is supplied at a constant marginal utility cost. Third, it is the numeraire good. Lastly, it provides an intuitive interpretation for the deflated dollars figures in terms of units of leisure.

3.3. Text-Based Product Similarity

The key data ingredient that we need, in order to estimate my model, is the matrix of product similarities $\mathbf{A}'\mathbf{A}$. The empirical counterpart of this object is provided by Hoberg and Phillips (2016, henceforth HP).

HP created a publicly-available database that provides product cosine similarities for the universe of public corporations in the United States. These cosine similarities originate from natural language processing (NLP) of 10-K filings, and are time-varying. A complete matrix of similarities is provided for every year, beginning in 1996. The 10-K is a regulatory form that must be filed by American public corporations with the U.S. Securities and Exchange Commission on a yearly basis. Item 1 of the 10-K is a long and detailed description of the product or service sold by the company. HP's product cosine similarities are constructed by comparing these textual product descriptions.

I briefly outline the construction of this dataset. HP start by building a vocabulary of 61,146 words that firms use to describe the characteristics of their products. Let us call the set of words comprising this vocabulary $\mathcal{V} = \{1, 2, ..., 61146\}$.[10]

[10] I report here verbatim the methodology description from the original paper by Hoberg and Phillips (2016):"[...] *In our main specification, we limit attention to nouns (defined by Webster.com) and proper nouns that appear in no more than 25 percent of all product descriptions in order to avoid common words. We define proper nouns as words that appear with the first letter capitalized at least 90 percent of the time in our sample of 10-Ks. We also omit common words that are used by more than 25 percent of all firms, and we omit geographical words including country and state names, as well as the names of the top 50 cities in the United States and in the world.* [...]"

Table 1: Variable Definitions and Mapping to Compustat

Panel A: Observed Variables

Notation	Description	Measurement	Source
$p_i q_i$	Revenues	Revenues	Compustat
TVC_i	Total Variable Costs	Costs of Goods Sold	Compustat
f_i	Fixed Costs	Selling, General and Administrative Costs	Compustat
$\mathbf{a}_i' \mathbf{a}_j$	Cosine Similarity	Word frequencies in 10-K Business Description	Hoberg and Phillips (2016)

Panel B: Identified Variables and Parameters

Notation	Description	Identification
δ_i	Marginal Cost Slope	$= 0$ in baseline model, see 6.1 for the extended model
α	Utility Weight on Common Characteristics	$= -\frac{\varepsilon_{\text{KQ}} \cdot p_{\text{K}}\, q_{\text{K}} + \varepsilon_{\text{QK}} \cdot p_{\text{Q}}\, q_{\text{Q}}}{2 \cdot \cos_{\text{KQ}}^{\text{HP}} \cdot \sqrt{p_{\text{K}} q_{\text{K}} - \text{TVC}_{\text{K}}} \cdot \sqrt{p_{\text{Q}} q_{\text{Q}} - \text{TVC}_{\text{Q}}}}$; $\begin{pmatrix} \text{K} = \text{Kellog's} \\ \text{Q} = \text{Quaker Oats} \end{pmatrix}$
q_i	Output	$= \sqrt{\pi_i / (1 + \delta_i/2)}$
c_i^0	Marginal Cost Intercept	$= h_i/q_i - \frac{1}{2}\delta_i q_i$
$(\mathbf{I} + \mathbf{\Sigma})$	$\partial \mathbf{p} / \partial \mathbf{q}$	$= (1-\alpha)\,\mathbf{I} + \alpha \mathbf{A}'\mathbf{A}$
$\mathbf{b}$	Demand Intercept	$= (2\mathbf{I} + \mathbf{\Delta} + \mathbf{\Sigma})\,\mathbf{q} + \mathbf{c}$

Based on this vocabulary, HP produce, for each firm i, a vector of word frequencies $\mathbf{v}_i$. Each dimension of this vector corresponds to a word in HP's vocabulary, and the corresponding coordinate is given by the number of times that word appears in firm i's 10-K product description:

$$\mathbf{v}_i = \begin{bmatrix} v_{i,1} \\ v_{i,2} \\ \vdots \\ v_{i,61146} \end{bmatrix} \tag{3.1}$$

Finally, the HP cosine similarity between firm i and firm j is defined as follows:

$$\cos_{ij}^{\text{HP}} \overset{\text{def}}{=} \frac{\mathbf{v}_i' \mathbf{v}_j}{\|\mathbf{v}_i\| \, \|\mathbf{v}_j\|} \tag{3.2}$$

The fact that all publicly-traded firms in the United States are required to file a 10-K form makes this data set unique, in that it covers the near entirety (97.8%) of the Compustat universe. HP use these cosine similarities to produce a dynamic industry classification, called TNIC, which they extensively validate.

Since their introduction in 2011, HP's industry classifications have become standard in the empirical corporate finance literature, where they have replaced NAICS and SIC for a variety of applications. A major reason for this methodological shift is that HP's dataset addressed an important limitation of traditional industry classifications. While these have often been used (for lack of better alternatives) to capture product market competition[11], it is well-known that they are based on the concept of *production process* similarity, not product similarity[12]. This is also one reason why, in the I.O. and Antitrust literature, NAICS and SIC are generally only used to estimate production functions[13].

There are other factors that differentiate HP's database from traditional industry classifications. While NAICS and SIC are binary (firms are either in the same industry or different industries), HP's database also provides continuous similarity scores ranging from zero to one, thus accommodating the inherent fuzziness of product market rivalries. While NAICS and SIC are seldom updated, HP's similarity scores are updated yearly. While NAICS and SIC are arbitrarily assigned (Chen et al., 2016 show that firms strategically manipulate their industry classifications), HP's similarity scores are rule-driven and incentive-compatible: executives face legal liability for misrepresenting company information in SEC filings.

I begin my empirical analysis by visualizing HP's dataset. To do so, I have to reduce the dimensionality of the dataset from 61,146 (the number of words in the HP's vocabulary) to two. I do so using the algorithm of Fruchterman and Reingold (1991, henceforth FR), which is widely used in network science to visualize weighted networks[14].

The result of this exercise is Figure 3: every dot in the graph is a publicly traded firm as of 2004. Firm pairs that have a high cosine similarity appear closer, and are joined by a thicker line. Conversely, firms that are more dissimilar are not joined, and are more distant. From the graph, we can see that the distribution of firms over the space of product characteristics is manifestly uneven: some areas are significantly more densely populated with firms than others. Also, the network displays a pronounced community structure:

[11] Before HP's data was published, Bloom, Schankerman and Van Reenen (2013) constructed cosine similarities to estimate R&D spillovers. They used Compustat Segments data, which is based on NAICS/SIC industries. This dataset's coverage of Compustat is insufficient to estimate my model (not enough firms/years available).

[12] See the following Bureau of Labor Statistics Guide.

[13] For example: DEU's method to compute markups uses production function estimates for NAICS industries.

[14] The algorithm models the network nodes as particles, letting them dynamically arrange themselves on a bidimensional surface as if they were subject to attractive and repulsive forces. One known shortcoming of this algorithm is that it is sensitive to the initial configurations of the nodes, and it can have a hard time uncovering the cluster structure of large networks. To mitigate this problem, and to make sure that the cluster structure of the network is properly displayed, I pre-arrange the nodes using the OpenOrd algorithm (which was developed for this purpose) before running FR.

Figure 3: Network Visualization of the Hoberg-Phillips Dataset

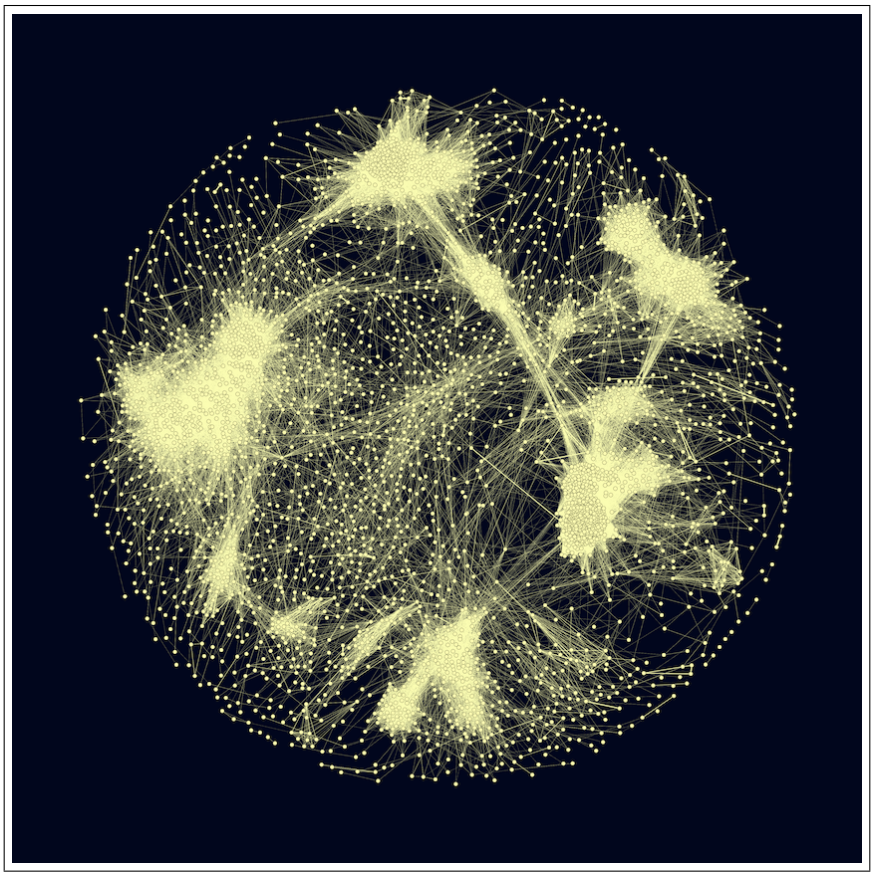


Figure Notes: The following diagram is a two-dimensional representation of the network of product similarities computed by Hoberg and Phillips (2016), which is used in the estimation of the model presented in Section 2. The data covers the universe of Compustat firms in 2004. Firm pairs that have thicker links are closer in the product market space. These distances are computed in a space that has approximately 61,000 dimensions. To plot this high-dimensional object over a plane, I applied the gravity algorithm of Fruchterman and Reingold (1991), which is standard in social network analysis.

large groups of firms tend to cluster in certain areas of the network.

3.4. Microeconometric Estimates of Demand Elasticity

In order to take the model to the data, we need to calibrate the free parameter α, which is the ratio between the inverse cross-price derivative $\partial p_i / \partial q_j$ and the cosine similarity $\cos_{ij}^{\text{HP}}$, and thus controls the overall size of the inverse cross-price elasticities $\partial \log p_i / \partial \log q_j$. We also need to validate the model, by comparing the GHL demand elasticities against estimates from the prior literature.

To calibrate α and validate the GHL elasticities, I manually collect microeconometric estimates of demand elasticities from three seminal demand estimation studies: Berry, Levinsohn and Pakes (1995)'s study of the automobile market, Nevo (2001)'s study of the ready-to-eat cereals market and Goeree (2008)'s study of the personal computer market.

Because many firms in Compustat are multi-product, in order to have some reasonable degree of compara-

bility between the microeconometric estimates and the corresponding GHL estimates I restrict my attention to firms that are (near) single-product that file 10-K forms that are covered in Compustat. Finally, I drop within-firm cross-price demand elasticities (which have no comparable object in my baseline model).

This procedure yields a database of 110 product-level demand elasticities that can be matched to firm pairs in the Compustat/Hoberg & Phillips sample, covering 7 firms and 17 firm-firm pairs. The firms covered are Ford, GM and Toyota for the Auto market; Kellogg's and Quaker Oats for the cereals market and Apple and Dell for the personal computer market.[15]

For all these product pairs, we have demand elasticities ($\partial \log q_i / \partial \log p_i$). For cereals only, I also obtain inverse cross-price elasticities ($\partial \log p_i / \partial \log q_i$) by inverting the matrix of elasticities. The inverse elasticities are not available for autos and PCs because the authors only report elasticities for a subset of the products, and thus the inverse matrix cannot be computed.

3.5. Identification

All of the unobserved variables in the model are identified subject to two parameters: (α) which controls the degree of horizontal differentiation between goods; and the diagonal matrix ($\mathbf{\Delta}$), which controls returns to scale. I will first show how to identify the remaining variables conditional on these two parameters and then I will illustrate my calibration procedure for α and $\mathbf{\Delta}$.

I measure the matrix $\mathbf{A}'\mathbf{A}$ using Hoberg and Phillips (2016)'s cosine similarity data. Given α, this provides an estimate of $\mathbf{\Sigma}$. We can then identify real output q_i from revenues and total variable cost data.

Proposition 10. *The physical output of granular firm i is identified as the following function of observables and the parameter δ_i:*

$$q_i \;=\; \sqrt{\frac{p_i q_i - \text{TVC}_i}{1 + \delta_i/2}} \qquad \text{if} \quad i \leq n \tag{3.3}$$

Proof. Appendix (G). □

If the model includes representative competitive firms ($i = n+1, n+2, ...$), the identification of q_i for these firms will be different. Specifically, the marginal cost pricing condition ($p_i = c_i$) implies that:

$$q_i \;=\; \sqrt{\pi_i + h_i} \qquad \text{if} \quad i > n \tag{3.4}$$

where $(\pi_i + h_i)$ is measured as the Value Added of private and foreign firms, which I compute using the OECD Trade in Value Added (TiVA) Dataset.

Having identified q_i, we can then pin down the vector of prices and the cost function intercepts:

$$p_i \;=\; \frac{p_i q_i}{q_i} \qquad c_i^0 \;=\; \frac{\text{TVC}_i}{q_i} - \frac{\delta_i}{2} q_i \tag{3.5}$$

Next, we state explicitly the identifying assumptions that allows us to identify the matrix $\mathbf{A}'\mathbf{A}$ given HP's cosine similarity data.

Assumption 1. *There are as many common characteristics as words – i.e. $\mathbf{v}_i$ and $\mathbf{a}_i$ are both vectors in $\mathbb{R}^m$.*

Assumption 2. *$\forall\, i$, the vectors $\mathbf{a}_i$ and $\mathbf{v}_i$ are collinear up to a permutation – that is $\mathbf{a}_i = \mathbb{P}\mathbf{v}_i\mathscr{C}_i$ for some $m \times m$ permutation matrix $\mathbb{P}$ and some strictly-positive scalar $\mathscr{C}_i$.*

[15] It makes sense to classify Apple as a single-product because the elasticities used date back to the '90s, when Apple had not yet entered the market for mobile devices.

What these two assumptions mean, intuitively, is that word frequencies in 10-K product descriptions ($\mathbf{v}_i$) can proxy for product characteristics loadings ($\mathbf{a}_i$). This is obviously a strong assumption, and one that needs to be validated empirically. However, it has powerful implications for identification.

Proposition 11. *Assumptions 1 and 2 imply that* $\mathbf{a}_i'\mathbf{a}_j \equiv \cos_{ij}^{\text{HP}}$.

Proof. Because there are as many words as common characteristics (m), we can re-label words so that $\mathbf{a}_i = \mathbf{v}_i\mathscr{C}_i$ (word 1 corresponds to characteristic 1, word 2 corresponds to characteristic 2 etc...). Because $|\mathbf{a}_i| = 1$ by construction, it must be that $\mathscr{C}_i = \|\mathbf{v}_i\|^{-1}$. Thus, equation (3.2) simplifies to $\mathbf{a}_i'\mathbf{a}_j$. □

We must clarify a crucial aspect of these set of assumptions: only *common* characteristics are being mapped to the vocabulary of Hoberg and Phillips. The idiosyncratic characteristics are instead assumed to be *unobserved*. This has important implications for the empirics. The presence of the idiosyncratic characteristics adds a degree of freedom to the demand system – the parameter α – which allows to calibrate the overall magnitude of the cross-price elasticities.

Having identified $\mathbf{A}'\mathbf{A}$, the matrix $\mathbf{\Sigma}$ is simply obtained using equation 2.15.. Finally, I identify the demand intercept b_i using equation (2.26):

$$\mathbf{b} \;=\; (2\mathbf{I} + \mathbf{\Delta} + \mathbf{\Sigma})\,\mathbf{q} + \mathbf{c}^0 \tag{3.6}$$

or, in the presence of a representative competitive firm:

$$\mathbf{b} \;=\; (\mathbf{I} + \mathbf{G} + \mathbf{\Delta} + \mathbf{\Sigma})\,\mathbf{q} + \mathbf{c}^0 \tag{3.7}$$

The last step required to take the model to the data is to to identify the scalar parameter α, which controls the elasticity of substitution among products. My strategy, for the baseline model, is to benchmark GHL against well-known demand estimation studies (Berry et al., 1995; Nevo, 2001; Goeree, 2008). All of these studies utilize the assumption that marginal cost is exogenous/flat. As a consequence, I will use this assumption as well for the baseline model (and will later relax it in Section E).

Assumption 3. *The marginal cost function is flat – that is,* $c_i \equiv c_i^0$ *for all firms* $i = 1, 2, ..., n$.

Conditional on this assumption, I will next show that we can immediately identify α, provided that we can observe the *inverse* cross-price demand elasticities for at least a product pair.:

Proposition 12. *Suppose that the inverse cross-price demand elasticity* $\varepsilon_{ij} = \frac{\partial \log p_i}{\partial \log q_j}$ *is observed for some firm pair* (K,Q). *Then,* α *is identified as the following function of observables:*

$$\alpha \;=\; -\frac{\varepsilon_{\text{KQ}} \cdot p_{\text{K}}\, q_{\text{K}} + \varepsilon_{\text{QK}} \cdot p_{\text{Q}}\, q_{\text{Q}}}{2 \cdot \cos_{\text{KQ}}^{\text{HP}} \cdot \sqrt{p_{\text{K}} q_{\text{K}} - \text{TVC}_{\text{K}}} \cdot \sqrt{p_{\text{Q}} q_{\text{Q}} - \text{TVC}_{\text{Q}}}} \tag{3.8}$$

Proof. Appendix G. □

Luckily, as hinted in 3.4, there exists a pair of firms for which we are able to obtain an estimate of the inverse cross-price elasticities from the previous literature: the firms in question are Kellogg's and Quaker Oats (hence the K and Q subscripts), and the corresponding inverse elasticities can be obtained by inverting the matrix of demand elasticities estimated by Nevo (2001) in his landmark study of ready-to-eat cereals.

By applying equation (3.8), I obtain a value of α equal to 0.12. Assumption 3 is relaxed in the extended model, which is presented in subsection (6.1).

4. Validation

This subsection is dedicated to validating Hoberg & Phillip's cosine similarity data and the GHL demand system, and is organized around five key questions about the model's empirical reliability.

4.1. Cluster Structure of the Network

As shown before, the network of product market rivalries generated by HP presents a characteristic "cluster" structure. It is natural to think of these macro-clusters of firms as product markets. If indeed HP's word frequencies are a good proxy for product characteristics (as previously assumed) we should expect to see a high degree of overlap between the network clusters in HP's data and some externally-defined definition of broad product markets.

To answer this question, we must first find an industry definition that better reflects product markets than NAICS and SIC (which, as discussed, reflect production process similarity and not product substitutability). Following Boller and Morton (2020), I use Standard & Poor's GIC broad industries. This provides a way to (independently) test the validity of Assumptions 1 and 2.

I test this implication in Appendix D: I reproduce Figure 3, coloring every dot (firm) according to its GIC product classification. Despite the fact that the graph was drawn *without* targeting GIC industries, there a very near perfect overlap between broad GIC industries and the network cluster structure, which is strong evidence that indeed HP's vocabulary of Hoberg captures product market interactions among firms.

4.2. Self-reported Product Market Rivalries

The previous validation exercise only shows that HP's data captures well broad industries, but how well does HP's data capture rivalry relationships among firms in narrower markets? The answer to this question is provided by Hoberg and Phillips (2016) in their seminal paper, where they use these cosine similarity scores to construct their own, product-centric industry classification - TNIC.

To validate TNIC's ability to identify product rivalries even at a granular level, HP use the fact that many (but not all) publicly-traded corporations self-report their closest peers/rivals in their 10-K form, and this information can be obtained from the Capital IQ database.

They show that TNIC (constructed to match NAICS and SIC in terms of granularity), strongly outperforms both NAICS and SIC in identifying self-reported rivalry relationships.

4.3. Microeconometric Demand Estimates

The previous exercises show that, in purely qualitative terms, HP's cosine similarity data performs strongly in identifying product market rivalries, in both broad markets as well as narrow markets. The next question is whether the GHL demand system, calibrated with HP's cosine similarity data, produces demand elasticities that are comparable in sign and magnitude to the microeconometric estimates from the previous empirical IO literature (previously described in 3.4).

To make a comparison, I convert the product-product microeconometric elasticities into firm-firm elasticities by taking firm-firm medians (again, taking care to exclude within-firm cross elasticities), and then match them to the corresponding GHL estimates for the earliest available year (1996 or 1997).

The result of this comparison is shown in Table 2. The first thing that we can observe from this table is that the sign of the GHL estimates matches that of the corresponding microeconometric estimate for every single firm-firm pair. This is not too surprising, since all the empirical papers providing the microeconometric estimates focus on narrow markets and close competitors.

Table 2: Demand Elasticities: Microeconometric Estimates (Untargeted) vs. GHL

			Demand Elasticity $\left(\frac{\partial q_i}{\partial p_j} \cdot \frac{p_j}{q_i}\right)$	
Market	**Firm** i	**Firm** j	**Micro Estimate**	**GHL** (*text-based*)
Auto	Ford	Ford	-4.320	-5.197
Auto	Ford	General Motors	0.034	0.056
Auto	Ford	Toyota	0.007	0.017
Auto	General Motors	Ford	0.065	0.052
Auto	General Motors	General Motors	-6.433	-4.685
Auto	General Motors	Toyota	0.008	0.005
Auto	Toyota	Ford	0.018	0.025
Auto	Toyota	General Motors	0.008	0.008
Auto	Toyota	Toyota	-3.085	-4.851
Cereals	Kellogg's	Kellogg's	-3.231	-1.770
Cereals	Kellogg's	Quaker Oats	0.033	0.023
Cereals	Quaker Oats	Kellogg's	0.046	0.031
Cereals	Quaker Oats	Quaker Oats	-3.031	-1.941
Computers	Apple	Apple	-11.979	-8.945
Computers	Apple	Dell	0.018	0.025
Computers	Dell	Apple	0.027	0.047
Computers	Dell	Dell	-5.570	-5.110

What *is* surprising from the table is that the GHL estimates also match with striking accuracy the magnitudes – not just for the own v/s cross price elasticities as a whole, but for individual firm pairs. It's also crucial to note that these demand elasticities are *untargeted* moments. All we used to calibrate α was the *inverse* cross-price demand elasticity between Kellogg's and Quaker Oats (which provides no guarantee of being able to match the corresponding demand elasticity, since the matrix of elasticities depends on the entire $\mathbf{\Sigma}$ matrix).

One feature of Table 2 that may makes it difficult to evaluate model fit using a single summary statistic (such as a correlation coefficient) is that it combines own demand elasticities (which are negative and large in absolute value) with cross-price demand elasticities (which are positive and small in absolute value); the correlation coefficient is extremely high (0.99) but this entirely due to the variation across own and cross-demand elasticities.

To construct a single, extremely-stringent summary statistic of model fit, I perform the following analysis. I take the log of the absolute value of each observation in the two right-most columns of Table 2. Then, I residualize the resulting two series on a dummy variable takes value 1 for own price elasticities ($i = j$), as well as on market fixed effects. Only after these manipulations, I consider the correlation of the two resulting series. The rationale for these manipulations is that we don't want to "give any points" to the model for matching the magnitude and sign of own and cross-price elasticities as a whole; we also don't want to give any points to the model for matching the variation at the market level (e.g. for nailing the average cross-price elasticity in autos as opposed to cereals). In sum, we want to evaluate the model solely on its ability to match *firm-firm* variation in elasticities *within markets* and *within own/cross* groups.

FIGURE 4: DEMAND ELASTICITIES: GHL VS. MICROECONOMETRIC ESTIMATES (UNTARGETED)

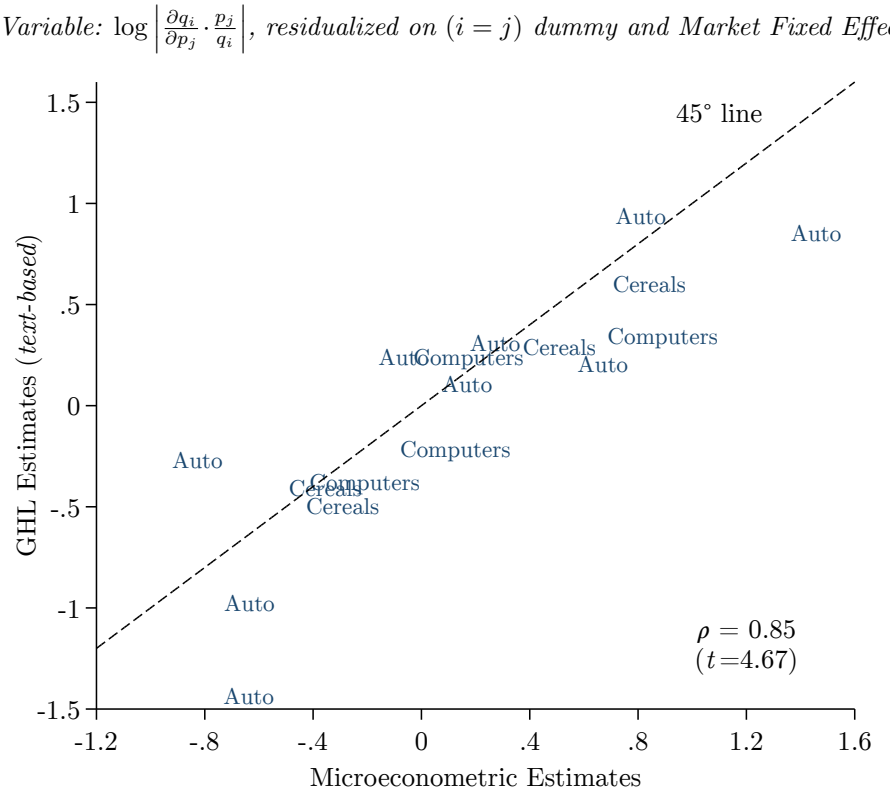


FIGURE NOTES: the scatter plot above illustrates GHL's ability to fit microeconometric estimates of demand elasticities. Each observation is a firm-firm pair and the label represents the relevant market. On both axes, the variable being plotted is the log-module of the demand elasticity $\left(\log\left|\frac{\partial q_i}{\partial p_j}\cdot\frac{p_j}{q_i}\right|\right)$, residualized on the "own/cross" dummy variable and market fixed effects. The GHL estimate is plotted on the vertical axis. The corresponding microeconometric estimate is plotted on the horizontal axis, and is a median of product-level demand elasticity for un-diversified firms.

4.4. Markups

GHL demand elasticities closely approximate (without directly targeting) the traditional IO micro estimates of demand elasticity and markups. One limitation, of course, is that such demand-based estimates are only available for a limited of industries. Recently De Loecker, Eeckhout and Unger (2020, DEU) were able to estimate markups for US public corporations using a supply-side approach that involves estimating the firms' production technology. It is thus natural to ask how the distribution of markups generated by the GHL model compare to those estimated by DEU.

For the baseline model with flat marginal cost, this comparison is omitted, for the trivial reason that, under the assumption of constant returns to the variable input, the markups implied by my baseline model coincide exactly, by construction, with those of DEU: both are equal to Revenues/COGS.

However, an important contribution of DEU is to be able to relax this assumption - by doing so they are able to obtain alternative measures of markups. I do so in a model extension that I present in 6.1. In the same subsection, I show that, even after relaxing the flat marginal cost assumption, my extended model generates markups that correlate extremely closely – both in the cross section and over time – with DEU's.

TABLE 3: WELFARE STATICS (2019)

	Scenario	*Cournot-Nash*	*Perfect Competition*	*Monopoly*	*Resource-Efficient*
		(1)	(2)	(3)	(4)
Welfare Statistic	**Notation**	$\mathbf{q}^{\Phi}$	$\mathbf{q}^{W}$	$\mathbf{q}^{\Pi}$	$\mathbf{q}^{H}$
Total Surplus (US$ trillions)	$W(\mathbf{q})$	13.771	15.516	10.969	14.990
Aggregate Profits (US$ trillions)	$\Pi(\mathbf{q})$	3.117	0.000	6.421	0.000
Consumer Surplus (US$ trillions)	$S(\mathbf{q})$	10.653	15.516	4.547	14.990
Total Surplus / Perfect Competition	$\frac{W(\mathbf{q})}{W(\mathbf{q}^{W})}$	0.888	1.000	0.707	0.966
Aggregate Profit / Total Surplus	$\frac{\Pi(\mathbf{q})}{W(\mathbf{q})}$	0.226	0.000	0.585	0.000
Consumer Surplus / Total Surplus	$\frac{S(\mathbf{q})}{W(\mathbf{q})}$	0.774	1.000	0.415	1.000

TABLE NOTES: The following table shows my estimates of aggregate profits, consumer surplus and total surplus in each of the counterfactuals scenarios presented in Section 5.

5. Empirical Findings

In this section, I present the results of the estimation of my model counterfactual exercises. My baseline estimates reflect the baseline model implementation that only includes granular firms (Compustat) and has flat marginal costs. In the next section, I explore various extensions of the model.

5.1. Policy Implementation of Counterfactuals

To study the *Perfect Competition* and the *Resource-Efficient* counterfactuals, I must make an assumption about how the social planner implements these counterfactual allocations. In these allocations, firms price at marginal cost thus do not generate any economic profits to cover their fixed costs. This means that, in the absence of any planner-mandated transfers from the households, aggregate profits would be negative; the representative household would voluntarily make payments to firms (instead of collecting dividends) and thus consumer surplus would be larger than total welfare.

Thus, in the policy implementation of these counterfactual, I make a "limited liability" assumption: I assume that the social planner ensures the non-negativity of aggregate profits by giving lump-sum subsidies to individual firms; these are, in return, financed by a lump sum tax imposed on the household. Thus, by construction, in the *Perfect Competition* and the *Resource-Efficient* counterfactuals, aggregate profits are zero and total surplus is equal to consumer surplus.

5.2. Welfare Statics

My first empirical exercise is to compute total surplus and break it down into profits and consumer surplus. This is done for both the observed equilibrium (which is assumed to be a Nash-Cournot equilibrium) and the counterfactuals considered in Section 2. These estimates are all shown in Table 3.

I estimate that the publicly-traded firms earn an aggregate economic profit of $5 trillion and produce an estimated total surplus of $9.1 trillion. Consumer surplus is therefore estimated to be about $4 trillion. About 55% of the total surplus produced is appropriated by the companies in the form of oligopoly profits. For context, the GDP of U.S. corporations in the same year (2017) is $11 trillion[16].

The first counterfactual that I consider, *Perfect Competition*, appears in the second column. By comparing this counterfactual with the Cournot-Nash allocation we can see that the welfare costs of oligopoly are significant. Under perfect competition, aggregate surplus is significantly higher – $10.2 trillion – hence, the deadweight loss amounts to about 11% of the total surplus.

While the effects of oligopoly on Pareto efficiency are significant, perhaps even more significant are the *distributional* effects. When firms price at marginal cost, a much larger share of the surplus goes to the consumer: $8.2 trillion, more than double than in the Cournot allocation. This amounts to 80% of the total surplus.

The next counterfactual I analyze, the *Monopoly* counterfactual, appears in the third column: it represents a scenario in which all firms are controlled by a single decision-maker that coordinates supply choices. In this allocation, aggregate surplus is significantly lower than in the Network Cournot equilibrium allocation: $8.2 trillion. Despite the decrease in aggregate welfare, profits are markedly higher: $5.7 trillion. Consequently, consumer surplus is reduced to just $2.5 trillion, a mere 33% of the total. One interpretation of this exercise is that policies that prevent coordination between firms (antitrust) have a large positive impact on welfare.

[16]The difference between GDP and total surplus is that total surplus does not include the value of labor input but it does include the value of inframarginal consumption. GDP, on the other hand, includes the value of labor input but not the inframarginal value of consumption. In this model each unit of labor is paid exactly its marginal disutility, hence there is no inframarginal value of leisure.

FIGURE 5: DISTRIBUTION OF PRODUCTIVITY AND PRODUCT MARKET CENTRALITY

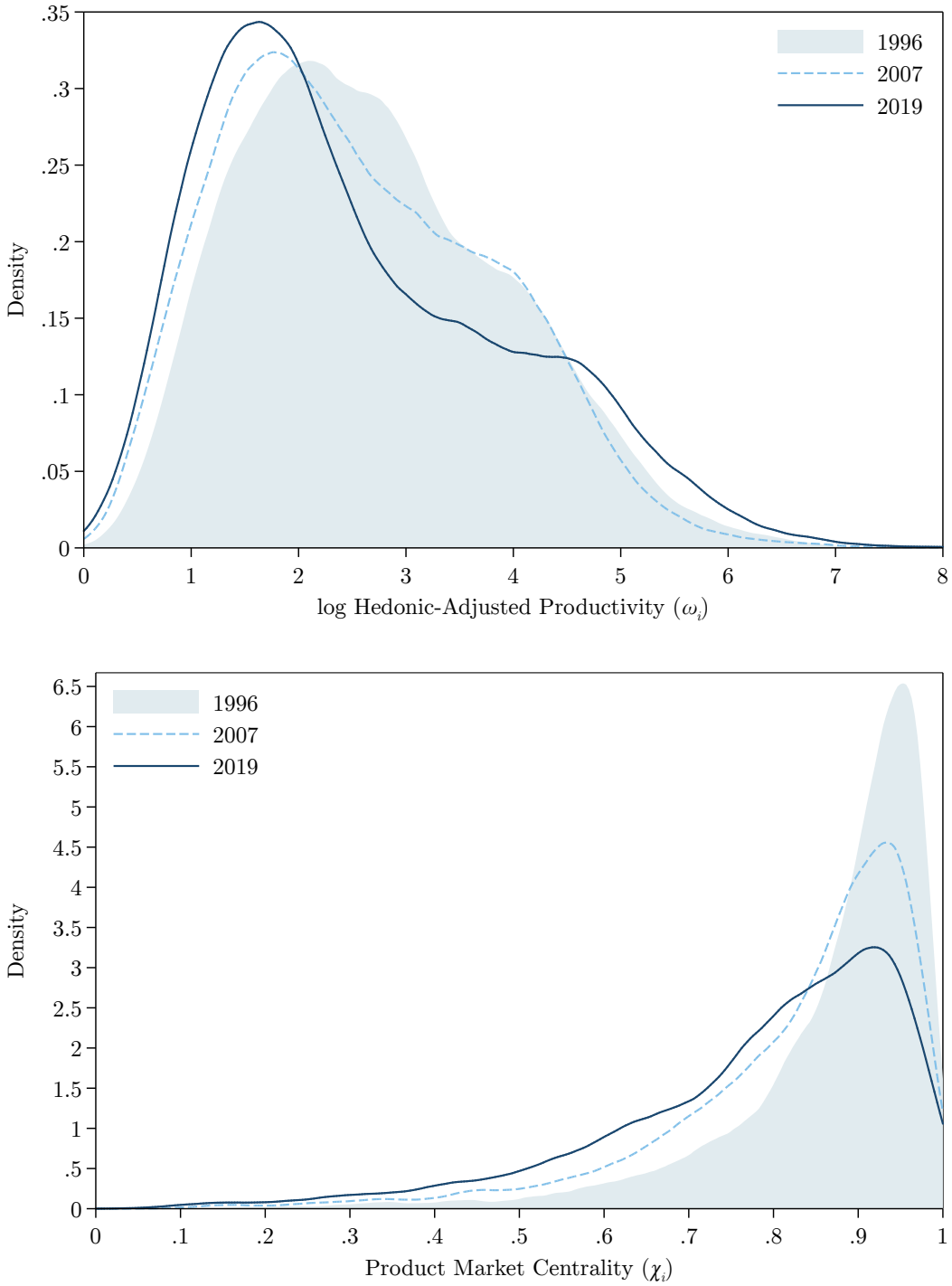


FIGURE NOTES: the figure above displays a kernel density estimate of the cross-sectional distribution of product market centrality (χ_i) in 1996 (grey area), 2007 (dotted light line) and 2019 (solid dark line). These estimates are based on the baseline model (with flat marginal cost).

Next, I consider the *Resource Efficient* counterfactual, in which the social planner maximizes total surplus subject to the constraint of not using more labor than the Cournot equilibrium.

In this scenario, markups across firms have been equalized, but not eliminated. By removing all dispersion in markups, this counterfactual targets the malallocative effects of oligopoly.

The total surplus produced in this counterfactual is \$9.9 trillion, about 3.3 percentage points lower than in perfect competition, and \$800 billion higher than the observed Cournot-Nash equilibrium. Most of the surplus produced – \$6.9 trillion, or 68% of the total – goes to the consumer; profits are reduced to \$3.1 trillion. Because labor is fixed, all the welfare gains with respect the Cournot equilibrium come from the *reallocation* of labor. Hence, an important take-away from this counterfactual is that a large share of the inefficiencies from oligopoly are driven by resource misallocation. A different way to say this is that the dispersion in markups (caused by oligopolistic competition with differentiated products) matters at least as much as the level of markups in determining the overall deadweight loss.

In sum, my measurements suggest that the oligopoly power of U.S. public firms has significant consequences for aggregate welfare, and that it impacts consumer welfare through two channels: it increases the dispersion of markups, generating resource misallocation which raises the deadweight loss; it also increases the level of markups, which in turn affects how surplus is shared between producers and consumers.

5.3. Hedonic-Adjusted Productivity and Centrality

By taking the model to the data, we recover the firm-level weighted market share ($\mathcal{M}_i$) hedonic-adjusted productivity (ω_i) and centrality (χ_i). The first empirical question we ask is: what is the distribution of these two statistics, and how does it change over time?

The upper panel of Figure 5 the cross-sectional distribution of $\mathcal{M}_i$ for the years 1996, 2007 and 2019 (estimated using kernel density). The data presents an enormous degree of dispersion in productivity. For the year 2019, a firm at the 75^{th} percentile is 7 times as productive as a firm at the 25^{th} percentile. The distribution also appears to be left-skewed. The other pattern that emerges rather clearly is a "superstar" effect (Autor et al., 2020): the dispersion appears to have increased between 1996 and 2019, and the distribution appears has become more bimodal over time. These patterns would lead to a progressive increase in industry concentration in a large class of models.

The lower panel of Figure 5 similarly displays the cross-sectional distribution of κ_i in 2019, 2007 and 1996. Again, there appears to be a significant heterogeneity in centrality. The distribution is markedly left-skewed, with most firms having a high degree centrality and thus behaving relatively competitively, while a smaller mass of firms displays low centrality and thus behave more like a monopolist. Between 1996 and 2007 and between 2007 and 2019, there distribution displays a very clear shift to the left: this is consistent with an overall increase in market power.

In sum, this decomposition of firm markup suggests that the "superstar" firm hypothesis and the "rise of oligopoly" hypothesis are not at all mutually exclusive. It appears instead that both phenomena are taking place simultaneously. The decomposition is a useful tool to disentangle the influence of these two forces on the distribution of markups.

5.4. Markups Growth Decomposition

While there's evidence in support of the fact that the markups of US corporations have increased in the past decades (De Loecker, Eeckhout and Unger, 2020), the reasons behind this increase are still very much up for debate: are higher markups the result of a right tail of highly-productive superstar firms pulling away from the rest the competition, or are they the result of a softening of product market competition (i.e. fewer substitutes)?

FIGURE 6: MARKUPS GROWTH DECOMPOSITION (1996-2019)

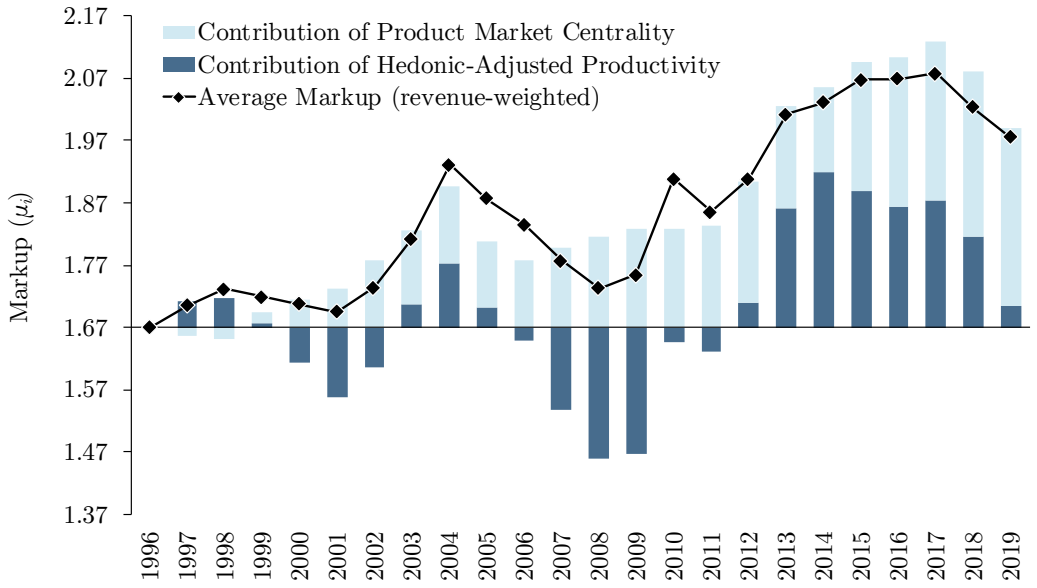


FIGURE NOTES: The figure decomposes the growth rate of the average markup (μ_i), weighted by revenues, using the formula in equation (5.2).

Equation (2.39) allows us to study this question by decomposing markups into two components reflecting these forces: by implementing the decomposition, I investigate whether the rise of markup is driven by ow product market centrality (χ_i) or hedonic-adjusted productivity (ω_i).

Defining the operator $\mathbb{E}(\cdot)$ as the (weighted) cross-sectional average of a variable, and letting $\text{cov}(\cdot)$ be the corresponding cross-sectional (weighted) covariance, we can write the following formula for the revenue-weighted markup, based on equation (2.39):

$$\mathbb{E}\mu_{it} \;=\; 1 + (1 - \mathbb{E}\chi_{it}) \cdot (\mathbb{E}\bar{\mu}_i - 1) - \text{cov}(\chi_{it}, \bar{\mu}_{i,t}) \tag{5.1}$$

where t represents the year of observation. Differentiating with respect to time, we obtain the following decomposition for the change in the revenue-weighted average markup:

$$\Delta\mathbb{E}(\mu_{it}) \quad \approx \quad \underbrace{\Delta\,\mathbb{E}\,\bar{\mu}_{i,t} \cdot \mathbb{E}(1 - \chi_{i,t-1})}_{\text{Contribution of Hedonic}-\text{Adjusted Productivity}} \quad - \quad \underbrace{\Delta\,\mathbb{E}(\chi_{i,t}) \cdot (\mathbb{E}\,\bar{\mu}_{i,t-1} - 1)}_{\text{Contribution of Product Market Centrality}} \tag{5.2}$$

where the discrepancy is determined by a higher-order/covariance term.

Figure 6 illustrates the implementation of this decomposition for the period 1995-2019. Over this period, we can see that the revenue-weighted average markup increased from 1.67 to 1.97. Relative to the minimum theoretical value of one, this value represents an increase of 44%. The decomposition shows that approximately 9/10 of this increase (.276) can be attributed to changes in product market centrality (χ_i), with the remaining share of the increase (.048) being driven by the contribution of hedonic-adjusted productivity (ω_i). The contribution of the higher-order term is, overall, negligible.

While it appears clearly, from these findings, that centrality drives the slow-moving upward trend in average markups, the hedonic-adjusted productivity component dominates the cyclical component of markups.

FIGURE 7: TOTAL SURPLUS OF US PUBLIC FIRMS (1996-2019)

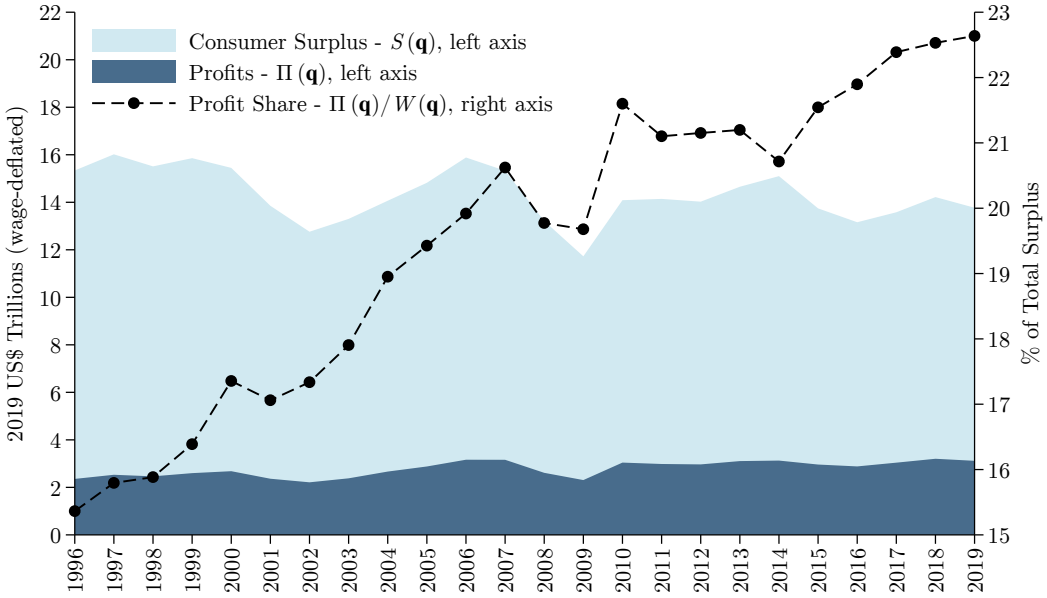


FIGURE NOTES: The figure above plots the evolution, between 1996 and 2019, of aggregate profits $\Pi(\mathbf{q})$, aggregate consumer surplus $S(\mathbf{q})$ and total surplus $W(\mathbf{q})$, as defined in the model from Section 2. Profits as a percentage of total surplus (Π/W, black dotted line) are shown on the right axis. These statistics cover the universe of the US publicly-listed corporations.

5.5. Time Trends of Profits, Consumer Surplus and the Deadweight Loss

HP's cosine similarity data is available starting from 1997. By mapping my model to Compustat data year by year, I can produce annual estimates of the welfare metrics previously presented. This allows me to study the welfare implications of the rising concentration of US industries. Most importantly, because my model leverages HP's time-varying product similarity data, these estimates account for how the product offering of US public firms changed over time. This is another contribution of this study.

In Figure 7, I plot aggregate consumer surplus S (the dark area) and profits Π (the light area) for every year between 1996 and 2019. The combined area represents total surplus W. I also plot, on the right axis (dotted black line), profits as a share of total surplus Π/W. All these statistics refer to the (observed) Cournot equilibrium.

The graph shows that the total surplus produced by US public corporations has nearly doubled between 1996 and 2019 from \$4.6 trillion to \$9.1 trillion between 1996 and 2019. Profits have increased more-than-proportionally with respect to consumer surplus – from about \$2.2 trillion to about \$5 trillion. Consumer surplus increased instead from \$2.2 trillion in 1997 to about \$4 trillion i 2017. As a consequence, the profit share of surplus has increased from about 50% of total surplus to nearly 56%. The consumer appears to capture a decreasing share of the surplus generated by public companies.

In Figure 8, I plot, over the same period, the percentage gain in total surplus from moving from the competitive equilibrium $\mathbf{q}^{\Phi}$ to the first best $\mathbf{q}^{W}$. This is the deadweight loss from oligopolistic behavior, and is plotted as the darker line. Its trend that mimic that of profit share of surplus: it increased from 8.5% (in 1997) to the current level of 11% (in 2017). This suggests that the impact of oligopoly on surplus creation has increased over time.

To investigate the role of resource misallocation in generating this trend I plot, in the same figure (light

FIGURE 8: DEADWEIGHT LOSS FROM OLIGOPOLY (1996-2019)

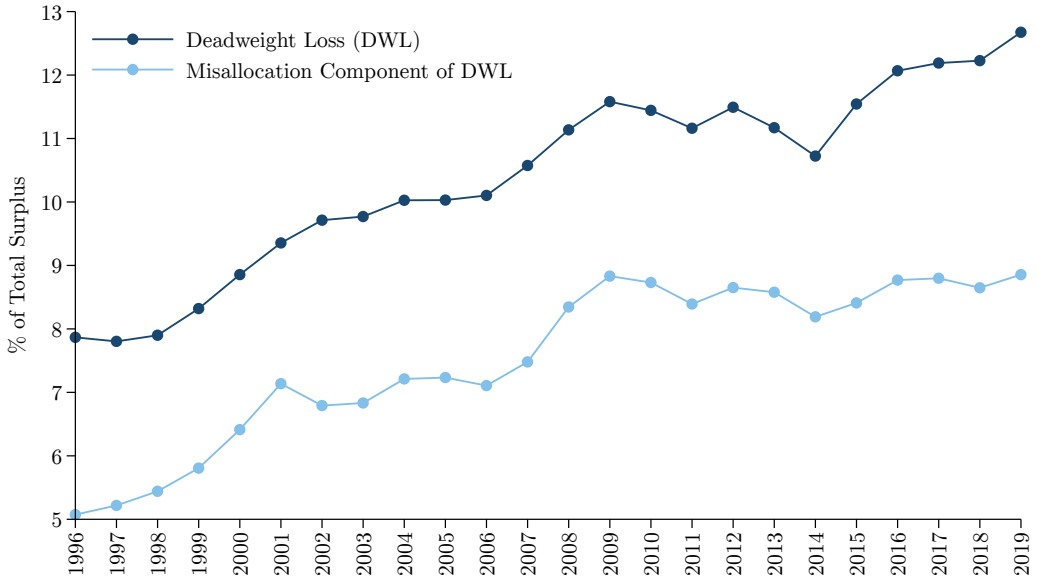


FIGURE NOTES: The following figure plots the estimated deadweight loss (DWL) from oligopoly, between 1996 and 2019. The darker line is the traditionally-defined DWL - the % difference in total surplus between the Cournot equilibrium and the First-Best scenario, while the lighter line is the % difference between the Cournot equilibrium and the Resource-Efficient Allocation.

line), the percentage difference in total surplus between the Cournot equilibrium and the *Resource Efficient* counterfactual: it has increased from 5 percentage points (in 1996) to 8.7 percentage points (in 2019). This implies that a significant share of the increase in the deadweight loss has been driven by increasing misallocation.

Overall, my findings are consistent with the interpretation that U.S. public firms have more oligopoly power than they had in 1996, and that this increase in oligopoly power has had a significant impact on both allocative efficiency and consumer welfare.

5.6. The Cross-Section of Mergers and Acquisitions

One factor that is frequently mentioned as a potential culprit for the declining competition is the (allegedly) lax merger enforcement by US competition authorities. Because my model allows to compute cross-price elasticities for millions of firm pairs, it can provide a fresh perspective on how the cross-section of mergers. An interesting question that we can ask is how do merging firms look like in terms of product substitutability, compared to firms that do not merge, and whether these differences have somewhat changed over time. To measure their degree product market interaction among merging companies I shall use the diversion ratio, which is one of the measures that FTC-DOJ Horizontal Merger Guidelines highlight as a useful statistic for merger analysis.[17]

The diversion ratio is defined as the change in quantity demanded of product i for a price change in product j that yields a unit decrease in the quantity demanded of product j. It is easy to see that in this model diversion ratios take a simple expression in terms of the inverse matrix $(\mathbf{I} + \mathbf{\Sigma})^{-1}$:

[17] We can alternatively use the cosine similarity: it makes little difference.

FIGURE 9: DIVERSION RATIOS OF MERGING FIRMS

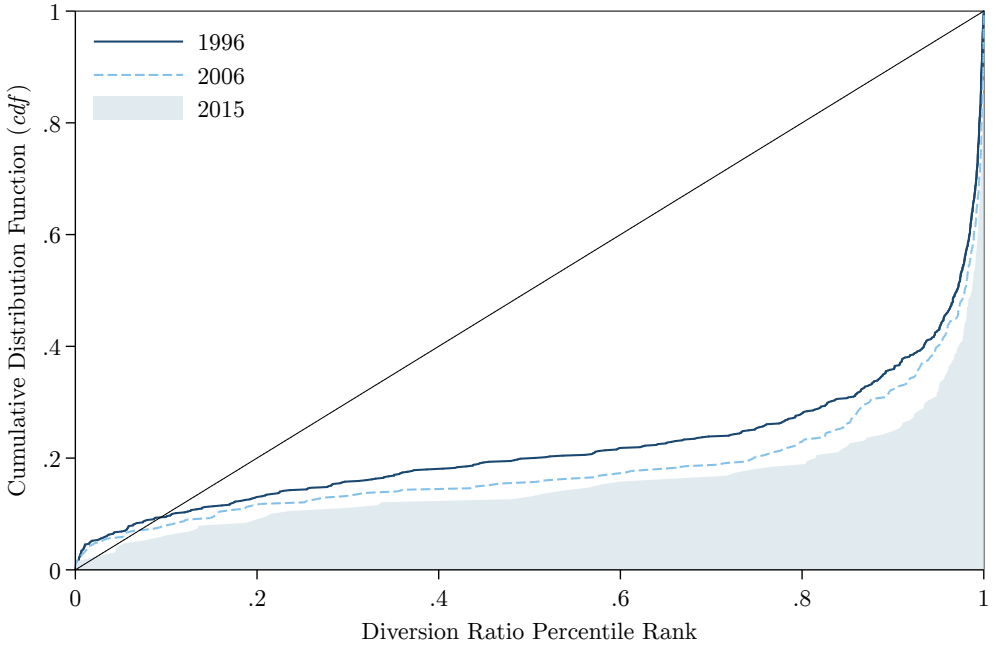


FIGURE NOTES: The figure above plots the cumulative distribution function of the diversion ratio percentile ranks of merging firms in 1996 (solid dark line), 2006 (dotted light line) and 2015 (grey area). These estimates are based on the baseline model with flat marginal cost. The 45° line displays a uniform distribution, which would be expected if selection into mergers was independent of diversion ratios.

$$\text{Diversion Ratio}_{ij} \quad \stackrel{\text{def}}{=} \quad \frac{\partial q_i}{\partial p_j}\left(\frac{\partial q_j}{\partial p_j}\right)^{-1} \quad = \quad \frac{(\mathbf{I}+\mathbf{\Sigma})^{-1}_{ij}}{(\mathbf{I}+\mathbf{\Sigma})^{-1}_{jj}} \tag{5.3}$$

In the next empirical exercise, I use the database of announced mergers between public firms constructed by Ewens, Peters and Wang (2019) and Phillips and Zhdanov (2013), which covers publicly-traded companies the period up to 2016. Then, for every pair of firms, I compute the diversion ratio for the preceding year, and the corresponding (within-year) percentile rank. Then, in Figure 9, I plot the cumulative distribution of the percentile ranks of merging firms – by year. I also compare it to a uniform distribution, which is what we would expect that distribution to look like if merging firms were randomly selected.

What clearly emerges from Figure 9 is that merging firm pairs are highly selected from the right tail of the distribution of diversion ratios. In other words, mergers tend to mostly happen among firms that interact strongly in product markets. It is also evident from the figure that this effect has become more pronounced over time: the distribution of percentile ranks of 2015 first-order stochastically dominates that of 2006, which in turn first-order stochastically dominates that of 1996. The typical merging firm pair in 1996 had a diversion ratio around the 97^{th} percentile. In 2015, the typical merging firm had a diversion ratio just shy of the 99^{th} percentile.

There doesn't seem to be a significant trend in the intensity of M&A activity among public corporations: the number of merging firms as a ratio of the total number of firm pairs has remained broadly stable at around 12 per million over this period.

Obviously, there is much that this analysis does not capture (potential synergies, vertical relationships,

Figure 10: Venture Capital Startup Exits by Type

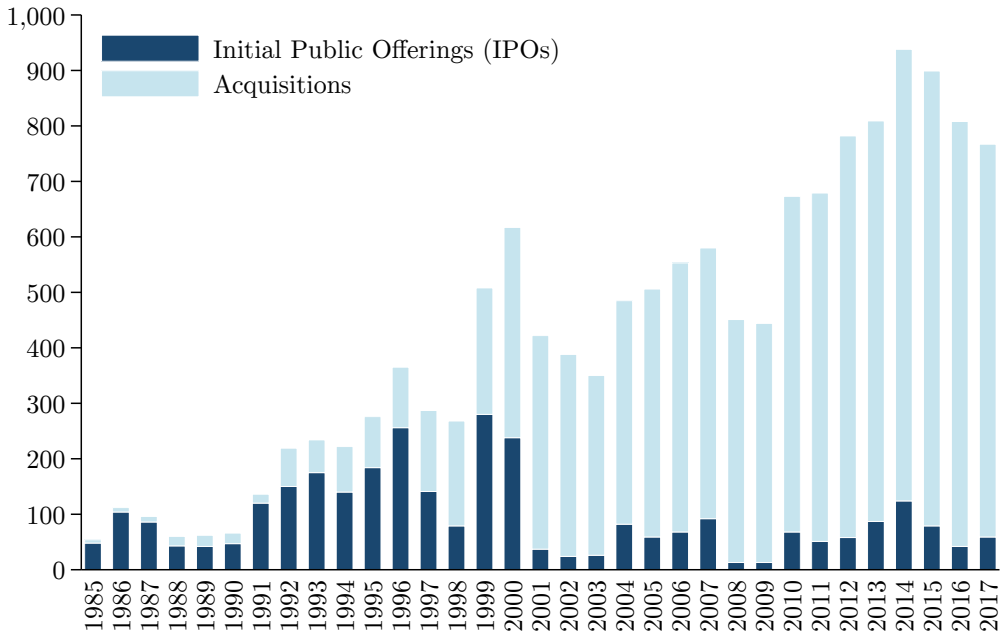


Figure Notes: the figure above plots the number of successful venture capital exits in the United States by year and type (Initial Public Offering v/s Acquisition). The data is sourced from the National Venture Capital Association (NVCA).

etc...) and we cannot simply infer from it that M&A activity has become more harmful for competition. Nonetheless, the analysis above clearly indicates that merger activity has become more concentrated, over time, among firms that are more likely to interact in the product market.

5.7. Firm Dynamics and the Role of Startups Acquisitions

One important driver of increase in oligopoly power (measured as the the left shift in the distribution of product centrality) is the increase in output and revenues concentration at the top end of the firm size distribution. In Compustat, this is reflected as a secular decline in the rate of IPOs (Kahle and Stulz, 2017) that has not been matched by a decrease in the rate of exit. One interesting aspect of the decline of IPOs is that it appears to be unrelated to the decline of the startup rate that has been measured in the broader economy (Decker et al., 2014). Far from declining, the number of startups that are backed by Venture Capital (VC), which make up the majority of startups that eventually become public companies, has boomed over this period.

Figure 10 plots the number of Venture Capital exits in the United States by year and type, for the period 1985-2017[18]. In the diagram, I separate IPOs from acquisitions. We can see from this figure that, at the beginning of the 1990s, the vast majority of VC exits were IPOs. However, since the mid-90s, there has been a dramatic shift toward acquisitions. One implication of this fact is that the decline of IPOs was not driven by a dearth of startups. Instead, the reason why IPOs have decreased is that they have been largely replaced by acquisitions. In a companion paper (Ederer & Pellegrino, forthcoming 2023), I investigate the role of startup acquisitions in the rise of oligopoly is investigated (using the very same model).

[18]This data is sourced from the National Venture Capital Association (NVCA).

6. Robustness and Extensions

In this section, I investigate the robustness of my empirical results to a variety of assumptions; I propose extensions to the model; and I discuss the limitations of this paper and potential follow-up work.

6.1. Variable Returns to Scale and Supply-Side Markups Estimation

Empirical IO economists use two broad strategies to estimate firm markups (De Loecker and Scott, 2016). The first relies on demand estimation, while the latter relies on production function estimation and was implemented by De Loecker et al. (2020) for the universe of publicly-listed companies. Thus far I have discussed the link between this paper and the demand estimation literature. However, this paper is also closely related to the literature on production function-based markup estimation. In fact, my model relies at least in part on supply-side methods for the identification of markups. This claim is based on the following two facts: first, my model utilizes the same data as DEU (revenues and costs) to estimate markups. Second, markups are identified, in my model, not by cosine similarities, but by knowledge of the of firm's production or cost function - in accordance with the supply-side markup estimation (De Loecker and Warzynski, 2012)

The supply-side approach to estimating markups relies not on a conduct assumption (Cournot, Bertrand profit maximization) but on the assumption that firms are cost minimizers. The first order condition for a cost-minimizing firm, with respect to some generic flexible input h_i (which in our case is labor) is:

$$p_i^h \;=\; \lambda_i \frac{\partial q_i}{\partial h_i} \tag{6.1}$$

where q_i is output, which depends on flexible input h_i, p_i^h is the price of that input, and λ_i is the Lagrange multiplier for the cost minimization problem. The markup μ_i is defined as the ratio between the output price (p_i) and the Lagrange multiplier (λ_i):

$$\mu_i \;=\; \frac{p_i}{\lambda_i} \tag{6.2}$$

Putting together the two equations above (substituting λ_i) and after a few manipulations, we obtain the following expression for the markup:

$$\mu_i \;=\; \frac{\partial \log q_i}{\partial \log h_i} \cdot \frac{p_i q_i}{p_i^h h_i} \tag{6.3}$$

now, the measure of revenues ($p_i q_i$) is exactly the same in my model as in that of De Loecker, Eeckhout and Unger (2020). Variable costs are also measured the same way (Costs of Goods Sold), although in my model the generic input associated with COGS is labelled "Labor".

Up to this formula, DEU's approach to retrieving markups is exactly the same as mine: if we could hypothetically observe/recover the elasticity $\partial \log q_i / \partial \log h_i$, the equation above would immediately yield the markup for either model with no assumptions about conduct. In practice, the elasticity $\partial \log q_i / \partial \log h_i$ has to estimated. In my baseline model we assumed a constant marginal cost. In this special case of the previous baseline model, where we have constant returns to the COGS variable input ($\partial \log q_i / \partial \log h_i = 1$), the two markups exactly coincide and can be estimated as Revenues/COGS without conduct assumptions or any econometrics.

What I do next is to relax the assumption that the marginal cost is exogenous (the MC function is flat). This is where my methodology diverges now from that of DEU. While DEU assume a Cobb-Douglas production function:

$$q_i \;\propto\; h_i^{\theta_i} \tag{6.4}$$

while I assume a quadratic cost function:

$$h_i \;=\; f_i + c_i^0 q_i + \frac{\delta_i}{2} q_i^2 \tag{6.5}$$

resulting in the following two expressions for the markup

$$\mu_i^{\text{DEU}} = \theta_i \cdot \frac{\text{Revenues}_i}{\text{COGS}_i} \; ; \qquad \mu_i^{\text{P}} = \frac{\text{COGS}_i}{\text{COGS}_i + (\delta_i/2)\, q_i^2} \cdot \frac{\text{Revenues}_i}{\text{COGS}_i} \tag{6.6}$$

Aside from the functional form, my approach differs from that of DEU in the strategy to estimate $\partial \log q_i / \partial \log h_i$. DEU estimate θ_i as part of a production function, using time series variation, without making an explicit conduct assumption.

In the extended model with non-constant marginal cost, I cannot estimate markups using DEU's approach, due to the fact that my model assumes a quadratic cost function, and this assumption is generally incompatible with DEU's that firms have a Cobb-Douglas production function.

For this reason, I use a different strategy: q_i is identified (as before) by equation (3.3), which relies on a conduct assumption (Cournot). Next, δ_i is instead identified by a set identification result that is specific to my quadratic cost function.

To estimate the slope of the marginal cost function, we make the following assumptions.

Assumption 4 (*relaxing Assumption 3*)**.** *The marginal cost function respects non-negativity* ($c_i \geq 0$) *and non-decreasing returns to scale* ($\partial c_i / \partial q_i \geq 0$).

These fairly uncontroversial assumptions are lead to a useful partial identification result, that applies to firms for whom observed revenues are at least twice total variables costs.

Proposition 13. *For all firms i such that $p_i q_i > 2 \cdot \text{TVC}_i$, we have* $\delta_i \in \left[0, \dfrac{2 \cdot \text{TVC}_i}{p_i q_i - 2 \cdot \text{TVC}_i}\right]$

Implying that, for a subset of firms, we can actually bound the slope of the marginal cost function (δ_i is unbounded for the remaining firms).

I follow DEU and assume that the parameter controlling the scale elasticity (δ_i) is constant for firms that belong to the same sector (defined as 2-digit NAICS). Then, it has to be the case that, at the sector level, the marginal cost slope is no larger than the lower of the upper bounds defined in Proposition 13. Formally, let us define $\mathcal{S}(i)$, the set of firms that make up the sector to which firm i belongs to, and let us call $\mathcal{F}$ the set of firms to which lemma 13 applies (whose revenues are at least twice total variable costs). Then, the marginal cost slope for all firms in set $\mathcal{S}(i)$ respects:

$$\delta_{\mathcal{S}(i)} \; \leq \; \min_{i' \in \mathcal{S}(i) \cap \mathcal{F}} \; \frac{2 \cdot \text{TVC}_i}{p_i q_i - 2 \cdot \text{TVC}_i} \tag{6.7}$$

I use this upper bound as my sector-level estimate of the slope of the marginal cost function. It makes sense to use an upper bound because, ceteris paribus, assigning a lower value to this parameter yields higher markups that are closer, in absolute value, to the flat marginal cost benchmark (in other words, we have already considered the lower bound $\delta_i = 0 \; \forall i$).

Note that we can measure markups as far back as the 1960s, since the data required is revenues and total variable costs.

The results of this markup estimation are displayed in Figure 11. The upper panel displays the revenue-weighted average markup for this extended model (with non-flat marginal cost) together with the corresponding series computed by DEU. The lower panel shows the cross-sectional correlation between the same two measure (we take the latest datapoint for each firm). I find that the markups from these two different approaches have an extremely high correlation, both in the aggregate time series and in the cross section. They display a similar upward trajectory after 1980, although (after that date) the average markup is higher for my model. In the cross-section, the main differences are: 1) markups from my extended model are bounded below by 1 (by profit maximization, my model cannot produce markups below; 2) at the upper hand of the distribution, markups from my model tend to be a bit lower.

FIGURE 11: MARKUPS: EXTENDED MODEL VS. DEU ESTIMATES (UNTARGETED)

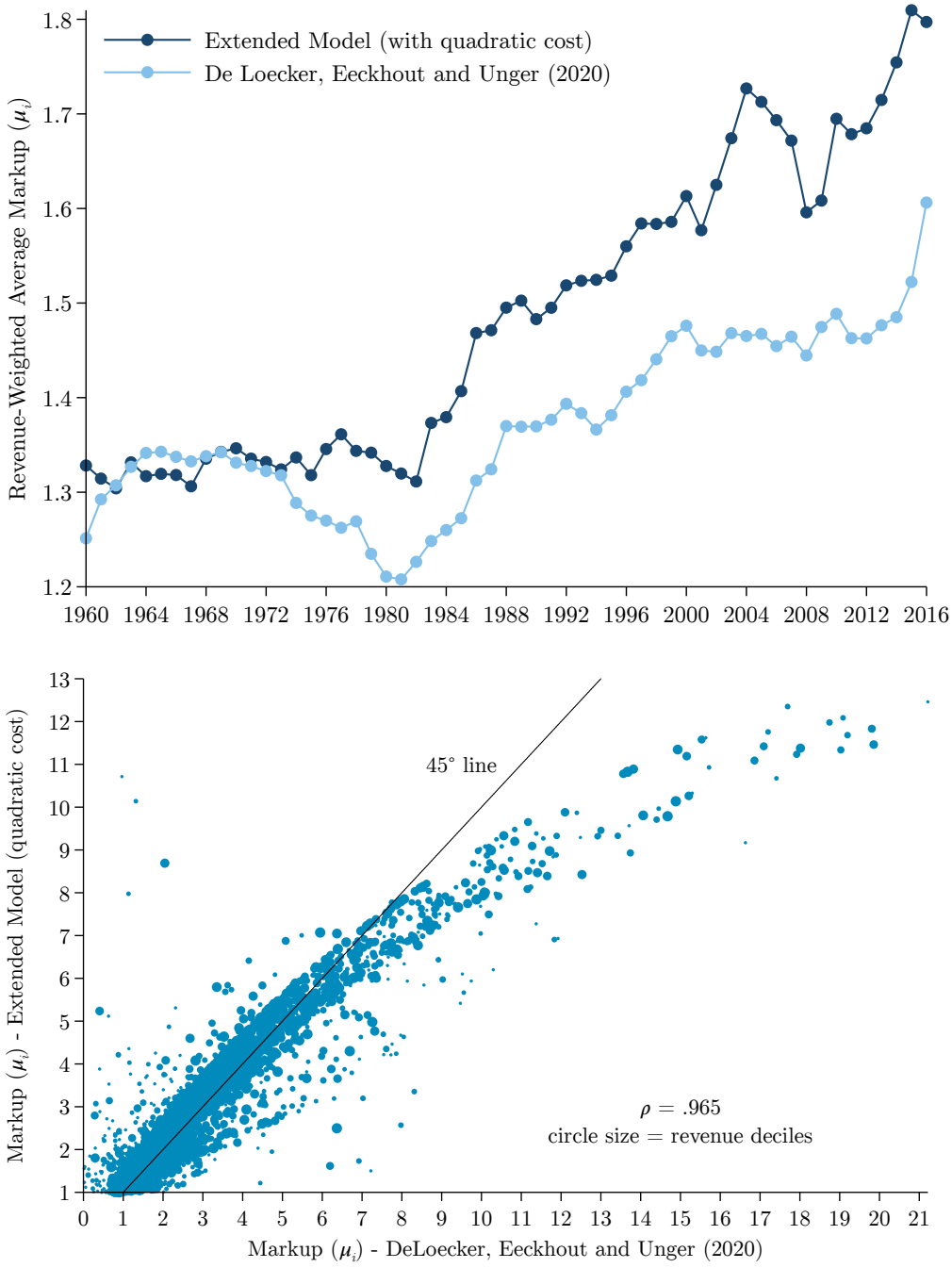


FIGURE NOTES: The figures above compare the markups from the extended model against the estimates of De Loecker, Eeckhout and Unger (2020). The upper panel presents the time series correlation (by averaging markups by year and weighting by revenue). In the lower panel, each circle is a firm (latest available observation), and the data encompasses all public firms in 1960-2016 for which markups can be computed according to both methodologies. The y-coordinate represents model markups, while the the x-coordinate represents the markup estimated by DEU.

This result is to be expected: as Bond et al. (xxx) show, when physical output is proxied using deflated revenues (as in the case of DEU) the supply-side approach to markup measurement (via production function estimation) suffers from a non-identification issue. This non-identification is expected to bias the level (but not necessarily the trend) of the average markup downwards towards one. Because this non-identification issue does not apply to my markup estimates (my approach does not rely on time-series production function estimation), we thus expect the average markup from my model to display a similar trend and a higher average value with respect to DEU's.

This alternative assumption on the marginal cost function yields alternate estimates for $\mathbf{\Delta}$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{c}$, $\mathbf{b}$ and S.

The welfare measurements and counterfactuals for this alternative model are presented in Appendix E.1, and they only differ slightly from the baseline. The profit share of total surplus increases from 16% (in 1996) to 24% (in 2019), while the deadweight loss increases from 6.9% (in 1996) to 11%.

6.2. Private and Foreign Firms

The obvious problem of estimating the model using publicly-traded firms only is that several relevant competitors are excluded from the analysis – namely, private and foreign firms. While the model also includes an outside good (which we labelled *leisure*), which is meant to capture unobserved options, there is nonetheless a clear concern that some of the empirical results obtained thus far may be an artifact of these data limitations.

There is no firm-level data (cosine similarities) for individual private and foreign firms, and thus we can't completely rule out this possibility. However, it is possible to show that the empirical results of Section (5) are highly robust to including private and foreign firms as a fringe of atomistic competitors.

As a robustness check, I use the aggregation result derived in Subsection 2.9 to include non-Compustat companies in the model by adding a set of representative firms that act competitively and whose size reflects the endogenous entry and exit of private and foreign firms in various macro-sectors of the economy.

Because revenue data for Compustat firms includes exported goods and services, in order to implement this extended model with private and foreign firms, I first use data from the Compustat Historical Segments database to construct a measure of the domestic sales share for Compustat companies, and I scale down all of Compustat firms' income statement data (revenues, costs) using the domestic sales share.

The next step is to estimate the share of US final demand served by private and foreign firms which – using equation (3.4) – identifies the equilibrium size of these representative firms. I use the OECD Trade in Value Added (TiVA) dataset to compute US final demand (value added) for 30 macro-sectors. These macro-sectors are briad ISIC v3.1 classifications, an international harmonized classification that is used by the OECD for its TiVA database. I then subtract from this estimate the percentage share of final demand served by US public firms, which I estimate by dividing the domestic sales revenues of Compustat firms by the final demand (computed as gross output, plus imports less exports).

After measuring the output of private and foreign firms in each of these macro-sectors, we must localize these representative firms in the product space. To accomplish this, I use an additional database developed by xxx: using textual descriptions of BEA industries, they construct measures of cosine similarity between products and sectors. I adapt these measures to ISIC v3.1 industries by developing a crosswalk between BEA and ISIC v3.1 classifications and averaging the cosine similarities at the level of the ISIC v3.1 broad industry.

The results from this alternate model including private and foreign firms are provided in Appendix E.2.

My empirical results only change slightly as a consequence of this modification: corporate profits, as a percentage of the surplus produced by public firms, increase from 15% in 1996 to 22.5% in 2019. The deadweight loss increases from 5.9% to 8%.

6.3. Non-Tradable Sectors

Another limitation of the empirical implementation of my model is that it includes sectors or industries that are non-tradable beyond narrow geographies. Consider grocery stores for example: grocery stores in New York do not compete with grocery stores in Los Angeles. The same applies to dialysis clinics, a sector that has seen significant consolidation and has been the object of much empirical research (Cutler et al., 2012; Eliason et al., 2020; Wollmann, 2020). The model, as implemented thus far, does not distinguish the geographic nature of competition in these sectors. Again, we worry that the empirical results that we obtained thus far may be an artifact of the inclusion of non-tradable sectors.

To address this concern, I implement yet another robustness check. I sort broad ISIC v.3.1 into tradable and non-tradable, and estimate yet another version of the model where non-tradable sectors are excluded.

Key welfare metrics from this alternative model are presented in Appendix 6.3, together with the list of ISIC v3.1 macro-sectors, classified into tradable and non-tradable. Over the sample period, the profit share of surplus increases from 18% (in 1996) to 25.4% (in 2019), while the deadweight loss increases from 9.9% to 14%.

6.4. Bertrand Competition

Thus far, we have always assumed that firms compete strategically by choosing output ($\mathbf{q}$). Let us now investigate how both the theoretical and the empirical model change if we instead assume that firms compete à la Bertrand (by setting prices). The model remains tractable as long as we assume that firms face an exogenous marginal cost.

To solve the Bertrand game, it helps to define $\mathbb{D}$, be the diagonal matrix that contains the diagonal entries of $\left(\mathbf{I}+\mathbf{\Sigma}\right)^{-1}$.

Proposition 14. *The Bertrand-Nash equilibrium (with exogenous marginal cost* $\mathbf{c}$*) is* $\mathbf{q}^{\Psi}$ *– the maximizer of the Bertrand potential* $\Psi\left(\cdot\right)$*:*

$$\mathbf{q}^{\Psi} \quad \overset{\text{def}}{=} \quad \arg\max_{\mathbf{q}} \Psi\left(\mathbf{q}\right) \tag{6.8}$$

where the Bertrand potential Ψ *is defined as:*

$$\Psi \overset{\text{def}}{=} \mathbf{q}'\left(\mathbf{b}-\mathbf{c}\right)-\frac{1}{2}\mathbf{q}'\left(\mathbf{I}+\mathbb{D}^{-1}+\mathbf{\Sigma}\right)\mathbf{q} \tag{6.9}$$

the Bertrand equilibrium allocation is:

$$\mathbf{q}^{\Psi} \quad = \quad \left(\mathbf{I}+\mathbb{D}^{-1}+\mathbf{\Sigma}\right)^{-1}\left(\mathbf{b}-\mathbf{c}\right) \tag{6.10}$$

Proof. Appendix G. □

Because $\mathbb{D}^{-1}$ is a diagonal matrix whose diagonal entries are between zero and one, we can see how Bertrand is a more "intense" form of competition than Cournot. In Appendix E, I juxtapose the empirical results from the Cournot and the Bertrand version of the model: empirically, the differences between Cournot and Bertrand appear to be minimal.

6.5. Labor Supply Elasticity

Next, I investigate how my empirical results change if I assume a completely-inelastic labor supply function. By definition, profit as a share of total surplus would be unchanged. The deadweight loss would instead become the total surplus difference between the Cournot equilibrium and the *Resource-Efficient* counterfactual, which we previously defined in Subsection 2.7. As can be seen in Table 3, this welfare difference is

smaller than the deadweight loss. Intuitively, this is because the labor supply (by definition) cannot respond to the removal of the oligopolistic distortions.

I compute this alternative measure of the deadweight loss (the percentage difference in total surplus between Cournot and *Resource Efficient*) over the period 1997-2017. I find that my core empirical results carry through: the level of this "alternative" deadweight loss is 5.2% in 1997, and it increases to 7.9% by 2017. In other words, the level of the deadweight loss is lower if we assume a fixed labor supply (as should be expected), but it increases more sharply (by half) over the 20-year period.

6.6. Complement Goods

In this subsection, I discuss how the model handles complement products. I begin by pointing out that, because the matrix $\mathbf{\Sigma}$ is non-negative by construction, the marginal utility from one unit of product j is always non-increasing in q_i - formally:

$$\frac{\partial^2 S}{\partial q_i \partial q_j} \;=\; -\sigma_{ij} \;\leq\; 0 \quad \forall\, i \neq j \tag{6.11}$$

In light of equation (6.11), it is tempting to jump to the conclusion that all products are by construction substitutes and that no pair of products are complements. That conclusion is, however, incorrect.

To understand why, we need to recall the textbook definition of substitution and complementarity. Two goods (i, j) are[19]:

$$\text{Complements} \quad \text{if} \quad \frac{\partial q_i}{\partial p_j} < 0 \qquad \text{Substitutes} \quad \text{if} \quad \frac{\partial q_i}{\partial p_j} > 0 \tag{6.12}$$

We intuitively expect this derivative to have the opposite sign of that in equation (6.11). In the case of CES, this intuition is correct. In the case of my model, however, this intuition fails. This is a consequence of the fact that the cross-price demand elasticity depends on the inverted matrix $(\mathbf{I} + \mathbf{\Sigma})^{-1}$, not on $\mathbf{\Sigma}$ itself. If the off-diagonal elements of $\mathbf{\Sigma}$ are not equal (here they are not) the off-diagonal elements of $-(\mathbf{I} + \mathbf{\Sigma})^{-1}$ will generally include positive as well as negative elements.[20] This implies that, in the empirical implementation of the model, many producer pairs are strategic complements.

For example, if we compute the vector of cross-price derivatives for car manufacturer General Motors in 2017, we will find that it includes several negative elements (i.e. complements), mostly corresponding to energy and consumer finance companies. This makes sense: intuitively, we expect higher oil prices, loan rates or insurance premia to adversely affect the residual demand for cars.

Hence, despite the property of the model described by equation (6.11), my model does produce strategic complementarity. Indeed, I argue that one of the strengths of the network Cournot model is its ability to produce a rich competitive environment that includes complement goods.

The model even allows for *physical* complements in the model. To do so, we modify the consumer utility to allow for cross-effect between common characteristics

$$U \;=\; \alpha\left(\mathbf{x}'\mathbf{b}^x - \frac{1}{2}\mathbf{x}'\mathbf{x}\right) + (1-\alpha)\left(\mathbf{y}'\mathbf{b}^y - \frac{1}{2}\mathbf{y}'\mathbf{y}\right) - H \tag{6.13}$$

[19] Usually, these derivatives refer to the *Hicksian* demand, to ensure that $\text{sign}\,(\partial q_i/\partial p_j) = \text{sign}\,(\partial q_j/\partial p_i)$ (see Kreps, 2012, section 11.6). This distinction can be ignored for my model, since the sign of this derivative is the same for the Marshallian and the Hicksian demand.

[20] It is fairly easy to come up with examples. Consider for instance three goods (1,2,3) and three common characteristics (A,B,C). If goods 1 and 3 load entirely on characteristic A and C respectively, and good 2 loads equally on all three characteristics, then it can be verified that goods 1 and 3 are strategic complements.

Then the matrix of inverse demand derivatives $\mathbf{\Sigma}$ takes the form

$$\mathbf{\Sigma} \overset{\text{def}}{=} \alpha\left(\mathbf{A}'\mathbf{M}\mathbf{A} - \mathbf{I}\right) \tag{6.14}$$

then, provided that $m_{kk'} < 0$ for some (k, k'), $\mathbf{\Sigma}$ will generally also include negative terms, which correspond to physical complements (an increase in the supply of good j increases the marginal utility of a unit of good i). However, because the cross-price elasticities are no longer written in terms of cosine similarities, the HP data is no longer sufficient to take the model to the data: we need to have data on characteristic loadings (that is, we need to observe the matrix $\mathbf{A}$, not just $\mathbf{A}'\mathbf{A}$).

6.7. Strategic Complementarities (and Oligopsony) in Input Markets

Next, we consider ways in which the model can be extended to allow for strategic interactions in input markets as well as product markets. First, notice that the more general model, with generic cost function $\mathbf{h}(\mathbf{q})$, already allows firms to interact strategically in input markets. There are various ways this could be accomplished parametrically. For example, a simple way to introduce technological spillovers in this model that preserves all of the closed-form solutions is to allow the cost function to incorporate cross-terms:

$$h_i \;=\; f_i + c_i^0 q_i + \frac{1}{2}\sum_{j=1}^{n} \delta_{ij} q_i q_j \tag{6.15}$$

that is - to allow the matrix $\mathbf{\Delta}$ to be non-diagonal: this allows marginal cost of firm i is to be affected by firm j's production scale.

Another extension that would allow for further strategic interactions in input markets is to introduce monopsony power: consider for example the following partial equilibrium variant of the model. Suppose that There are $\ell = 1, .., L$ types of workers (sorted by geography, occupation, etc...) that act as wage-takers. Firms produce using a Leontief production function, so that labor inputs vector $\mathbf{h}$ is required to produce output vector $\mathbf{q}$:

$$\mathbf{h} \;=\; \mathbf{T}\mathbf{q} \tag{6.16}$$

Each dimension of $\mathbf{h}$ corresponds to a type of worker and each dimension of $\mathbf{q}$ is (again) a firm. Assume that the labor supply function is linear, so that ℓ's wage is:

$$\mathcal{W}_\ell \;=\; \xi_\ell h_\ell \tag{6.17}$$

The total cost of production, for firm i, is equal to

$$\text{TC}_i \;=\; \sum_{\ell=1}^{L} \mathcal{W}_\ell t_{\ell i} q_i \;=\; \sum_{\ell=1}^{L}\sum_{j=1}^{n} \xi_\ell t_{\ell j} t_{\ell i} q_i q_j \tag{6.18}$$

Also, assume that the consumer utility is given by:

$$U(\mathbf{x}, \mathbf{y}, \mathcal{O}) \quad \overset{\text{def}}{=} \quad \alpha\left(\mathbf{x}'\mathbf{b}^x - \frac{1}{2}\cdot\mathbf{x}'\mathbf{x}\right) + (1-\alpha)\left(\mathbf{y}'\mathbf{b}^y - \frac{1}{2}\cdot\mathbf{y}'\mathbf{y}\right) - \mathcal{O} \tag{6.19}$$

where $\mathcal{O}$ is an outside good. Then the vector of profits is equal to:

$$\pi \;=\; \text{diag}(\mathbf{q})\left(\mathbf{p} - \mathbf{T}'\mathbf{\Xi}\mathbf{T}\mathbf{q}\right) \tag{6.20}$$

The equilibrium Cournot equilibrium with oligopoly *and* oligopsony is then:

$$\mathbf{q}^\Phi \;=\; \left(2\mathbf{I} + \mathbf{\Sigma} + \mathbf{T}'\mathbf{\Xi}\mathbf{T}\right)^{-1}\mathbf{b} \tag{6.21}$$

There is unfortunately no data (yet) to estimate this extended model, and thus its quantification is left for future research.

6.8. Limitations and Future Work

This model (like every other model) has limitations and leaves out certain aspects of market power that might be relevant to the current debate on antitrust policy.

For example, one important assumption that I make in my model (in order to preserve tractability) is that all firms are final goods firms. In other words, input-output linkages between individual firms are not part of the model. This might lead to underestimating the welfare costs of oligopoly, if input-output linkages result in double marginalization. One interesting extension of the model would be to modify the firm's production function to allow firms to use other firm's output as inputs. In order to take this extended model to the data, granular input-output data on firm-to-firm relationships would be required.

Another important restriction of my model is that it treats the firms' position in the product characteristics space as fixed. The assumption that product characteristics are exogenous is standard in the demand estimation literature (Berry et al., 1995; Nevo, 2001); however, it makes policy simulations less robust in the long-run. This is because, given enough time, firms may be able to endogenously change their product portfolios. Some recent work in the IO literature (see Fan, 2013; Wollmann, 2018) has considered endogenous product characteristics. If product characteristics data can be obtained (HP only provide cosine similarities), another interesting direction for future research is to generalize my model by endogenizing the firms' position in the characteristics space.

7. Conclusions

In this study, I have presented a new general equilibrium model of oligopolistic competition with hedonic demand and differentiated products, with the objective of measuring the welfare consequences of rising oligopoly in the United States from 1996 to 2019. To estimate my model, I used a data set (recently developed by Hoberg and Phillips, 2016) of product cosine similarities that covers all public firms in the United States on a yearly basis. Through the lens of my model, these similarity scores are used to retrieve the cross-price elasticity of demand for every pair of publicly traded firms.

My measurements suggest that oligopoly has a considerable and growing effect on aggregate welfare. In particular, I estimate that, if all publicly traded firms were to behave as atomistic competitors, the total surplus produced by this set of companies would increase by 12.7 percentage points. Consumer welfare would increase even more dramatically—by 45%—as surplus is completely reallocated from producers to consumers. I find that a large share of the deadweight loss caused by oligopoly (8.9 percentage points) can be attributed to resource misallocation—that is, a significant share of the deadweight losses could theoretically be recovered by a benevolent social planner, even if we assumed labor to be inelastically supplied.

I also find evidence that the deadweight losses increases significantly in the presence of collusion: consolidating firm ownership in the hands of one owner, who induces firms to coordinate, would depress aggregate surplus by an additional 20.3 percentage points. Consumer surplus would drop even more, with a projected decrease of 57.3%. Overall, my analysis of firm-level data suggests that there is evidence of sizable welfare distortions due to oligopoly power.

By mapping my model to firm-level data for every year between 1996 and 2019, I find that, while both the profits earned by U.S. public corporations and the corresponding consumer surplus have increased over this period, profits have increased at a significantly faster pace: consequently, the share of surplus appropriated by firms in the form of profits has increased substantially (from 15.4% to 22.6%). Consistent with this finding, I estimate that the welfare costs of oligopoly, defined as the percentage increase in surplus that is obtained

by moving to the competitive outcome, have increased (from 7.8% to 12.7%). Overall, my estimates are consistent with the hypothesis that the observed secular trends in markups and concentration have resulted in an increased welfare loss, particularly at the expense of the consumer.

The model also disentangle the role of two factors that have contributed to secular rise of markups (previously documented by De Loecker, Eeckhout and Unger, 2020): productivity (adjusted for quality) and product market centrality (which captures the availability of substitute products). The first drives the reallocation of market shares towards high-markup firms, while the latter drives the within-firm growth component. This finding provides a proximate explanation for the rise of markups, and imposes significant constraints to theories that seek to provide a deep explanation: a credible theoretical account for the secular rise of markups should accommodate both of these forces.

The model allows me to compute a number of novel counterfactuals that are relevant to competition policy, and to shed additional light on these ongoing trends.

This paper contributes—both methodologically and empirically—to a growing literature in macroeconomics and finance that is devoted to incorporating heterogeneity, imperfect competition and Industrial Organization methods in general equilibrium models. In particular, it shows that combining firm financials with measures of similarity based on natural-language processing of regulatory filings offers a promising avenue to model product differentiation and imperfect substitutability at the macroeconomic level: it affords the opportunity to impose a less arbitrary structure on the degree of substitution across sectors and firms.

In addition to the theory contribution, this paper provides measurements that align with a growing body of empirical work on rising market power (De Loecker, Eeckhout and Unger, 2020) and the anti-competitive effects of merger activity (Cunningham et al., 2018; Wollmann, 2019).

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The network origins of aggregate fluctuations," *Econometrica*, 80, 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2017): "Microeconomic origins of macroeconomic tail risks," *American Economic Review*, 107, 54–108.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020): "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 135, 645–709.

Axtell, R. L. (2001): "Zipf Distribution of US Firm Sizes," *Science*, 1818–1820.

Baker, J. B., and T. F. Bresnahan (1985): "The gains from merger or collusion in product-differentiated industries," *The Journal of Industrial Economics*, 427–444.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's who in networks. Wanted: The key player," *Econometrica*, 74, 1403–1417.

Baqaee, D. R., and E. Farhi (2020): "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying technology spillovers and product market rivalry," *Econometrica*, 81, 1347–1393.

Boller, L., and F. S. Morton (2020): "Testing the theory of common stock ownership,"Technical report, National Bureau of Economic Research.

Bonacich, P. (1987): "Power and centrality: A family of measures," *American journal of sociology*, 92, 1170–1182.

Carvalho, V. M. (2014): "From micro to macro via production networks," *Journal of Economic Perspectives*, 28, 23–48.

Carvalho, V. M., M. Nirei, Y. Saito, and A. Tahbaz-Salehi (2020): "Supply chain disruptions: Evidence from the great east japan earthquake," *The Quarterly Journal of Economics, forthcoming*.

Carvalho, V. M., and A. Tahbaz-Salehi (2019): "Production networks: A primer," *Annual Review of Economics*, 11, 635–663.

Chen, H., L. Cohen, and D. Lou (2016): "Industry window dressing," *The Review of Financial Studies*, 29, 3354–3393.

Covarrubias, M., G. Gutiérrez, and T. Philippon (2020): "From Good to Bad Concentration? US Industries over the past 30 years," *NBER Macroeconomics Annual*, 34, 1–46.

Cunningham, C., F. Ederer, and S. Ma (2018): "Killer acquisitions," *Available at SSRN 3241707*.

Cutler, D., L. S. Dafny, and C. J. Ody (2012): "How does competition impact the quality of health care? a case study of the us dialysis industry."

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*.

De Loecker, J., and P. T. Scott (2016): "Estimating market power: Evidence from the US Brewing Industry," working paper, National Bureau of Economic Research.

De Loecker, J., and F. Warzynski (2012): "Markups and firm-level export status," *The American Economic Review*, 102, 2437–2471.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014): "The role of entrepreneurship in US job creation and economic dynamism," *Journal of Economic Perspectives*, 28, 3–24.

Edmond, C., V. Midrigan, and D. Y. Xu (2018): "How costly are markups?" working paper, National Bureau of Economic Research.

Einav, L., and J. Levin (2010): "Empirical industrial organization: A progress report," *Journal of Economic Perspectives*, 24, 145–62.

Eliason, P. J., B. Heebsh, R. C. McDevitt, and J. W. Roberts (2020): "How acquisitions affect firm behavior and performance: Evidence from the dialysis industry," *The Quarterly Journal of Economics*, 135, 221–267.

Epple, D. (1987): "Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products," *Journal of political economy*, 95, 59–80.

Ewens, M., R. H. Peters, and S. Wang (2019): "Measuring intangible capital with market prices,"Technical report, National Bureau of Economic Research.

Fan, Y. (2013): "Ownership consolidation and product characteristics: A study of the US daily newspaper market," *American Economic Review*, 103, 1598–1628.

Fruchterman, T. M., and E. M. Reingold (1991): "Graph drawing by force-directed placement," *Software: Practice and experience*, 21, 1129–1164.

Goeree, M. S. (2008): "Limited information and advertising in the US personal computer industry," *Econometrica*, 76, 1017–1074.

Grullon, G., Y. Larkin, and R. Michaely (2019): "Are US Industries Becoming More Concentrated?" *Review of Finance*, 23, 697–743.

Hoberg, G., and G. Phillips (2016): "Text-based network industries and endogenous product differentiation," *Journal of Political Economy*, 124, 1423–1465.

Hopenhayn, H. A. (1992): "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1127–1150.

Kahle, K. M., and R. M. Stulz (2017): "Is the US public corporation in trouble?" *Journal of Economic Perspectives*, 31, 67–88.

Katz, L. (1953): "A new status index derived from sociometric analysis," *Psychometrika*, 18, 39–43.

Khan, L. (2018): "The new Brandeis movement: America's antimonopoly debate."

Kreps, D. M. (2012): *Microeconomic Foundations I: Choice and Competitive Markets*: Princeton University Press.

Kwon, S. Y., Y. Ma, and K. Zimmermann (2021): "100 Years of Rising Corporate Concentration," *Available at SSRN 3936799*.

Lancaster, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132–157.

Monderer, D., and L. S. Shapley (1996): "Potential games," *Games and economic behavior*, 14, 124–143.

Moulton, B. R. et al. (2001): "The expanding role of hedonic methods in the official statistics of the United States," *BEA Papers*, 14.

Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307–342.

Phillips, G. M., and A. Zhdanov (2013): "R&D and the Incentives from Merger and Acquisition Activity," *The Review of Financial Studies*, 26, 34–78.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, 82, 34–55.

Syverson, C. (2004): "Market structure and productivity: A concrete example," *Journal of Political Economy*, 112, 1181–1222.

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

Ushchev, P., and Y. Zenou (2018): "Price competition in product variety networks," *Games and Economic Behavior*, 110, 226–247.

Wollmann, T. G. (2018): "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *American Economic Review*, 108, 1364–1406.

——— (2019): "Stealth consolidation: Evidence from an amendment to the Hart-Scott-Rodino Act," *American Economic Review: Insights*, 1, 77–94.

——— (2020): "How to get away with merger: Stealth consolidation and its real effects on US healthcare,"Technical report, National Bureau of Economic Research.

Online Appendices

Product Differentiation and Oligopoly: a Network Approach

Bruno Pellegrino, UMD

A. Derivation of the Cournot Potential

In the Network Cournot model, each firm i chooses its own output level to maximize its own profit by taking as given the output of every other firm:

$$q_i^* = \arg\max_{q_i} \pi\left(q_i; \bar{\mathbf{q}}_{-i}\right) \tag{A.1}$$

where $\bar{\mathbf{q}}_{-i}$ is the vector of output for every firm except i. The upper bar sign $\bar{\cdot}$ indicates that i treats the quantity supplied by other firms as fixed. The system of first order conditions for this problem is

$$0 = \ b_i - c_i\left(q_i\right) - 2q_i - \alpha \sum_{j \neq i} \left(\mathbf{a}_i' \mathbf{a}_j\right) \bar{q}_j \tag{A.2}$$

which can be expressed, in vector form, as:

$$0 = \mathbf{b} - \mathbf{c} - 2\mathbf{q} - \mathbf{\Sigma}\bar{\mathbf{q}} \tag{A.3}$$

This system of reaction functions defines a vector field $\mathbf{q}\left(\bar{\mathbf{q}}\right)$ which represents the firms' best response as a function of every other firms' strategy. To find the Cournot-Nash Equilibrium, we look for the fixed point $\mathbf{q}^*$ such that $\mathbf{q} = \bar{\mathbf{q}} = \mathbf{q}^{\Phi}$. Plugging this inside the equation above yields the first order condition that is needed to maximize the potential function $\Phi\left(\mathbf{q}\right)$:

$$0 = \mathbf{b} - \mathbf{c}\left(\mathbf{q}^{\Phi}\right) - \left(2\mathbf{I} + \mathbf{\Sigma}\right)\mathbf{q}^{\Phi} \tag{A.4}$$

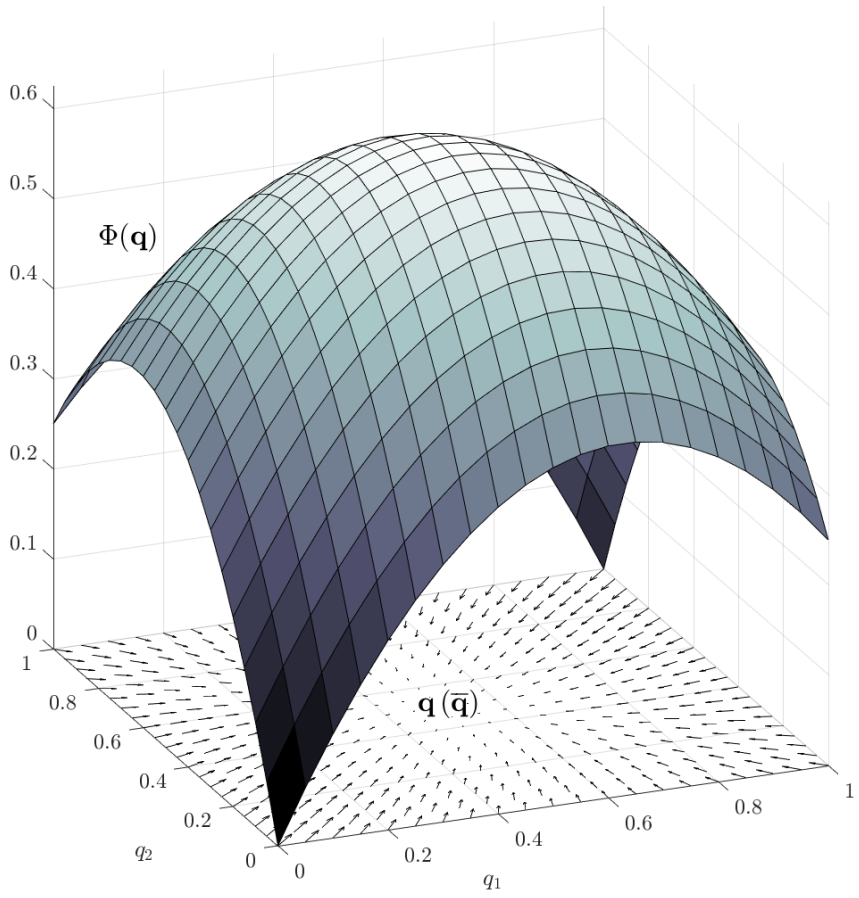
which clarifies why the maximizer of the potential function solves the Network Cournot game. The potential $\Phi\left(\mathbf{q}\right)$ is then obtained as the solution to the following system of partial differential equations

$$\nabla\, \Phi\left(\mathbf{q}\right) \ = \ \mathbf{b} - \mathbf{c}\left(\mathbf{q}\right) - \left(2\mathbf{I} + \mathbf{\Sigma}\right)\mathbf{q} \tag{A.5}$$

which equates the gradient of the potential function to the linear system of Cournot reaction functions.

The relationship between the potential and the Cournot-Nash equilibrium is represented graphically, for the two-firm case, in Figure 12. The arrows represent the vector field defined by the firms' reaction functions. The potential function is defined to be the scalar-valued function whose gradient coincides with this vector field. A game is a potential game if the vector field defined by the players' reaction functions is a conservative field - that is, if it is the gradient of some scalar function. We call that function the game's potential.

Figure 12: Graphing the Cournot Potential for the Two-firm Case



B. Volumetric Normalization and Idiosyncratic Characteristics

In this appendix I discuss two distinctive features of my demand system, namely: 1) the volumetric normalizations that I apply to products and characteristics; 2) the presence of idiosyncratic characteristics. First, I will show mathematically that the utility function in equation (2.9) is actually quite general, in the sense that we can derive it by starting from a much more form by applying some welfare-invariant normalizations. Second, I will show that, while idiosyncratic characteristics do restrict the model somewhat theoretically, they also make the demand system, from an empirical perspective more *flexible*. What I mean by more flexible is that it will have an additional degree of freedom to match empirical cross-price demand elasticities.

B.1. Volumetric Normalization

Let us re-write the utility function with maximum flexibility:

$$U(\mathbf{x},\mathbf{y},H) \quad \stackrel{\text{def}}{=} \quad \tau_1 \mathbf{x}'\mathbf{b}^x - \frac{\tau_2}{2}\cdot \mathbf{x}'\mathbf{M}^x\mathbf{x} + \frac{\tau_3}{2}\mathbf{y}'\mathbf{b}^y - \frac{\tau_4}{2}\mathbf{y}'\mathbf{M}^y\mathbf{y} - \tau_5 H \tag{B.1}$$

$$\mathbf{x} = \mathbf{A}^x\mathbf{q} \qquad \text{and} \qquad \mathbf{y} = \mathbf{A}^y\mathbf{q} \tag{B.2}$$

where $\mathbf{M}^x$ is a diagonalizable matrix, $\mathbf{M}^y$ and $\mathbf{A}^y$ are diagonal matrices, and where we do *not* require $\|\mathbf{a}^x_k\| = 1$ and $\mathbf{a}^y_{ii} = 1$, but we do require that:

$$\left\| \left[\left(\frac{\tau_2}{\alpha}\mathbf{M}^x \right)^{-\frac{1}{2}} \mathbf{A}^x \right]_i \right\| \;=\; \left[\left(\frac{\tau_4}{1-\alpha}\mathbf{M}^y \right)^{-\frac{1}{2}} \mathbf{A}^y \right]_{ii} \tag{B.3}$$

I show that, up to some change in volumetric units, we can rewrite the above utility in the following form

$$U(\mathbf{x},\mathbf{y},H) \quad \stackrel{\text{def}}{=} \quad \alpha\left(\mathbf{x}'\mathbf{b}^x - \frac{1}{2}\mathbf{x}'\mathbf{x}\right) - (1-\alpha)\left(\mathbf{y}'\mathbf{b}^y - \frac{1}{2}\mathbf{y}'\mathbf{y}\right) - H \tag{B.4}$$

$$\mathbf{x} = \mathbf{A}\mathbf{q}, \qquad \|\mathbf{a}_k\| = 1 \;\forall k \qquad \text{and} \qquad \mathbf{y} = \mathbf{q} \tag{B.5}$$

First, notice that the following utility

$$U(\tilde{\mathbf{x}},\tilde{\mathbf{y}},H) \;=\; \frac{\tau_1}{\tau_5}\mathbf{x}'\mathbf{b}^x - \frac{\tau_2}{2\tau_5}\cdot \mathbf{x}'\mathbf{M}^x\mathbf{x} + \frac{\tau_3}{\tau_5}\mathbf{y}'\mathbf{b}^y - \frac{\tau_4}{2\tau_5}\mathbf{y}'\mathbf{M}^y\mathbf{y} - H \tag{B.6}$$

describes the exact same preferences as equation (B.1), therefore we can use it instead at no loss. Next, because $\mathbf{M}^x$ is diagonalizable, we can write it as:

$$\mathbf{M}^x \;=\; \mathbf{M}_1'\mathbf{M}_2\mathbf{M}_1 \tag{B.7}$$

where $\mathbf{M}_2$ is diagonal. Let us now define the following alternative vectors of characteristics:

$$\tilde{\mathbf{x}} \;\stackrel{\text{def}}{=}\; \sqrt{\frac{\tau_2}{\alpha\tau_5}}\mathbf{M}_2^{\frac{1}{2}}\mathbf{M}_1\mathbf{x} \tag{B.8}$$

$$\tilde{\mathbf{y}} \;\stackrel{\text{def}}{=}\; \sqrt{\frac{\tau_4}{(1-\alpha)\,\tau_5}}\,(\mathbf{M}^y)^{\frac{1}{2}}\,\mathbf{y} \tag{B.9}$$

We can now re-write the utility in terms of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$:

$$U(\tilde{\mathbf{x}},\tilde{\mathbf{y}},H) \;=\; \frac{\tau_1\sqrt{\alpha}}{\sqrt{\tau_2\tau_5}}\tilde{\mathbf{x}}'\mathbf{M}_1^{-1}\mathbf{M}_2^{-\frac{1}{2}}\mathbf{b}^x - \frac{\alpha}{2}\cdot\tilde{\mathbf{x}}'\tilde{\mathbf{x}} + \frac{\tau_3\sqrt{1-\alpha}}{\sqrt{\tau_2\tau_5}}\tilde{\mathbf{y}}'\,(\mathbf{M}^y)^{-\frac{1}{2}}\,\mathbf{b}^y - \frac{1-\alpha}{2}\tilde{\mathbf{y}}'\tilde{\mathbf{y}} - H \tag{B.10}$$

Next, define:

$$\tilde{\mathbf{b}}^x \stackrel{\text{def}}{=} \frac{\tau_1}{\sqrt{\alpha\tau_2\tau_5}}\mathbf{M}_1^{-1}\mathbf{M}_2^{-\frac{1}{2}}\mathbf{b}^x \tag{B.11}$$

$$\tilde{\mathbf{b}}^y \stackrel{\text{def}}{=} \frac{\tau_3}{\sqrt{(1-\alpha)\,\tau_2\tau_5}}\left(\mathbf{M}^y\right)^{-\frac{1}{2}}\mathbf{b}^y \tag{B.12}$$

and re-write the utility function in the concise desired form:

$$U\left(\tilde{\mathbf{x}},\tilde{\mathbf{y}},H\right) = \alpha\left(\tilde{\mathbf{x}}'\tilde{\mathbf{b}}^x - \frac{1}{2}\cdot\tilde{\mathbf{x}}'\tilde{\mathbf{x}}\right) + (1-\alpha)\left(\tilde{\mathbf{y}}'\tilde{\mathbf{b}}^y - \frac{1}{2}\tilde{\mathbf{y}}'\tilde{\mathbf{y}}\right) - H \tag{B.13}$$

$$\text{with} \quad \hat{\mathbf{x}} = \tilde{\mathbf{A}}^x\mathbf{q} \quad \text{and} \quad \tilde{\mathbf{y}} = \tilde{\mathbf{A}}^y\mathbf{q} \tag{B.14}$$

By construction, $\tilde{\mathbf{A}}^x$ and $\tilde{\mathbf{A}}^y$ are equal to:

$$\tilde{\mathbf{A}}^x \stackrel{\text{def}}{=} \sqrt{\frac{\alpha\tau_5}{\tau_2}}\mathbf{M}_1^{-1}\mathbf{M}_2^{-\frac{1}{2}}\mathbf{A}^x \quad \text{and} \quad \tilde{\mathbf{A}}^y \stackrel{\text{def}}{=} \sqrt{\frac{(1-\alpha)\,\tau_5}{\tau_4}}\left(\mathbf{M}^y\right)^{-\frac{1}{2}}\mathbf{A}^y \tag{B.15}$$

Note that $\tilde{\mathbf{A}}^y$ is diagonal, and that restriction (B.3) implies:

$$\|\tilde{\mathbf{a}}_i^x\| \equiv \tilde{a}_{ii}^y \tag{B.16}$$

which is the reason why such restriction equates to imposing that the relative loading on common v/s idiosyncratic characteristics is the same for all goods and thus can be captured by using a single α for all goods.

Also, note that have have not yer imposed that $\tilde{\mathbf{a}}_i^x$ have unitary norm, and thus $\tilde{\mathbf{A}}^y$ is not necessarily an identity matrix. Therefore the final step is to show that the restriction

$$\|\tilde{\mathbf{a}}_i^x\| = a_{ii}^y = 1 \quad \forall i \tag{B.17}$$

is without loss of generality. We change the volumetric units of our goods by defining:

$$\hat{\mathbf{q}} \stackrel{\text{def}}{=} \tilde{\mathbf{A}}^y\mathbf{q} \tag{B.18}$$

so that

$$\tilde{\mathbf{x}} = \hat{\mathbf{A}}\hat{\mathbf{q}} \quad \text{with} \quad \hat{a}_{ij} \stackrel{\text{def}}{=} \frac{\tilde{a}_{ij}}{a_{ii}^y} \tag{B.19}$$

this implies finally that

$$\|\hat{\mathbf{a}}_i^x\| = 1\ \forall i \quad \text{and} \quad \tilde{\mathbf{y}} = \hat{\mathbf{q}} \tag{B.20}$$

implying that we can re-write the utility function in terms of $\hat{\mathbf{q}}$

$$U\left(\hat{\mathbf{q}},H\right) = \alpha\left(\hat{\mathbf{q}}'\hat{\mathbf{A}}\tilde{\mathbf{b}}^x - \frac{1}{2}\cdot\hat{\mathbf{q}}'\hat{\mathbf{A}}'\hat{\mathbf{A}}\hat{\mathbf{q}}\right) + (1-\alpha)\left(\hat{\mathbf{q}}'\tilde{\mathbf{b}}^y - \frac{1}{2}\hat{\mathbf{q}}'\hat{\mathbf{q}}\right) - H \tag{B.21}$$

where $\hat{\mathbf{A}}'\hat{\mathbf{A}}$ is now by definition a cosine similarity matrix.

Notice that we still have one degree of freedom left, which is the choice of the unit of measurements of labor. That degree of freedom allows us to impose labor as the numeráire (change the units of labor so that one labor unit = 1\$).

B.2. The Role of Idiosyncratic Characteristics

In this Appendix, I further clarify the role of idiosyncratic characteristics in GHL demand. As previously mentioned, from a *theoretical* point of view, the presence of idiosyncratic characteristics is somewhat redundant, because we can just "fold" them inside the matrix of common characteristics. The case $\alpha < 1$ is actually less general, because it *imposes* the presence of idiosyncratic characteristics, and it also imposes that all products load *equally* on their respective idiosyncratic characteristics. It is thus the presence of idiosyncratic characteristics (and not the volumetric normalizations) that restricts the model.

Now, I argue, while the idiosyncratic characteristics do limit the generality of the model from a *theoretical perspective,* from an *empirical perspective* they provide a much valuable additional degree of freedom. The key to this important (non-trivial) distinction is the identifying assumption, made explicit in sub-section 3.5, that only the *common* characteristics are mapped to the Hoberg-Phillips words/characteristics space. Idiosyncratic characteristics are instead *unobserved*.

This has important implications for the overall magnitude of the elasticities: by introducing an additional set of unobserved idiosyncratic characteristic that is not contained inside HP's similarity data, we have an additional degree of freedom that allows to better control the overall magnitude of the cross-price elasticities.

If we exclude the presence of idiosyncratic characteristics ($\alpha = 1$), the price-quantity derivative $\partial p_i / \partial q_j$ is just the cosine similarity between i and j:

$$\partial p_i / \partial q_j \;=\; \mathbf{a}_i' \mathbf{a}_j \tag{B.22}$$

this implies that the cross-price elasticities are entirely pinned down by HP's cosine similarities. In other way, there is no way to control how high or low cross-price elasticities are overall: we must fully trust not only the relative magnitude of HP's cosine similarities, but also their *level*. While I know for a fact that HP's dataset was constructed with great care and has been validated in every possible way, I think that would be too restrictive for most observers.

When we add unobserved idiosyncratic unobserved characteristics, the price-quantity derivative for $i \neq j$ becomes

$$\partial p_i / \partial q_j \;=\; \alpha\, \mathbf{a}_i' \mathbf{a}_j \tag{B.23}$$

By calibrating the parameter α, I can make sure that the magnitude of the recovered elasticities is broadly consistent with previous estimates from the IO literature.

C. Nash-Cournot Equilibrium and Network Centrality

In this appendix, I provide additional details on the relationship between equation (2.26), which describes the equilibrium size of firms in my model, and the measures of network centrality developed by Katz (1953) and Bonacich (1987), which are widely used in the social networks literature.

Under the restriction that the cost function is quadratic, the game played by the firms from Section 2 is a type of linear-quadratic network game. Ballester, Calvó-Armengol and Zenou (2006, henceforth BCZ) show that players' equilibrium actions and payoffs in this class of games depends on their centrality in the network.

In the game played the firms that populate by model, the adjacency matrix of the network over which the game is played, is given by the matrix $(-\mathbf{\Sigma})$. This matrix appears in the quadratic term of all the welfare functions (profits, total surplus and the Cournot potential).

Before discussing how the linkage extends to my model, I am going to formally define the metric of centrality.

Definition 11 (Katz-Bonacich Centrality)**.** For a weighted network with adjacency matrix $\mathbf{M}$, we define the vector of centralities $\mathbf{k}$, a function of M with parameters $(\eta, \mathbf{g})$:

$$\begin{aligned} \mathbf{k}\left(\mathbf{M}; \eta, \mathbf{g}\right): \qquad \mathbf{k} &= \eta\, \mathbf{M}\mathbf{k} + \mathbf{g} \\ &= \left(\mathbf{I} - \eta \mathbf{M}\right)^{-1} \mathbf{g} \end{aligned} \tag{C.1}$$

The Katz-Bonacich centrality is defined recursively: a node receives a higher centrality score the higher is the centrality of the nodes it is connected to.

The Nash-Bonacich linkage extends to my model. Suppose that the vector of marginal costs $(\mathbf{c})$ is exogenous: then, the Cournot-Nash equilibrium allocation of the model presented in Section 2 (equation 2.26) can be easily verified to coincide with the vector of Katz-Bonacich centralities, with parametrization $\left(\frac{1}{2},\ \frac{1}{2}\left(\mathbf{b} - \mathbf{c}\right)\right)$:

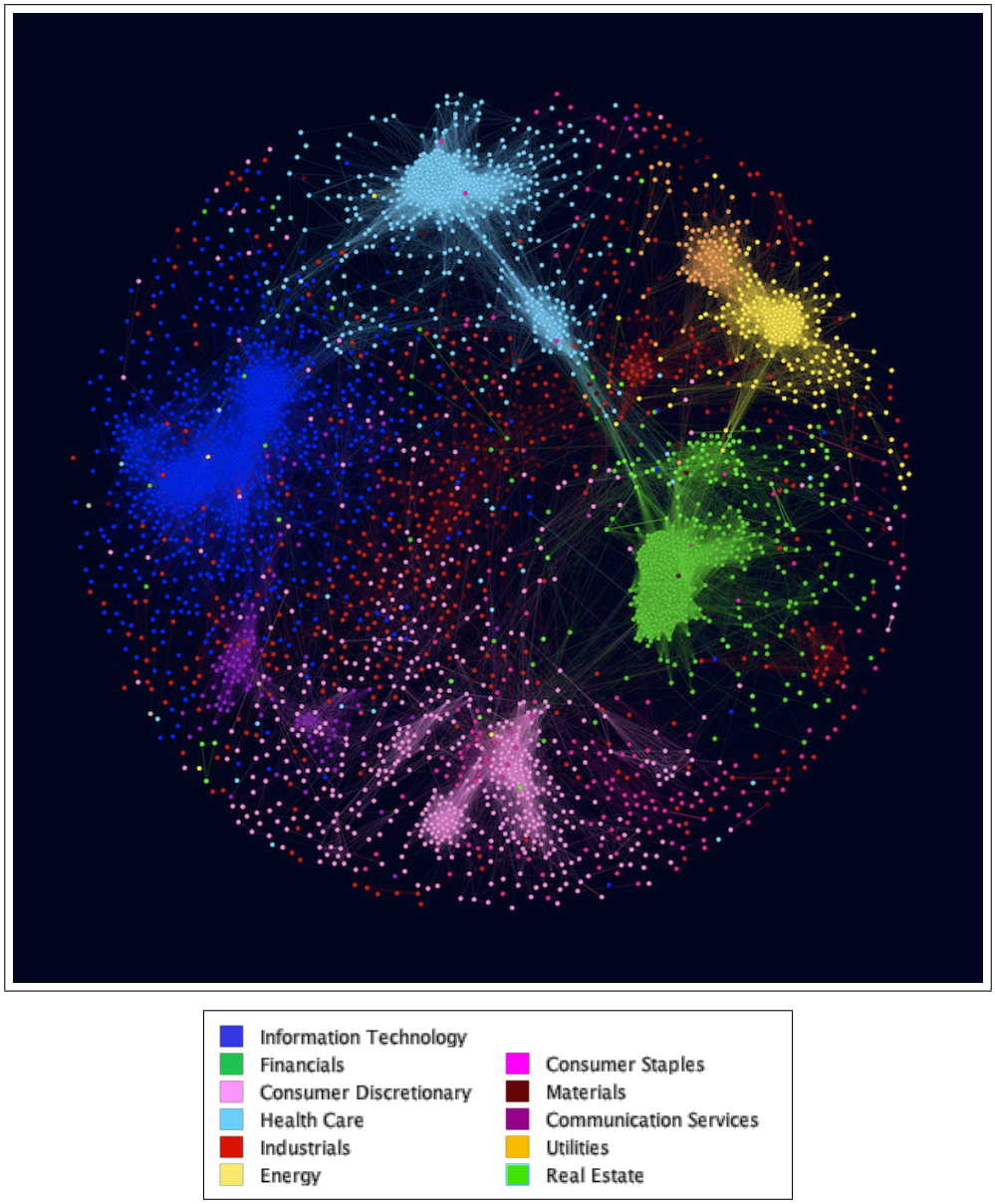
$$\mathbf{q}^{\Phi} \ \equiv\ \mathbf{k}\left(-\mathbf{\Sigma};\ \frac{1}{2},\ \frac{1}{2}\left(\mathbf{b} - \mathbf{c}\right)\right) \tag{C.2}$$

Notice that the Cournot game is played over a *negatively-weighted* network (the adjacency matrix is $-\mathbf{\Sigma}$). The consequence is that the interpretation of $\mathbf{q}^{\Phi}$, as a measure of centrality, is reversed: a higher Katz-Bonacich centrality actually reflects a more peripheral position in the network with positive weights $\mathbf{\Sigma}$.

D. Independent Validation of the Hoberg & Phillips Dataset

In this appendix, I validate independently the text-based product similarity measures of Hoberg and Phillips (2016). In the figure below I produce a graph similar to that of Figure 3, while coloring different nodes according to the respective firm's GIC economic sector. The figure shows that there is significant overlap between the macro clusters of the network of product similarity and the broad GIC sectors. To produce this visualization, the dimensionality of data has been reduced from 61,000 to 2; yet, the overlap is nonetheless very clearly visible. The GIC sectors were *not* targeted in producing this graph.

FIGURE 13: VISUALIZATION OF THE PRODUCT SPACE (ALTERNATE COLORING)



E. Model Extensions: Empirical Results

E.1. Quadratic Cost Function (decreasing returns to scale)

Figures 14 and 15 replicate, respectively, Figures 7 and 8 for the extended model where firms have a non-flat marginal cost function (discussed in 6.2).

Figure 14: Profit as % of Total Surplus (1996-2019)

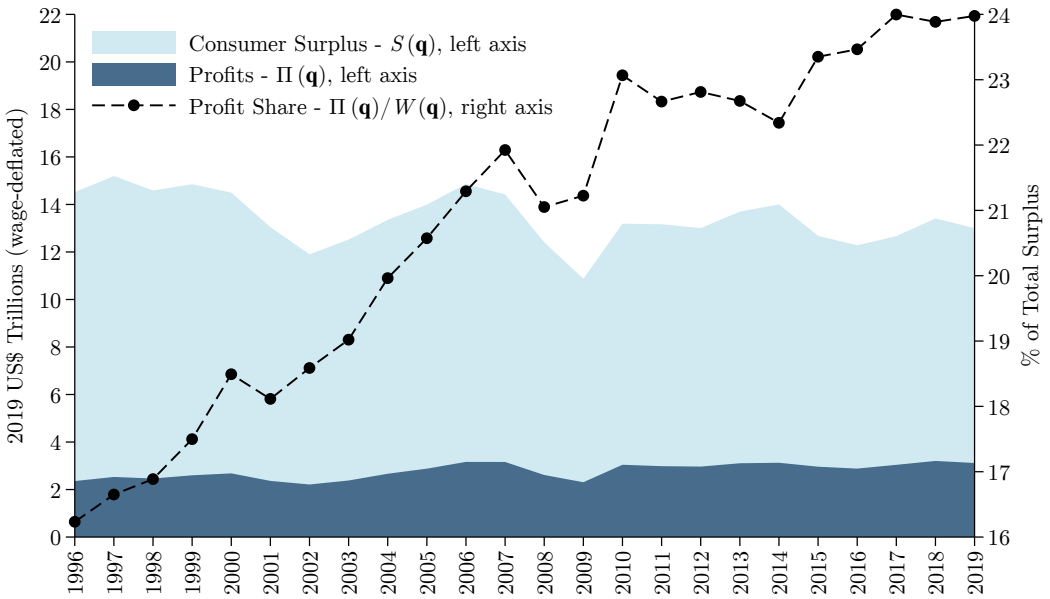
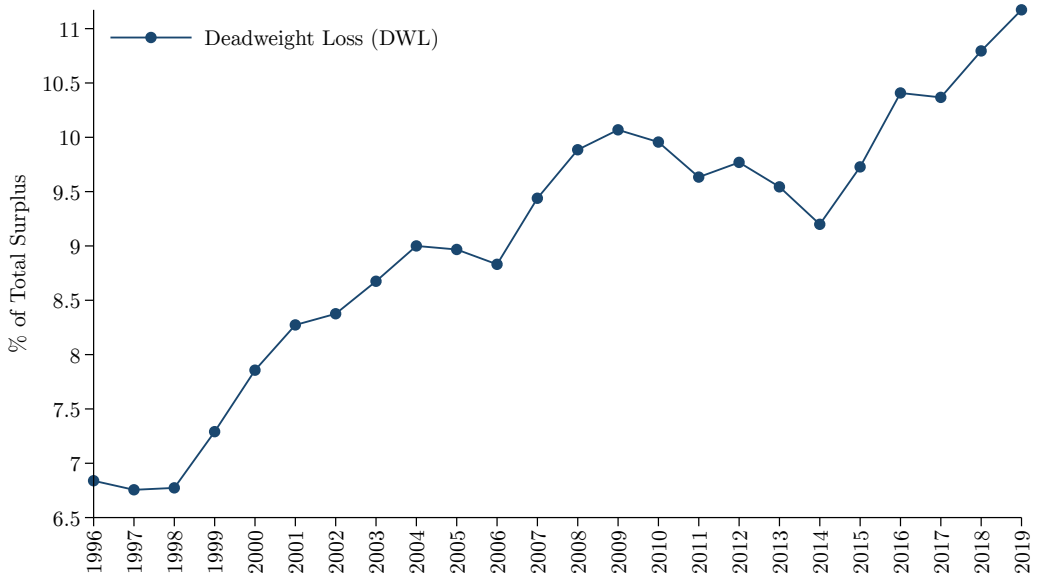


Figure 15: Deadweight Loss from Oligopoly (1996-2019)



E.2. Alternative Model with Private and Foreign Firms

Figures 16 and 17 replicate, respectively, Figures 7 and 8 for the extended model that includes a competitive fringe of atomistic firms that represent private and foreign firms (discussed in 6.2). For comparability, the surplus measures only include granular oligopolistic firms (Compustat).

Figure 16: Profit as % of Total Surplus (1996-2019)

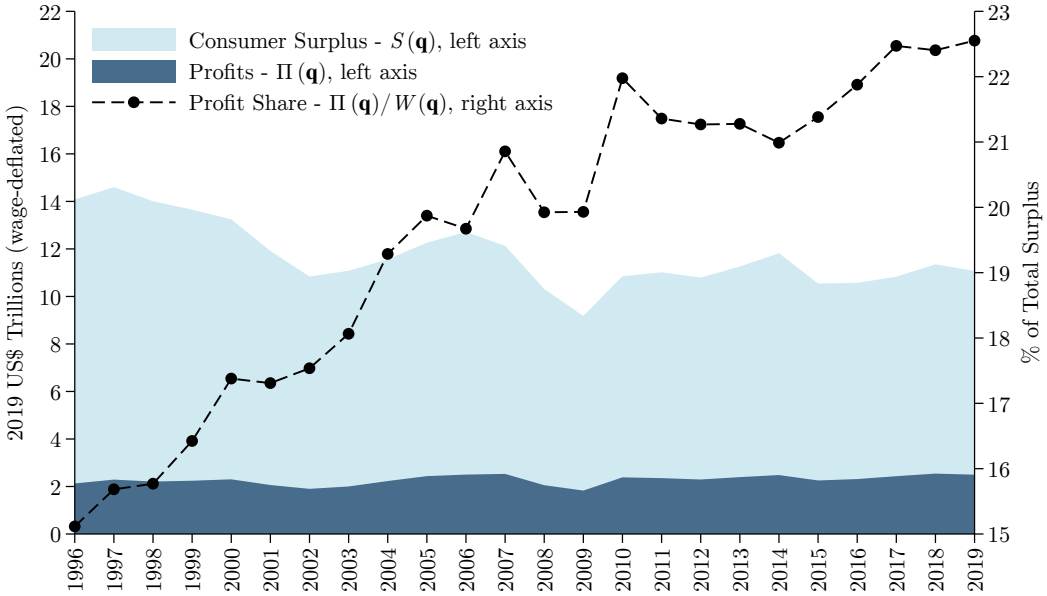
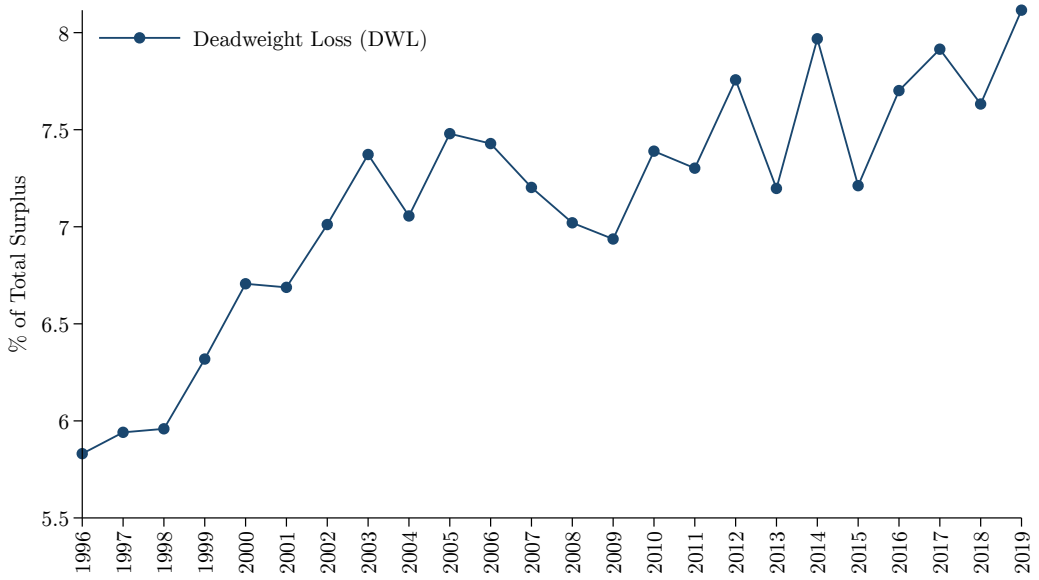


Figure 17: Deadweight Loss from Oligopoly (1996-2019)



E.3. Alternative Model: Tradable Sectors Only

Figures 18 and 19 replicate, respectively, Figures 7 and 8 for the model where non-tradable sectors are excluded (discussed in 6.3). For comparability, the surplus measures only include granular oligopolistic firms (Compustat). The list of broad sectors, with an indication of whether they are considered tradable for this model extension, is found in Table 4 (directly below).

Table 4: List of Macro-sectors (Tradable and Non-tradable)

Code	ISIC v3.1 Macro-Sector	Non-Tradable
1-5	Agriculture and Fishing	
10-14	Mining and Quarrying	
15-16	Food, Beverages and Tobacco	
17-19	Paper, Pulp and Publishing	
20	Wood and Wood Products	
21-22	Pulp, Paper and Publishing	
23	Coke, Refined Petroleum Products and Nuclear Fuel	
24	Chemicals and Chemical Products	
25	Rubber and Plastic Products	
26	Other Non-Metallic Mineral Products	
27	Basic Metals	
28	Fabricated Metal Products	
29	Machinery and Equipment	
30-33	Electrical and Optical Products	
34	Motor Vehicles and related	
35	Other Transport Equipment	
36-37	Other Manufacturing and Recycling	
40-41	Utilities	✓
45	Construction	
50-52	Wholesale and Retail Trade	✓
55	Hotels and Restaurants	✓
60-64	Transport, Logistics and Communication	
65-67	Financial Services	
70	Real Estate	
71	Renting of Machinery and Equipment	
72	Computer-Related Activities	
73-74	Other Professional Services	
80	Education	✓
85	Health Services and Social Work	✓
90-93	Other Community, Social and Personal Services	

Figure 18: Profit as % of Total Surplus (1996-2019)

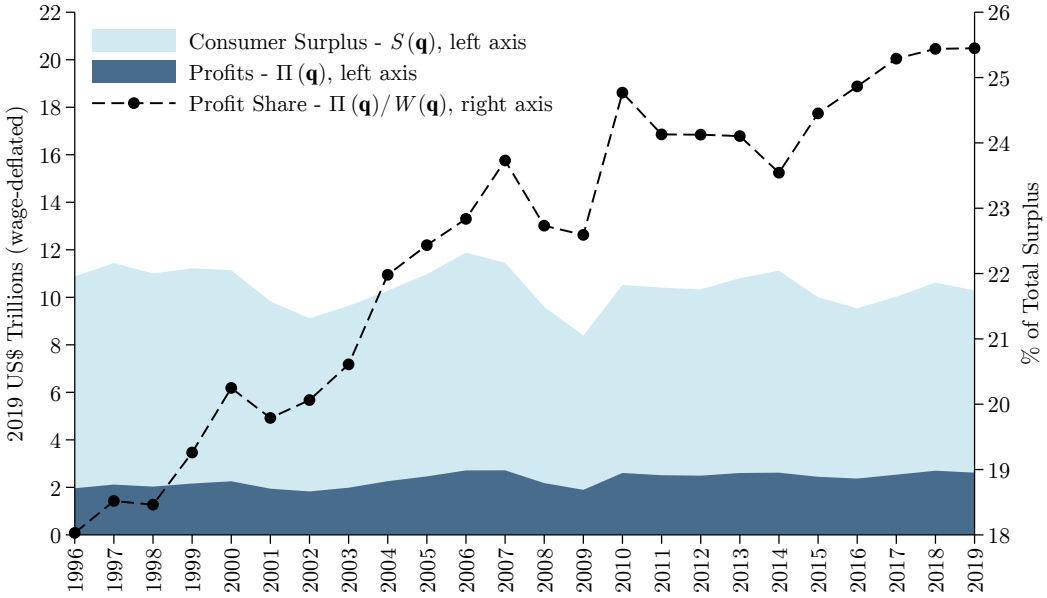
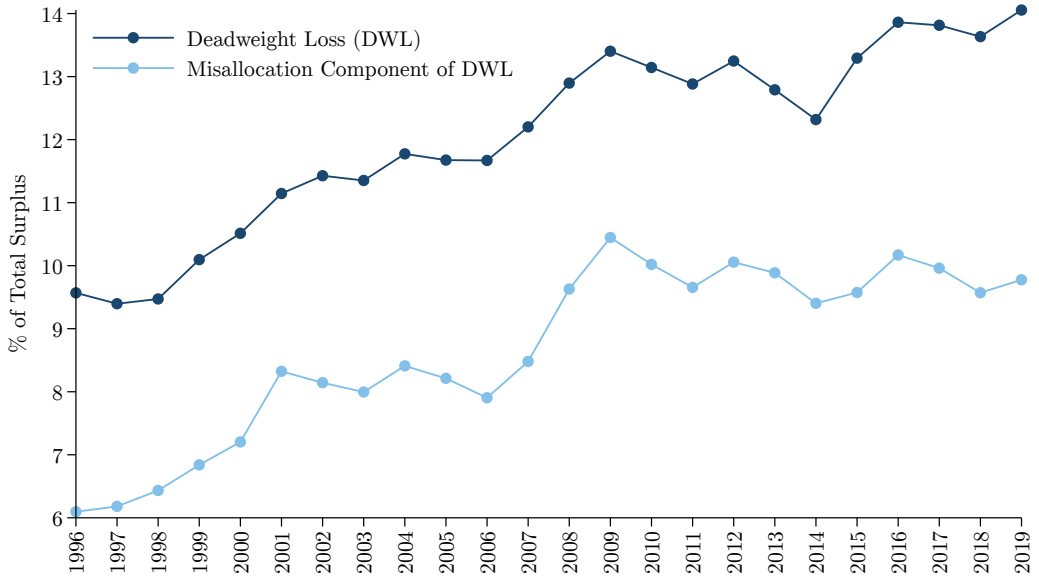


Figure 19: Deadweight Loss from Oligopoly (1996-2019)



E.4. Alternative Model with DRS, Private and Foreign Firms, Tradable Sectors Only

Figures 20 and 21 replicate, respectively, Figures 7 and 8 for the alternative model with decreasing returns to scale; a competitive fringe of atomistic firms that represent private and foreign firms and where non-tradable sectors are excluded. For comparability, the profit and surplus measures only include granular oligopolistic firms (Compustat).

FIGURE 20: PROFIT AS % OF TOTAL SURPLUS (1996-2019)

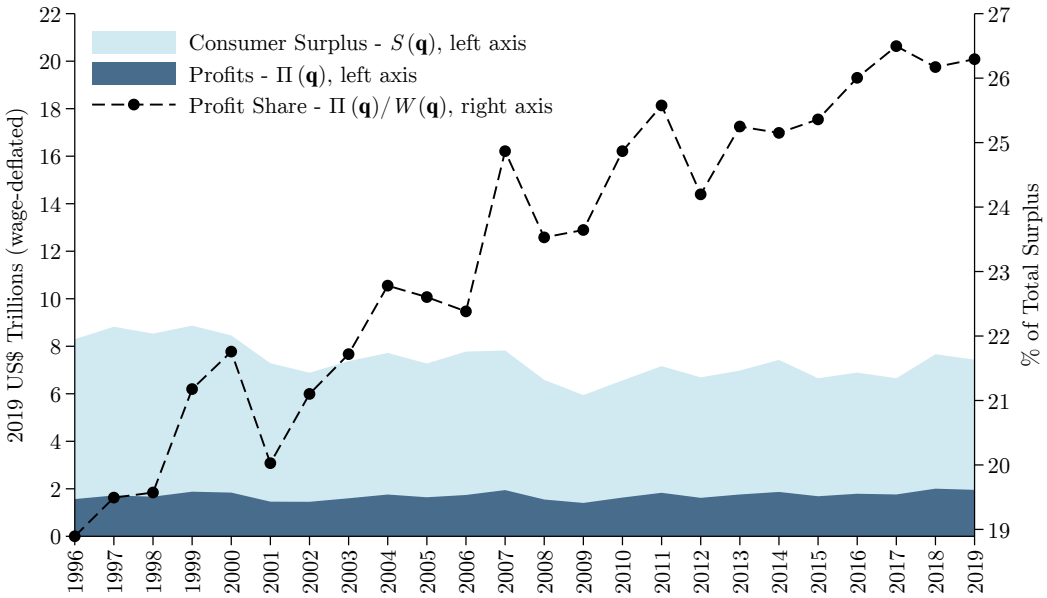
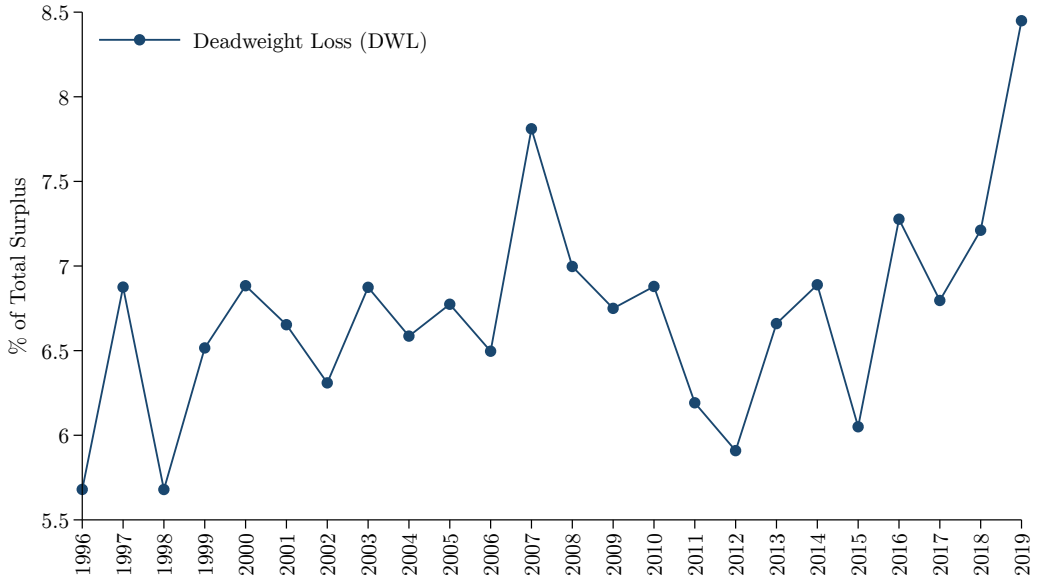


FIGURE 21: DEADWEIGHT LOSS FROM OLIGOPOLY (1996-2019)



E.5. Comparison of Cournot Oligopoly and Bertrand Oligopoly

Figures 22 and 23 compare, respectively, the profit share (of total surplus) and the deadweight loss of the baseline Cournot model against its Bertrand counterpart, presented in subsection 6.4. Figure 22 replicates (for both) the dotted black line of Figure 7, while Figure 23 does the same with the blue line of Figure 8.

Figure 22: Profit as % of Total Surplus (1996-2019)

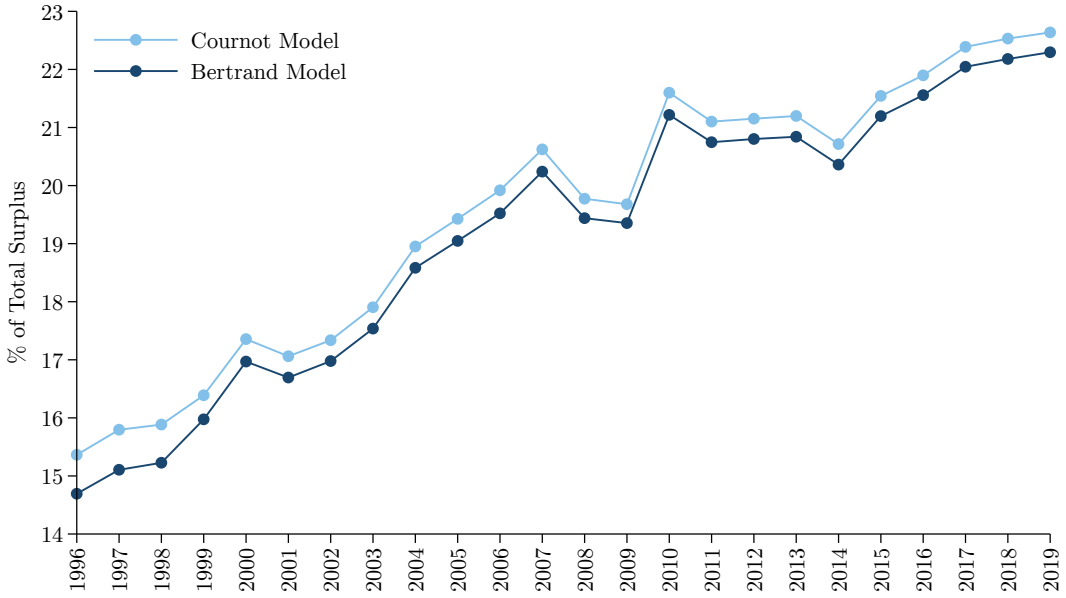
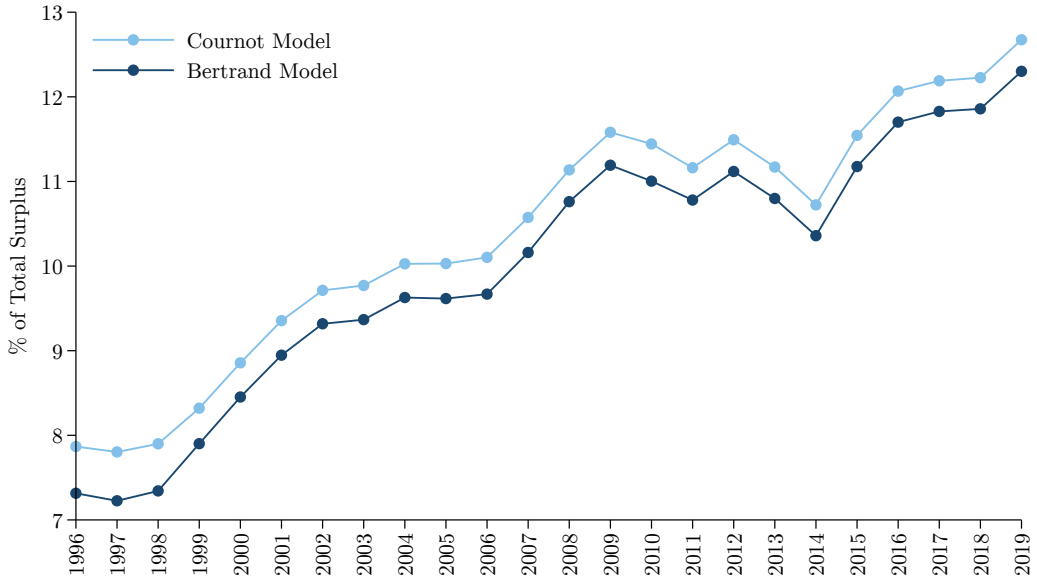


Figure 23: Deadweight Loss from Oligopoly (1996-2019)



F. Symmetric Benchmark Linear Demand Model

In this appendix, I quantify the heterogeneity in cosine/similarities and show practically how ignoring this heterogeneity can distort the measurement of other objects of the model. I do so by comparing my baseline model to a benchmark linear demand model with a symmetric cosine similarity matrix – that is, where:

$$\sigma_{ij} \;=\; \bar{\sigma} \;\forall\,(i,j) \tag{F.1}$$

$\bar{\sigma}$ is a scalar that we pick to make the benchmark model welfare-equivalent. Because total consumer surplus is $\frac{1}{2}\mathbf{q}'\mathbf{\Sigma}\mathbf{q}$, the value of $\bar{\sigma}$ that makes the two models welfare-equivalent is:

$$\bar{\sigma} \;=\; \frac{\mathbf{q}'\mathbf{\Sigma}\mathbf{q}}{\mathbf{q}'\left(\mathbf{1}\mathbf{1}'-\mathbf{I}\right)\mathbf{q}} \tag{F.2}$$

Now, for both the baseline and the benchmark symmetric model, the vector of demand intercepts $\mathbf{b}$ (which captures product quality) is identified by the following equation

$$b_i \;=\; p_i + q_i + d_i \tag{F.3}$$

where d_i is the output-weighted degree centrality of firm i. $\mathbf{p}$ and $\mathbf{q}$ are unchanged, because the identification of $\mathbf{q}$ and $\mathbf{c}$ does not depend on $\mathbf{\Sigma}$. In the benchmark symmetric model, the term $(\bar{\sigma}q_i + d_i)$ is the same across all firms; therefore, it is easy to see from the equation above that any unobserved heterogeneity in $\mathbf{\Sigma}$, summarized by the degree centrality d_i, is absorbed one-for-one by b_i.

That means that, by comparing the distribution of d_i and b_i across the two models, we can get a sense of what is missed by a model with no heterogeneity. The attached Figure compares the distribution of these two variables across the two models: the baseline model and the alternative "symmetric" model where $\sigma_{ij} = \bar{\sigma}$. The upper panel shows the cross-sectional distribution of the d_i. In the baseline model there is vast heterogeneity among firms in centrality. In the benchmark symmetric model all firms have virtually the same d_i.

The lower panel displays the same graph for the log of b_i, with a vertical solid line marking the degree centrality for the symmetric model. What we find is that, under the alternative model, we recover a completely different distribution of b_i: not only the mean, variance and skewness are off, but even the range is different: it is easy to show that b_i is bounded below by d_i, and the fact that the all firms have the same value (let's call it $\bar{d}$) for d_i implies that $\bar{d}$ provides a sharp lower boundary for the range of b_i.

The intuition is that, given the profits observed in the data, firms that are less "central" do not need to be very productive, while firms that are very productive do not need to very peripheral. The interpretation of these findings is that the ignored heterogeneity in σ is not just absorbed by b_i, but it's actually lost (less variation in $d_i \rightarrow$ less variation in b_i). Hence, ignoring variation in σ has real consequences.

where d_i is the degree centrality of firm i (defined in subsection 2.5). By ignoring the variation in cosine similarities (which are summarized at the firm level by d_i) we lose a lot of the variation in b_i as well. The intuition is that firms that are highly-central in the network of rivalries need to be producing goods that provide a lot of utility (high b_i) otherwise they wouldn't be active in equilibrium. By the same token, firms with low centrality don't need to be very productive to be active.

The implication is that ignoring the heterogeneity in cross-price elasticities is not harmless. If we lose variation in centrality it doesn't simply show up as heterogeneity in productivity: there is a real loss of information that occurs.

FIGURE 24: CROSS-SECTIONAL DISTRIBUTION OF DEGREE CENTRALITY AND DEMAND INTERCEPT (2019)

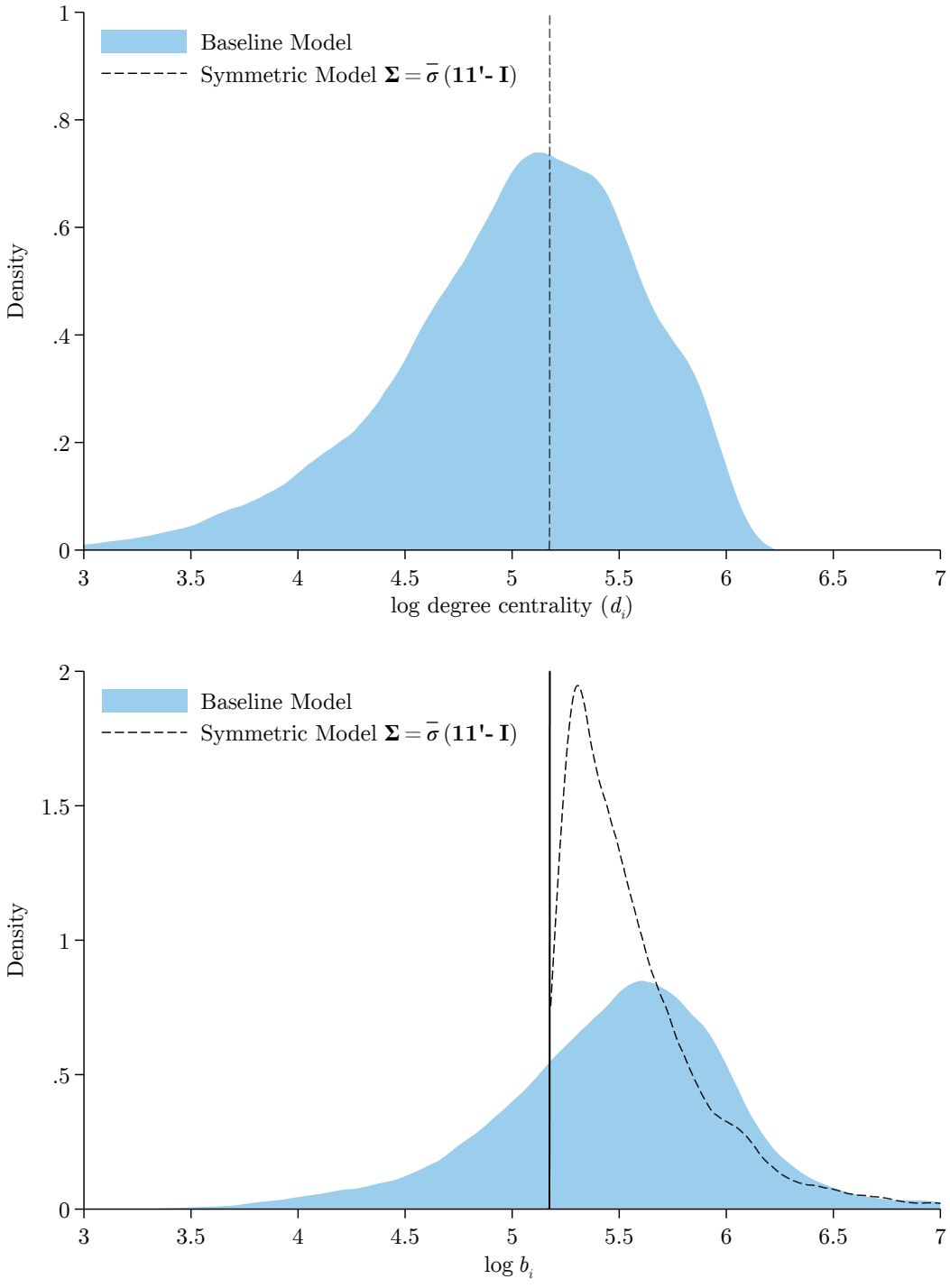


FIGURE NOTES: The figure above presents kernel density estimates of the cross-sectional distribution of the (log of) degree centrality (d_i - upper panel) and the demand intercept/quality (b_i - lower panel). Each observation is a firm, data is from the year 2019. The light blue area represents the distribution for the baseline model, while the dotted black line represents the distribution for an alternative model where $\mathbf{\Sigma} = \bar{\sigma}(\mathbf{11}' - \mathbf{I})$.

G. Proofs & Derivations

Proof to Lemma 1. Divide both sides of equation (2.34) by c_i and subtract one to obtain:

$$\frac{p_i}{c_i} \;=\; \frac{b_i + c_i - d_i}{2c_i} \tag{G.1}$$

Because $\sigma_{ij} \geq 0$ for all (i, j) and $q_j \geq 0$ for all j, $d_i \geq 0$, and the lemma follows. □

Proof to Proposition 2. Start from equation (2.17), evaluated at the Cournot equilibrium, and subtract $(\mathbf{I} + \mathbf{\Delta})\,\mathbf{q}^{\Phi}$ from both sides to obtain:

$$\mathbf{p}^{\Phi} - \mathbf{c} - (\mathbf{I} + \mathbf{\Delta})\mathbf{q}^{\Phi} \;=\; \mathbf{b} - \mathbf{c} - (2\mathbf{I} + \mathbf{\Sigma} + \Delta)\,\mathbf{q}^{\Phi} \tag{G.2}$$

next, notice that, by equation (2.26), the right hand side is zero. Move $-\mathbf{q}^{\Phi}$ to the right hand side with a positive sign:

$$\mathbf{p}^{\Phi} - \mathbf{c} - \mathbf{\Delta}\mathbf{q}^{\Phi} \;=\; \mathbf{q}^{\Phi} \tag{G.3}$$

what is left on the left hand side is $\mathbf{p}^{\Phi}$ minus the vector of marginal costs. This is the left-hand side of of the equation (2.31). To prove the right hand side, re-write equation (2.26) as:

$$(2\mathbf{I} + \mathbf{\Sigma})\,\mathbf{q} \;=\; \mathbf{b} - \mathbf{c} \tag{G.4}$$

move $\mathbf{d} = \mathbf{\Sigma}\mathbf{q}$ to the right hand side and divide both sides by two. □

Proof to Proposition 3. Rewrite equation (2.36) as

$$\mu_i - 1 = \frac{1}{2}\,(1 - \chi_i)\,(\omega_i - 1) \tag{G.5}$$

plug the expression for $\omega_i - 1 = 2\,(\bar{\mu}_i - 1)$ to obtain:

$$\mu_i - 1 = (1 - \chi_i)\,(\bar{\mu}_i - 1) \tag{G.6}$$

add 1 to both sides, take the (-1) out of the parentheses and simplify to obtain equation □

Proof to Proposition 4. Notice that the left hand side of equation (2.31) implies:

$$r_i \;=\; (p_i - c_i)\,q_i \;=\; q_i^2 \tag{G.7}$$

$$\text{then} \qquad \frac{r_i}{s_i} \;=\; \frac{2q_i^2}{q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i q_j} \;=\; \frac{2q_i}{q_i + \sum_{j \neq i} \sigma_{ij} q_i} \;\equiv\; 2\mathcal{M}_i \tag{G.8}$$

□

Proof to Proposition 7. We write the first order condition of firm z with respect to the output of product line i:

$$0 \;=\; \frac{\partial \varpi_z}{\partial q_i} \;=\; \sum_{j=1}^{n} o_{ij} o_{iz} \cdot \frac{\partial \pi_j}{\partial q_i} \;=\; b_i - c_i - 2q_i - \sum_{j=1}^{n} (1 + \kappa_{ij})\, \sigma_{ij} q_j \tag{G.9}$$

in vector form:

$$\mathbf{0} \;=\; \mathbf{b} - \mathbf{c} - (2\mathbf{I} + \mathbf{\Sigma} + \mathbf{K} \circ \mathbf{\Sigma})\, \mathbf{q} \tag{G.10}$$

the potential function $\mathbf{\Phi}$ is simply the solution to the differential equation that equates the gradient $\nabla \mathbf{\Phi}(\mathbf{q})$ to the right-hand side of equation (G.10). □

Proof to Proposition 9. Let p_{n+1} be the price of the good sold by the atomistic firms (recall we assumed they produce a homogeneous product). Because firms behave competitively and price at marginal cost, it must be the case that:

$$q(\zeta) \;=\; p_{n+1} \cdot \zeta \tag{G.11}$$

hence labor supply and profits are given by:

$$h(\zeta) \;=\; \frac{p_{n+1}^2}{2} \cdot \zeta; \qquad \pi(\zeta) \;=\; \frac{p_{n+1}^2}{2} \cdot \zeta \tag{G.12}$$

Because the atomistic firms pay an entry cost of one unit of labor, the productivity cutoff for entry, which we call $\zeta_{\min}$, is given by:

$$1 \;=\; \frac{p_{n+1}^2}{2} \cdot \zeta_{\min} \tag{G.13}$$

Let us now compute aggregate output and aggregate labor

$$\begin{array}{rclcrclcrcl} q_{n+1} & = & \int_{\zeta_{\min}}^{\infty} q(\zeta)\, \mathrm{d}F(\zeta) & = & p_{n+1} \int_{\zeta_{\min}}^{\infty} \zeta \cdot f(\zeta)\, \mathrm{d}\zeta & = & \sqrt{\frac{2}{\zeta_{\min}}} \int_{\zeta_{\min}}^{\infty} \frac{\beta - 1}{\zeta^{\beta}}\, \mathrm{d}z & = & \sqrt{2} \zeta_{\min}^{\frac{1}{2} - \beta} \\ h_{n+1} & = & \int_{\zeta_{\min}}^{\infty} h(\zeta)\, \mathrm{d}F(\zeta) & = & \frac{p_{n+1}^2}{4} \int_{\zeta_{\min}}^{\infty} \zeta \cdot f(\zeta)\, \mathrm{d}\zeta & = & \frac{1}{\zeta_{\min}} \int_{\zeta_{\min}}^{\infty} \frac{\beta - 1}{\zeta^{\beta}}\, \mathrm{d}\zeta & = & \zeta_{\min}^{-\beta} \end{array} \tag{G.14}$$

By writing the productivity cutoff $\zeta_{\min}$ in terms of aggregate output q_{n+1} and plugging it in the expression for aggregate cost h_{n+1}, we find the aggregate cost function:

$$h_{n+1} \;=\; \left[\left(\frac{q_{n+1}}{\sqrt{2}} \right)^{\frac{1}{\frac{1}{2} - \beta}} \right]^{(-\beta)} \;=\; \left(\frac{q_{n+1}}{\sqrt{2}} \right)^{\frac{2\beta}{2\beta - 1}} \tag{G.15}$$

by taking the limit $\beta \to 1^{+}$, we can see that the expression above converges to the quadratic form:

$$h_{n+1} \;=\; \frac{q_{n+1}^2}{2} \tag{G.16}$$

□

Proof to Proposition 10. The left hand side of equation (2.31) can be written as:

$$q_i^2 \;=\; p_i q_i - c_i q_i \tag{G.17}$$

Then we replace $c_i q_i = \text{TVC}_i - \delta_i q_i^2/2$, move $(\delta_i q_i/2)$ to the right hand side and obtain:

$$(1+\delta_i/2)\, q_i^2 \;=\; p_i q_i - \text{TVC}_i \tag{G.18}$$

and finally solve for q_i by dividing both sides by $(1+\delta_i/2)$ and taking the square root. □

Proof to Proposition 12. We start by re-writing equation (2.18) as:

$$\alpha \cdot \cos^{\text{HP}}_{\text{KQ}} \cdot q_{\text{K}}\, q_{\text{Q}} \;=\; -\varepsilon_{\text{KQ}}\, p_{\text{K}} q_{\text{K}} \tag{G.19}$$

Then, substituting $q_{\text{K}}, q_{\text{Q}}$ using (3.3), with $(\delta_{\text{K}} = \delta_{\text{Q}} = 0\ \forall i)$, we can write:

$$\alpha \cdot \cos^{\text{HP}}_{\text{KQ}} \cdot \sqrt{p_{\text{K}} q_{\text{K}} - \text{TVC}_{\text{K}}} \sqrt{p_{\text{Q}} q_{\text{Q}} - \text{TVC}_{\text{Q}}} \;=\; -\varepsilon_{\text{KQ}}\, p_{\text{K}} q_{\text{K}} \tag{G.20}$$

We can then re-write the same equation for (j,i) instead of (i,j). We sum left hand sides and the right hand sides to obtain:

$$2 \cdot \alpha \cdot \cos^{\text{HP}}_{\text{KQ}} \cdot \sqrt{p_{\text{K}} q_{\text{K}} - \text{TVC}_{\text{K}}} \sqrt{p_{\text{Q}} q_{\text{Q}} - \text{TVC}_{\text{Q}}} \;=\; -\left(\varepsilon_{\text{KQ}}\, p_{\text{K}} q_{\text{K}} + \varepsilon_{\text{KQ}}\, p_{\text{Q}} q_{\text{Q}}\right) \tag{G.21}$$

Finally, we rearrange with α on the left hand side to obtain equation (3.8). □

Proof to Proposition 13. The left bound simply is a restatement of the non-DRS assumption. The right bound comes from the requirement that $c_i \geq 0$. We simply re-write it in terms of revenues, costs and δ_i:

$$c_i \;=\; \frac{\text{TVC}_i}{q_i} - \frac{\delta_i}{2} q_i \;=\; \frac{\text{TVC}_i}{\sqrt{\frac{p_i q_i - \text{TVC}_i}{1+\delta_i/2}}} - \frac{\delta_i}{2}\sqrt{\frac{p_i q_i - \text{TVC}_i}{1+\delta_i/2}} \geq 0 \tag{G.22}$$

and rearrange the right hand side to obtain:

$$\delta_i \leq \frac{2 \cdot \text{TVC}_i}{p_i q_i - 2 \cdot \text{TVC}_i} \tag{G.23}$$

□

Proof to Proposition 14. Start by writing the vector of economic profits in terms of the price vector $\mathbf{p}$:

$$\pi \quad = \quad \text{diag}\left(\mathbf{p}-\mathbf{c}\right)\left(\mathbf{I}+\mathbf{\Sigma}\right)^{-1}\left(\mathbf{b}-\mathbf{p}\right) \tag{G.24}$$

Now let $\mathbb{D}$ and $\mathbb{O}$ be, respectively, the matrices containing the diagonal and off-diagonal elements of $\left(\mathbf{I}+\mathbf{\Sigma}\right)^{-1}$ so that:

$$\left(\mathbf{I}+\mathbf{\Sigma}\right)^{-1} \quad = \quad \mathbb{D}+\mathbb{O} \tag{G.25}$$

Then we can write:

$$\pi \quad = \quad \text{diag}\left(\mathbf{p}-\mathbf{c}\right)\left[\mathbb{D}\left(\mathbf{b}-\mathbf{p}\right)+\mathbb{O}\left(\mathbf{b}-\bar{\mathbf{p}}\right)\right] \tag{G.26}$$

and take the first order condition firm-by-firm by taking the price vector of other firms (denoted by the upper bar) as given:

$$0 \quad = \quad \left(\mathbf{I}+\mathbf{\Sigma}\right)^{-1}\left(\mathbf{b}-\mathbf{p}\right)-\mathbb{D}\mathbf{p}+\mathbb{D}\,\mathbf{c} \tag{G.27}$$

we can re-write this system in terms of $\mathbf{q}$:

$$0 \quad = \quad \left(\mathbf{I}+\mathbb{D}^{-1}+\mathbf{\Sigma}\right)\mathbf{q}-\left(\mathbf{b}-\mathbf{c}\right) \tag{G.28}$$

the corresponding Bertrand potential is

$$\Psi \quad = \quad \mathbf{q}'\left(\mathbf{b}-\mathbf{c}\right)-\frac{1}{2}\mathbf{q}'\left(\mathbf{I}+\mathbb{D}^{-1}+\mathbf{\Sigma}\right)\mathbf{q} \tag{G.29}$$

and the Bertrand equilibrium is:

$$\mathbf{q}^{\Psi} \quad = \quad \left(\mathbf{I}+\mathbb{D}^{-1}+\mathbf{\Sigma}\right)^{-1}\left(\mathbf{b}-\mathbf{c}\right) \tag{G.30}$$

□

H. Additional Graphs

Figure 25: Average Markup vs. Profit Rate

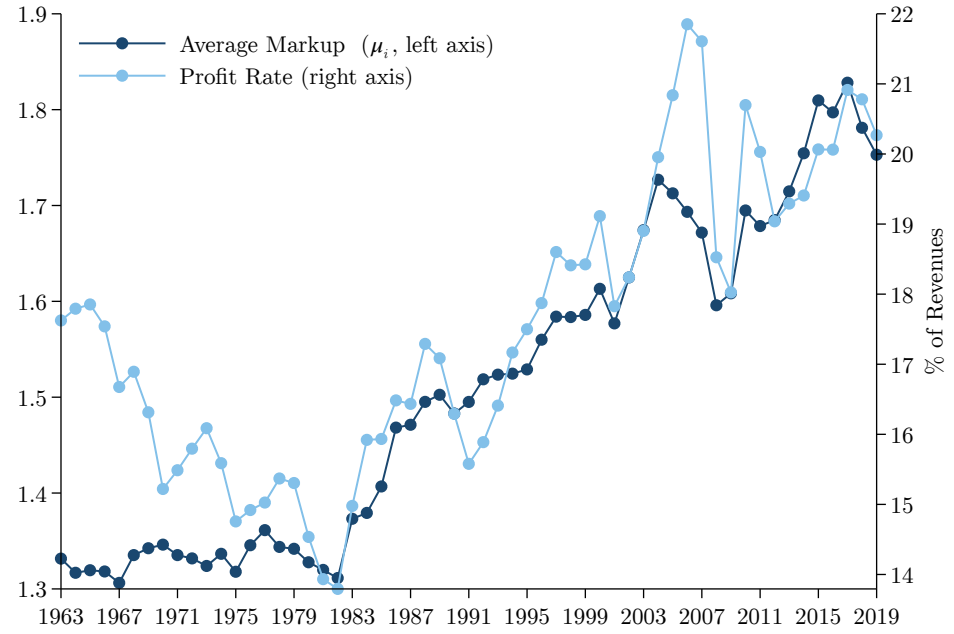


Figure 26: Operating vs. Pre-Tax Profit Rates

