

Online Appendix

PRODUCT DIFFERENTIATION AND OLIGOPOLY: A NETWORK APPROACH

Bruno Pellegrino (Columbia GSB)

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:

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

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

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

which can be expressed, in vector form, as:

$$0 = \mathbf{b} - \mathbf{c} - 2\mathbf{q} - \Sigma \bar{\mathbf{q}} \quad (\text{A.3})$$

This system of reaction functions defines a vector field $\mathbf{q}(\bar{\mathbf{q}})$ 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 $\Phi(\mathbf{q})$:

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

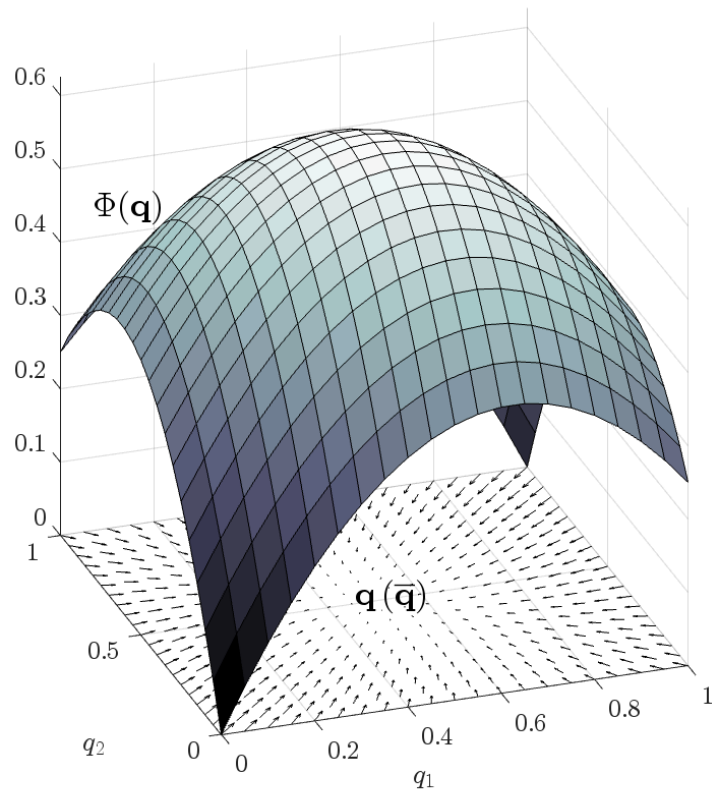
which clarifies why the maximizer of the potential function solves the Network Cournot game. The potential $\Phi(\mathbf{q})$ is then obtained as the solution to the following system of partial differential equations

$$\nabla \Phi(\mathbf{q}) = \mathbf{b} - \mathbf{c}(\mathbf{q}) - (2\mathbf{I} + \Sigma) \mathbf{q} \quad (\text{A.5})$$

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 1. 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 1: GRAPHING THE COURNOT POTENTIAL FOR THE TWO-FIRM CASE



B Generality of the Utility Function

In this appendix I investigate mathematically the generality of the utility function in equation (9). Specifically, I show that it can be derived from a more flexible functional form, by applying a series of welfare-invariant changes of volumetric units.

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

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

$$\mathbf{x} = \mathbf{A}^x \mathbf{q} \quad \text{and} \quad \mathbf{y} = \mathbf{A}^y \mathbf{q} \quad (\text{B.2})$$

$\mathbf{M}^x$ is a diagonalizable matrix, implying its “square root” matrix $(\mathbf{M}^x)^{\frac{1}{2}}$ exists and that we can write $\mathbf{M}^x$ in terms of such square root matrix as follows:

$$\mathbf{M}^x = \left[(\mathbf{M}^x)^{\frac{1}{2}} \right]' (\mathbf{M}^x)^{\frac{1}{2}} \quad (\text{B.3})$$

$\mathbf{M}^y$ is a positive diagonal matrix, we do *not* require $\|\mathbf{a}_k^x\| = 1$ and the matrix $\mathbf{A}^y$ is defined as follows:

$$a_{ii}^y \stackrel{\text{def}}{=} \sqrt{\frac{\tau_2}{\tau_4} \cdot \frac{\alpha}{1-\alpha} \cdot m_{ii}^y} \cdot \left\| \left[(\mathbf{M}^x)^{-\frac{1}{2}} \mathbf{A}^x \right]_i \right\| \quad (\text{B.4})$$

the intuitive meaning of this restriction (which will be easy to see later on) is that all goods have the same relative loading on the common v/s the idiosyncratic characteristics. It ensures that, after normalization, we only need a single common loading parameter for α all goods (as opposed a different one for each good).

Our objective is to show that, up to some changes in volumetric units, we can rewrite the above utility in the following form

$$U(\mathbf{x}, \mathbf{y}, H) \stackrel{\text{def}}{=} \alpha \left(\mathbf{x}' \mathbf{b}^x - \frac{1}{2} \mathbf{x}' \mathbf{x} \right) - (1-\alpha) \left(\mathbf{y}' \mathbf{b}^y - \frac{1}{2} \mathbf{y}' \mathbf{y} \right) - H \quad (\text{B.5})$$

$$\mathbf{x} = \mathbf{A} \mathbf{q}, \quad \|\mathbf{a}_k\| = 1 \quad \forall k \quad \text{and} \quad \mathbf{y} = \mathbf{q} \quad (\text{B.6})$$

First, notice that the following utility

$$U(\mathbf{x}, \mathbf{y}, H) = \frac{\tau_1}{\tau_5} \mathbf{x}' \mathbf{b}^x - \frac{\tau_2}{2\tau_5} \cdot \mathbf{x}' \mathbf{M}^x \mathbf{x} + \frac{\tau_3}{\tau_5} \mathbf{y}' \mathbf{b}^y - \frac{\tau_4}{2\tau_5} \mathbf{y}' \mathbf{M}^y \mathbf{y} - H \quad (\text{B.7})$$

describes the exact same preferences as equation (B.1), therefore we can use it instead at no loss. Let us now define the following alternative vectors of characteristics:

$$\tilde{\mathbf{x}} \stackrel{\text{def}}{=} \sqrt{\frac{\tau_2}{\alpha\tau_5}} (\mathbf{M}^x)^{\frac{1}{2}} \mathbf{x} \quad \text{and} \quad \tilde{\mathbf{y}} \stackrel{\text{def}}{=} \sqrt{\frac{\tau_4}{(1-\alpha)\tau_5}} (\mathbf{M}^y)^{\frac{1}{2}} \mathbf{y} \quad (\text{B.8})$$

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

$$U(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, H) = \frac{\tau_1 \sqrt{\alpha}}{\sqrt{\tau_2 \tau_5}} \cdot \tilde{\mathbf{x}}' (\mathbf{M}^x)^{-\frac{1}{2}} \mathbf{b}^x - \frac{\alpha}{2} \cdot \tilde{\mathbf{x}}' \tilde{\mathbf{x}} + \frac{\tau_3 \sqrt{1-\alpha}}{\sqrt{\tau_2 \tau_5}} \cdot \tilde{\mathbf{y}}' (\mathbf{M}^y)^{-\frac{1}{2}} \mathbf{b}^y - \frac{1-\alpha}{2} \tilde{\mathbf{y}}' \tilde{\mathbf{y}} - H \quad (\text{B.9})$$

Next, define:

$$\tilde{\mathbf{b}}^x \stackrel{\text{def}}{=} \frac{\tau_1}{\sqrt{\alpha\tau_2\tau_5}} (\mathbf{M}^x)^{-\frac{1}{2}} \mathbf{b}^x \quad \text{and} \quad \tilde{\mathbf{b}}^y \stackrel{\text{def}}{=} \frac{\tau_3}{\sqrt{(1-\alpha)\tau_2\tau_5}} (\mathbf{M}^y)^{-\frac{1}{2}} \mathbf{b}^y \quad (\text{B.10})$$

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

$$U(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, H) = \alpha \left(\tilde{\mathbf{x}}' \tilde{\mathbf{b}}^x - \frac{1}{2} \cdot \tilde{\mathbf{x}}' \tilde{\mathbf{x}} \right) + (1 - \alpha) \left(\tilde{\mathbf{y}}' \tilde{\mathbf{b}}^y - \frac{1}{2} \tilde{\mathbf{y}}' \tilde{\mathbf{y}} \right) - H \quad (\text{B.11})$$

$$\text{with} \quad \hat{\mathbf{x}} = \tilde{\mathbf{A}}^x \mathbf{q} \quad \text{and} \quad \tilde{\mathbf{y}} = \tilde{\mathbf{A}}^y \mathbf{q} \quad (\text{B.12})$$

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

$$\tilde{\mathbf{A}}^x \stackrel{\text{def}}{=} \sqrt{\frac{\alpha \tau_5}{\tau_2}} (\mathbf{M}^x)^{-\frac{1}{2}} \mathbf{A}^x \quad \text{and} \quad \tilde{\mathbf{A}}^y \stackrel{\text{def}}{=} \sqrt{\frac{(1 - \alpha) \tau_5}{\tau_4}} (\mathbf{M}^y)^{-\frac{1}{2}} \mathbf{A}^y \quad (\text{B.13})$$

Note that $\tilde{\mathbf{A}}^y$ is also a diagonal matrix, and the definition of $\mathbf{A}^y$ (equation (B.4)) implies that:

$$\|\tilde{\mathbf{a}}_i^x\| \equiv \tilde{a}_{ii}^y \quad (\text{B.14})$$

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.

Note that have not yet imposed that $\tilde{\mathbf{a}}_i^x$ has unit length, and thus $\tilde{\mathbf{A}}^y$ is not necessarily an identity matrix. Therefore the final step is to show that the restriction

$$\|\tilde{\mathbf{a}}_i^x\| = a_{ii}^y = 1 \quad \forall i \quad (\text{B.15})$$

is without loss of generality. To achieve this, we change the volumetric units of our goods by defining:

$$\hat{\mathbf{q}} \stackrel{\text{def}}{=} \tilde{\mathbf{A}}^y \mathbf{q} \quad (\text{B.16})$$

so that

$$\tilde{\mathbf{x}} = \hat{\mathbf{A}} \hat{\mathbf{q}} \quad \text{with} \quad \hat{a}_{ij} \stackrel{\text{def}}{=} \frac{\tilde{a}_{ij}}{a_{ii}^y} \quad (\text{B.17})$$

this implies finally that

$$\|\hat{\mathbf{a}}_i^x\| = 1 \quad \forall i \quad \text{and} \quad \tilde{\mathbf{y}} = \hat{\mathbf{q}} \quad (\text{B.18})$$

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

$$U(\hat{\mathbf{q}}, H) = \alpha \left(\hat{\mathbf{q}}' \hat{\mathbf{A}} \tilde{\mathbf{b}}^x - \frac{1}{2} \cdot \hat{\mathbf{q}}' \hat{\mathbf{A}}' \hat{\mathbf{A}} \hat{\mathbf{q}} \right) + (1 - \alpha) \left(\hat{\mathbf{q}}' \tilde{\mathbf{b}}^y - \frac{1}{2} \hat{\mathbf{q}}' \hat{\mathbf{q}} \right) - H \quad (\text{B.19})$$

where $\hat{\mathbf{A}}' \hat{\mathbf{A}}$ is now by construction 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. This degree of freedom allows us to pick labor as the numéraire (one labor unit = 1\$).

C The Role of Idiosyncratic Characteristics

In this Appendix, I clarify the role of idiosyncratic characteristics in GHL demand. From a *theoretical* point of view, the presence of idiosyncratic characteristics is somewhat redundant, because they can just be seen as a special case of the common characteristics (i.e. we could simply include the entries of vector $\mathbf{x}$ inside $\mathbf{y}$ at no loss and add n rows to the $\mathbf{A}$ matrix).

Now, I argue, while redundant from a *theoretical perspective*, from an *empirical perspective* the idiosyncratic characteristics make a significant difference, by providing a valuable additional degree of freedom (the param-

eter α). Key to this is the assumption, made explicit in sub-section II.D, that the words contained Hoberg & Phillips's vocabulary are always mapped to the /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 :

$$\partial p_i / \partial q_j = \mathbf{a}'_i \mathbf{a}_j \quad (\text{C.1})$$

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*.

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

$$\partial p_i / \partial q_j = \alpha \mathbf{a}'_i \mathbf{a}_j \quad (\text{C.2})$$

By calibrating the parameter α , we can ensure that the magnitude of the recovered elasticities is consistent with state-of-the-art estimates from the IO literature (see Section (III)).

D Nash-Cournot Equilibrium and Network Centrality

In this appendix, I provide additional details on the relationship between equation (29), 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 I is a type of linear-quadratic network game. Ballester et al. [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 1 (Katz-Bonacich Centrality). For a weighted network with adjacency matrix $\mathbf{M}$, we define the vector of centralities $\mathbf{k}$, a function of $\mathbf{M}$ with parameters $(\eta, \mathbf{g})$:

$$\begin{aligned} \mathbf{k}(\mathbf{M}; \eta, \mathbf{g}) : \quad \mathbf{k} &= \eta \mathbf{M} \mathbf{k} + \mathbf{g} \\ &= (\mathbf{I} - \eta \mathbf{M})^{-1} \mathbf{g} \end{aligned} \quad (\text{D.1})$$

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 I (equation 29) can be easily verified to coincide with the vector of Katz-Bonacich centralities, with parametrization $(\frac{1}{2}, \frac{1}{2}(\mathbf{b} - \mathbf{c}))$:

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

Notice that the Cournot game is played over a *negatively-weighted* network (the adjacency matrix is $-\Sigma$). 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 Σ .

E Consumer Surplus Disaggregation

In this appendix, I provide the theoretical background for my definition of the firm-level consumer surplus $s_i(\mathbf{q})$. I also provide a mathematical derivation. As remarked in Section I, Hotelling [1938] showed that, for the n -goods case, Consumer Surplus is given by the following line integral (of a vector field):

$$S(\mathbf{q}) = \int_{\mathcal{Q}} \sum_{i=1}^n \text{MCS}_i(\bar{\mathbf{q}}) d\bar{q}_i = \int_0^1 \sum_{i=1}^n \left(b_i - \bar{q}_i - p_i - \sum_{j \neq i} \sigma_{ij} \bar{q}_j \right) \bar{q}'_i(t) dt \quad (\text{E.1})$$

where $\mathcal{Q}$ is a piece-wise smooth curve from $\mathbf{0}$ to $\mathbf{q}$ and $\bar{\mathbf{q}} : [0, 1] \rightarrow \mathcal{Q}$ is a bijective parametrization of $\mathcal{Q}$ such that $\bar{\mathbf{q}}(0) = \mathbf{0}$ and $\bar{\mathbf{q}}(1) = \mathbf{q}$. The vector valued function

$$\left[\text{MCS}_1(\bar{\mathbf{q}}) \quad \text{MCS}_2(\bar{\mathbf{q}}) \quad \cdots \quad \text{MCS}_n(\bar{\mathbf{q}}) \right]' \quad (\text{E.2})$$

is n -dimensional vector field.

Now, for a generic n -dimensional vector field $\left[F_1(\bar{\mathbf{q}}) \quad F_2(\bar{\mathbf{q}}) \quad \cdots \quad F_n(\bar{\mathbf{q}}) \right]'$ and a generic curve $\mathcal{C}$ from $\mathbf{q}_0$ to $\mathbf{q}_1$ and parametrization $\bar{\mathbf{q}} : [0, 1] \rightarrow \mathcal{C}$, it is well-known from the theory of vector calculus [see Marsden and Tromba, 2012, pages 358-373] that if we exchange the order of summation-integration we can write the line integral as:

$$\int_{\mathcal{C}} \sum_{i=1}^n F_i(\bar{\mathbf{q}}) d\bar{q}_i = \int_0^1 F_1(\bar{\mathbf{q}}) \bar{q}'_1(t) dt + \int_0^1 F_2(\bar{\mathbf{q}}) \bar{q}'_2(t) dt + \dots + \int_0^1 F_n(\bar{\mathbf{q}}) \bar{q}'_n(t) dt \quad (\text{E.3})$$

The integral $\int_0^1 F_i(\bar{\mathbf{q}}) \bar{q}'_i(t) dt$ is called the *contribution* of component i , and its interpretation depends on the physical setting. In classical mechanics (the canonical application), i would be a direction in an n -dimensional space, $\bar{\mathbf{q}}$ represents the coordinates of a particle in space, the vector-valued function

$$\left[F_1(\bar{\mathbf{q}}) \quad F_2(\bar{\mathbf{q}}) \quad \cdots \quad F_n(\bar{\mathbf{q}}) \right]' \quad (\text{E.4})$$

represents a force field (gravity, magnetic, etc...), and the line integral $\int_{\mathcal{C}} \sum_{i=1}^n F_i(\bar{\mathbf{q}}) d\bar{q}_i$ represents the total work done on the particle by the force field along the path from point $\mathbf{q}_0$ to point $\mathbf{q}_1$. If the force field is conservative (energy is conserved), the integral is path-independent. That is, the total work done will only depend on the endpoints of the curve $\mathcal{C}$ ($\mathbf{q}_0$ and $\mathbf{q}_1$), and not on the path taken by the particle to go from one endpoint to another.

In classical mechanics, the contribution of component i $\left(\int_0^1 F_i(\bar{\mathbf{q}}) \bar{q}'_i(t) dt \right)$ is interpreted as the total work done on the particle only along direction i .

In our consumer demand setting, each coordinate i represents a good, instead of a direction in space. The vector-valued function

$$\left[\text{MCS}_1(\bar{\mathbf{q}}) \quad \text{MCS}_2(\bar{\mathbf{q}}) \quad \cdots \quad \text{MCS}_n(\bar{\mathbf{q}}) \right]' \quad (\text{E.5})$$

represents the marginal consumer surplus induced by an additional unit of every good i when bundle $\bar{\mathbf{q}}$ is already being consumed (as opposed to a force field). The line integral $\int_{\mathcal{Q}} \sum_{i=1}^n \text{MCS}_i(\bar{\mathbf{q}}) d\bar{q}_i$ is consumer surplus. Because we derived MCS_i from a well-defined utility function, the MCS vector field is conservative by construction. As a consequence, the consumer surplus line integral is path-independent (the change in

Figure 2: Domain of Integration for Consumer Surplus

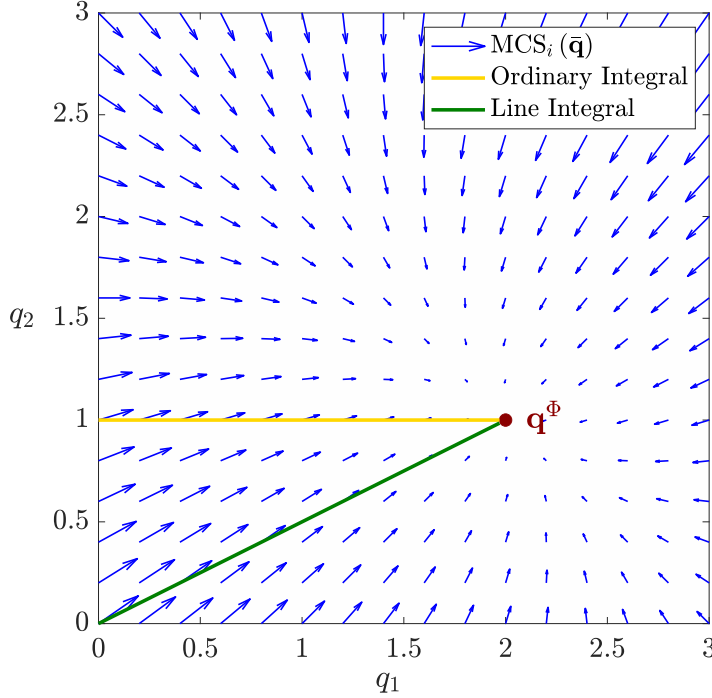


FIGURE NOTES: this graph represents marginal consumer surplus (MCS) as a vector field for the two-goods case, and illustrates the difference in computing consumer surplus for good 1 as an ordinary integral (fixing the consumption of good 2 at the equilibrium quantity) or as a line integral. $\sigma_{ij} = .3$, $(b_1 - p_1) = 2.3$ and $(b_2 - p_2) = 1.6$.

consumer surplus from one bundle to another is always the same, it does not depend along what path the consumer adjusts from one bundle to another).

Applying the identity in E.3 to consumer surplus we can thus write:

$$S(\mathbf{q}) = \int_0^1 \text{MCS}_1(\bar{\mathbf{q}}) \bar{q}'_1(t) dt + \int_0^1 \text{MCS}_2(\bar{\mathbf{q}}) \bar{q}'_2(t) dt + \dots + \int_0^1 \text{MCS}_n(\bar{\mathbf{q}}) \bar{q}'_n(t) dt \quad (\text{E.6})$$

The contribution of good i to total consumer surplus is therefore:

$$s_i(\mathbf{q}) \stackrel{\text{def}}{=} \int_0^1 \text{MCS}_i(\bar{\mathbf{q}}) \bar{q}'_i(t) dt \quad (\text{E.7})$$

Notice that the consumer surplus for the one good case (equation 44) is just a special case of the equation above. Intuitively, the use of the path integral means that we are computing the change in consumer surplus as we move from the null bundle $(0, 0, \dots, 0)$ to bundle $[q_1, q_2, \dots, q_{i-1}, q_i, q_{i+1}, \dots, q_n]'$. With the ordinary integral, we would be moving from bundle $[q_1, q_2, \dots, q_{i-1}, 0, q_{i+1}, \dots, q_n]'$ to bundle $[q_1, q_2, \dots, q_{i-1}, q_i, q_{i+1}, \dots, q_n]'$. This difference is illustrated in Figure 2.

We can now write total consumer surplus as the sum of the contributions from each individual good:

$$S(\mathbf{q}) = \sum_{i=1}^n s_i(\mathbf{q}) \quad (\text{E.8})$$

Now, one challenge we still face (in defining s_i) is that while total consumer surplus S is path-independent, the contribution of good i (s_i) is path-dependent. This poses the problem of along what curve $\mathcal{Q}$ we should integrate. Considering that $\mathcal{Q}$ represents (implicitly) a relaxation of the consumer's budget constraint (an increase in income or a drop in prices), there are two natural options. The first is the straight line from $\mathbf{0}$ to $\mathbf{q}$: this curve represents a linear drop in prices from $\bar{\mathbf{p}} = \mathbf{b}$ (the unique set of prices that induce the consumer to demand the null bundle) to $\bar{\mathbf{p}} = \mathbf{p}$. Then the curve $\mathcal{Q}$ is defined by the parametrization $\bar{\mathbf{q}} = \mathbf{q}t$.

The natural alternative is to integrate along the Engel curve: this would simulate instead an increase in income from zero. Because preferences are not homothetic, this would result in a different measure of firm-level consumer surplus. In this paper, I utilize the first definition for computability reasons: before reaching the interior optimal bundle, the Engel curve would hit many corner bundles. While this second measure of consumer surplus is mathematically well-defined, it is computationally infeasible. Another reason to prefer the straight line approach, in this setting, is that from the point of view of the consumer, income is endogenous (the consumer chooses how much to work), while prices are exogenous (the consumer is a price-taker). It is therefore natural to consider a change in prices.

Using the straight line parametrization $\bar{\mathbf{q}} = \mathbf{q}t$, we have:

$$\begin{aligned} s_i(\mathbf{q}) &\stackrel{\text{def}}{=} \int_0^1 \left[b_i - p_i - \left(q_i t + \sum_{j \neq i} \sigma_{ij} q_j t \right) \right] q_i dt = \\ &= \left[b_i - p_i - \left(q_i + \sum_{j \neq i} \sigma_{ij} q_j \right) \int_0^1 t dt \right] q_i = \\ &= \left[b_i - p_i - \frac{1}{2} \left(q_i + \sum_{j \neq i} \sigma_{ij} q_j \right) \right] q_i = \frac{1}{2} \left(q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i q_j \right) \end{aligned} \quad (\text{E.9})$$

If we were to use instead the ordinary integral, we would instead have to fix the level of output of every other firm (say, at $\tilde{\mathbf{q}}_{-i}$). Unless we arbitrarily pick this level (say, $\tilde{\mathbf{q}}_{-i} = \frac{1}{2}\mathbf{q}_{-i}$), this would yield a measure of firm-level consumer surplus that does not satisfy aggregation:

$$\int_0^{q_i} \text{MCS}_i(\bar{q}_i, \tilde{\mathbf{q}}_{-i}) d\bar{q}_i = q_i \left[b_i - p_i - \left(\frac{1}{2} q_i + \sum_{j \neq i} \sigma_{ij} \tilde{q}_j \right) \right] = \frac{1}{2} q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i (q_j - \tilde{q}_j) \quad (\text{E.10})$$

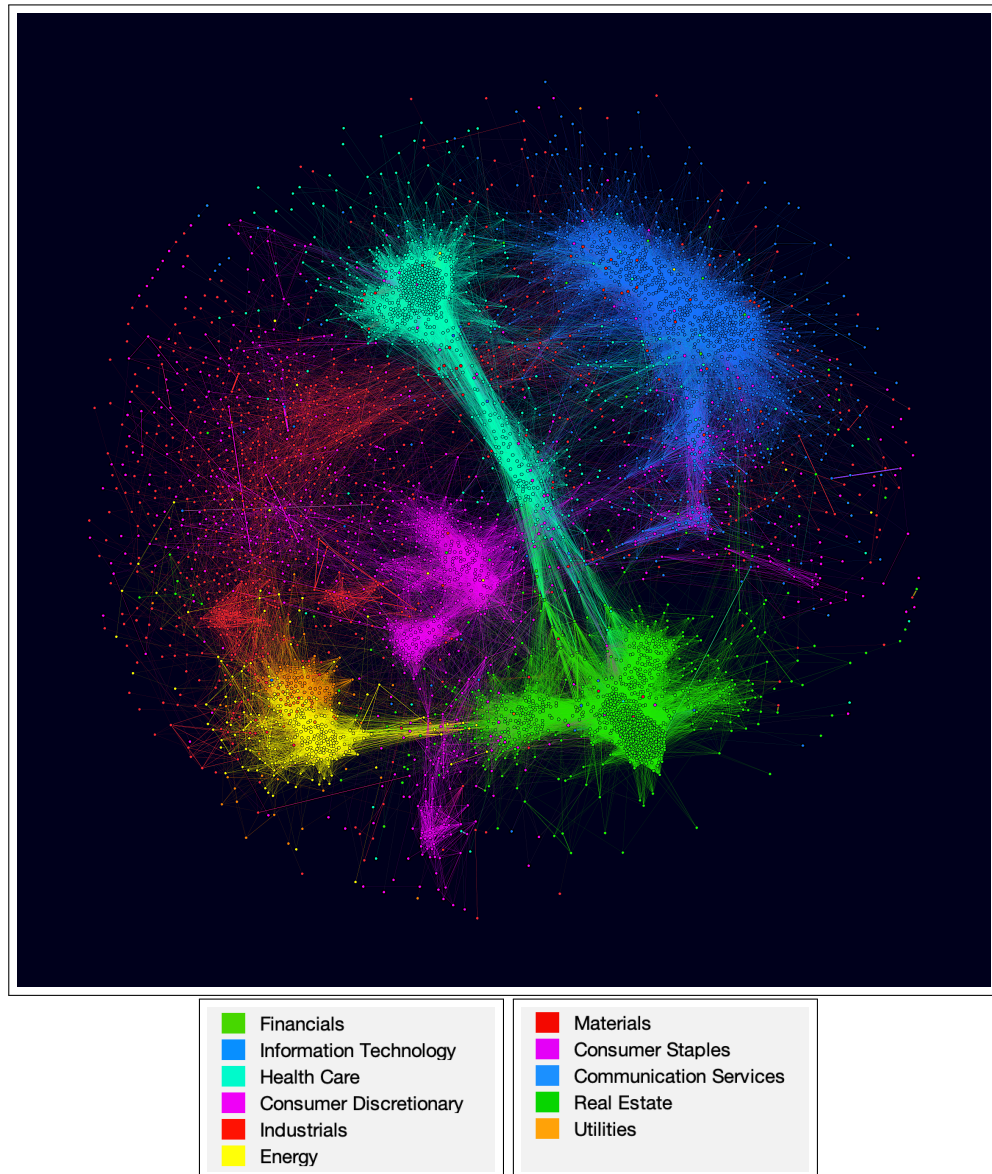
$$\sum_{i=1}^n \int_0^{q_i} \text{MCS}_i(\bar{q}_i, \tilde{\mathbf{q}}_{-i}) d\bar{q}_i = S(\mathbf{q}) + \sum_{i=1}^n \sum_{j \neq i} \sigma_{ij} q_i \left(\frac{1}{2} q_j - \tilde{q}_j \right) \quad (\text{E.11})$$

Again, this is a consequence of the fact that we are not considering a move from the null bundle to $\mathbf{q}$, but we are considering only adjusting the quantity consumed of good i . It is not a measure of the total surplus contributed by good i . Instead, it is a measure of the marginal consumer surplus contributed by good i when we fix the consumption of every other good at $\tilde{\mathbf{q}}_{-i}$.

F 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 2, 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 3: VISUALIZATION OF THE PRODUCT SPACE (ALTERNATE COLORING)



G Model Extensions: Empirical Results

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

Figures 4 and 5 replicate, respectively, Figures 5 and 6 for the extended model where firms have a non-flat marginal cost function (discussed in V.B).

FIGURE 4: PROFIT AS % OF TOTAL SURPLUS (1995-2021)

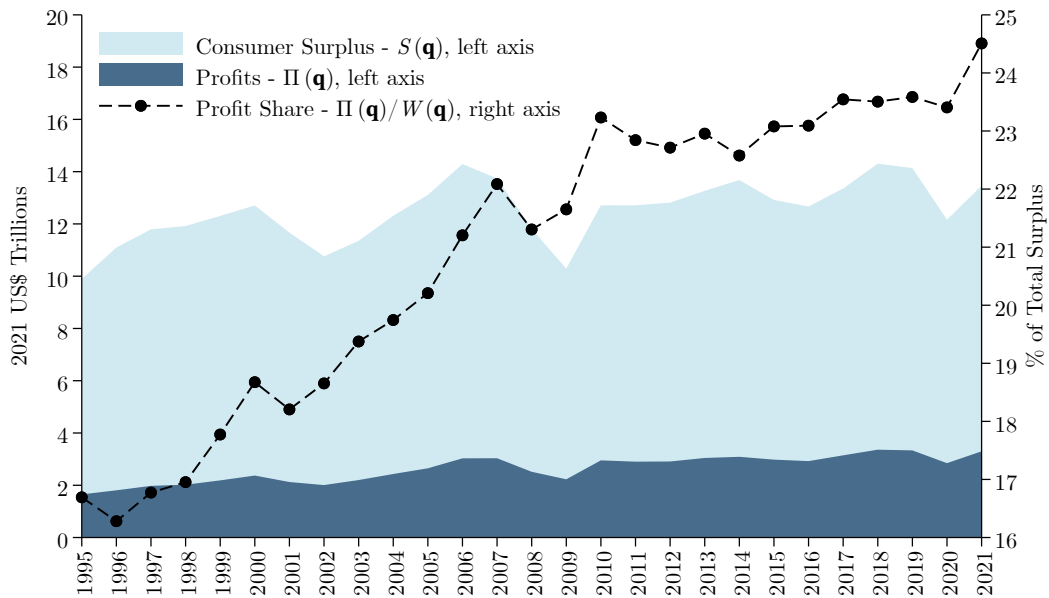
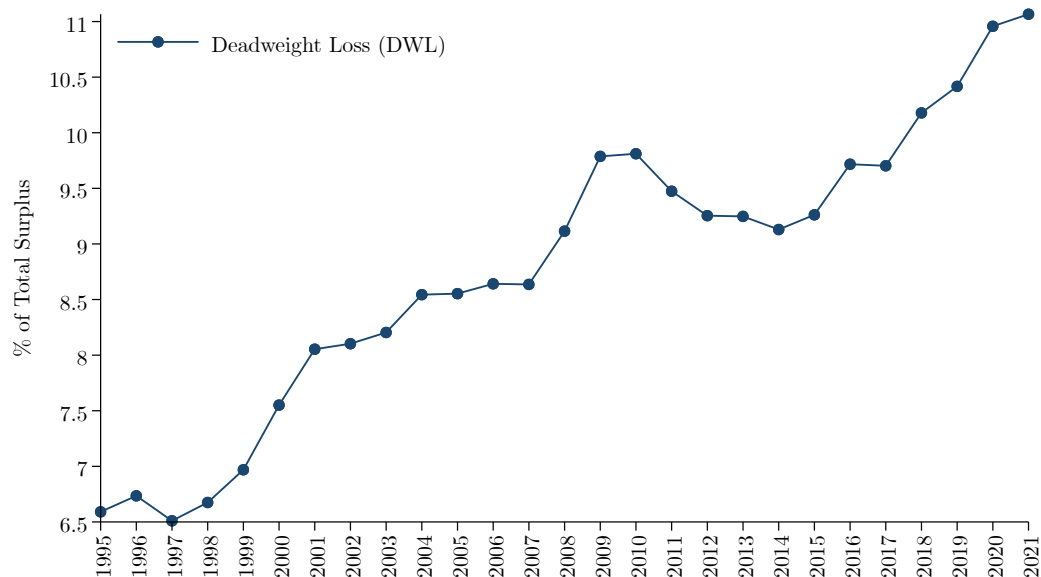


FIGURE 5: DEADWEIGHT LOSS FROM OLIGOPOLY (1995-2021)



G.2 Alternative Model with Private and Foreign Firms

Figures 6 and 7 replicate, respectively, Figures 5 and 6 for the extended model that includes a competitive fringe of atomistic firms that represent private and foreign firms (discussed in V.B). For comparability, the surplus measures only include granular oligopolistic firms (Compustat).

FIGURE 6: PROFIT AS % OF TOTAL SURPLUS (1995-2021)

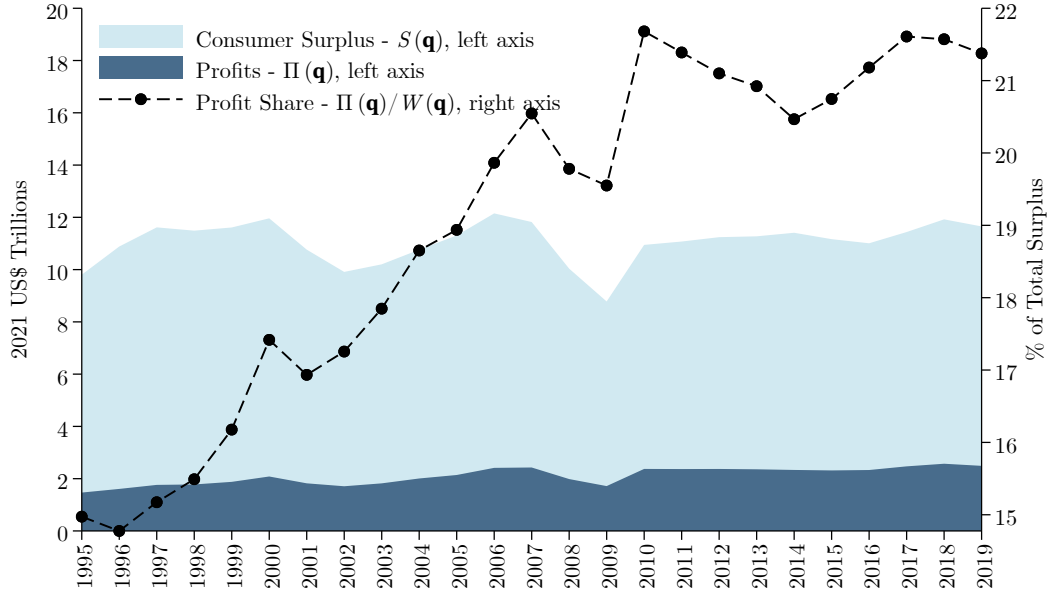
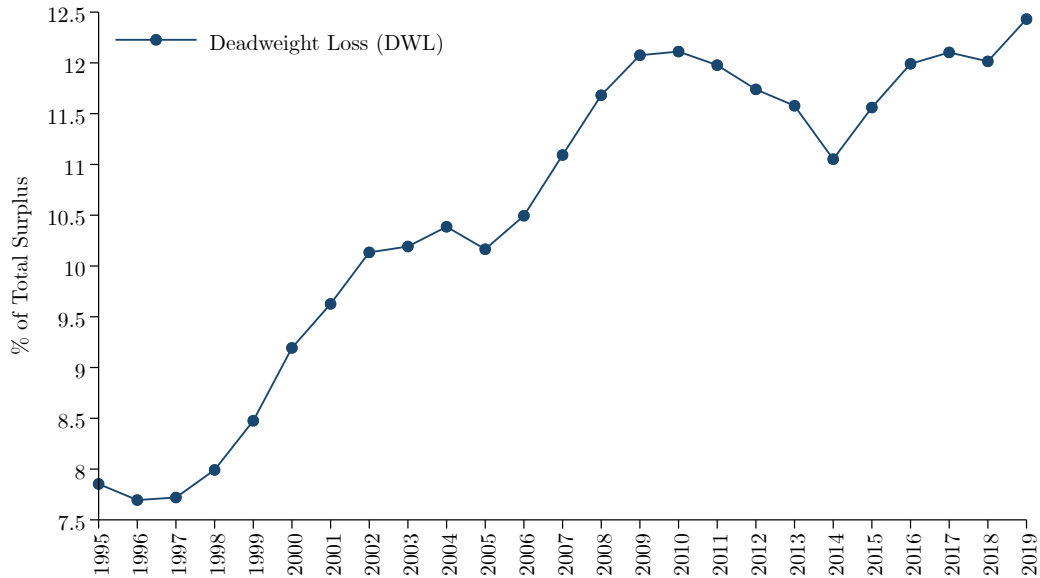


FIGURE 7: DEADWEIGHT LOSS FROM OLIGOPOLY (1995-2021)



G.3 Alternative Model: Tradable Sectors Only

Figures 8 and 9 replicate, respectively, Figures 5 and 6 for the model where non-tradable sectors are excluded (discussed in V.C). 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 1 (directly below).

TABLE 1: 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 8: PROFIT AS % OF TOTAL SURPLUS (1995-2021)

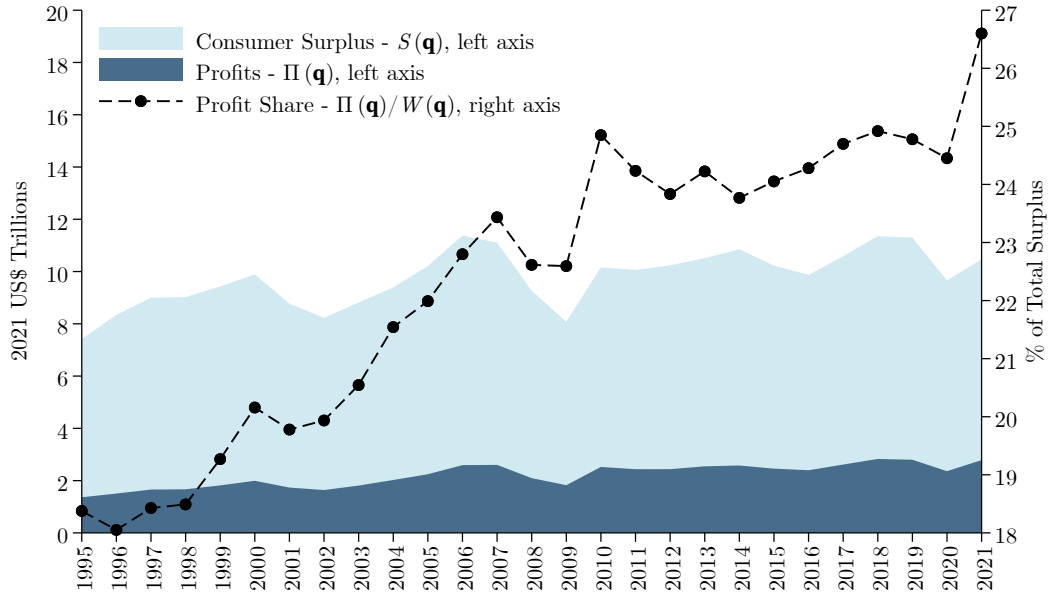
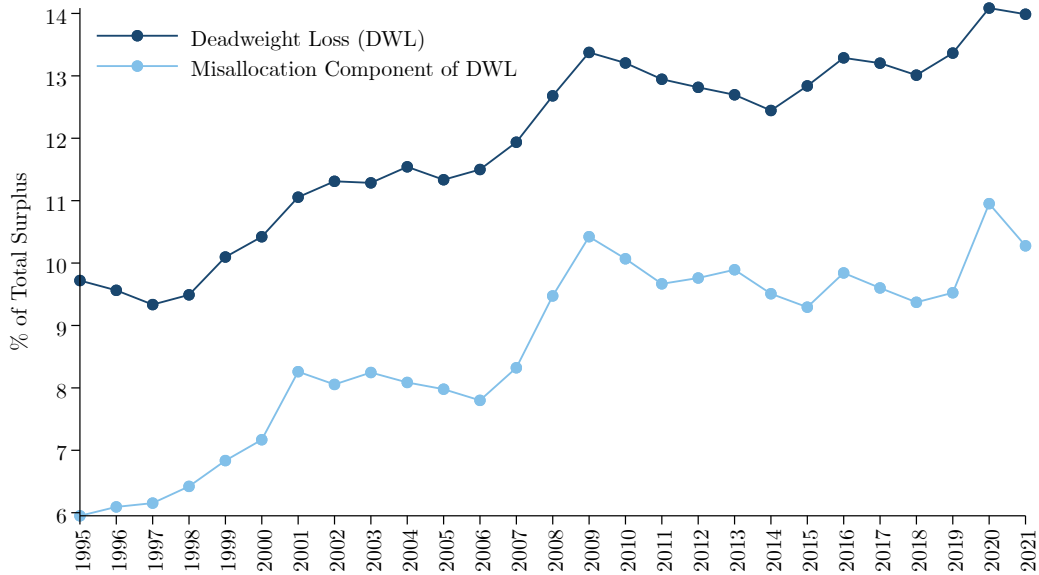


FIGURE 9: DEADWEIGHT LOSS FROM OLIGOPOLY (1995-2021)



G.4 Alternative Model with Multi-Product Firms

Figures 10 and 11 replicate, respectively, Figures 5 and 6 for the alternative model with multi-product firms described in I.G (data starts in 1999).

FIGURE 10: PROFIT AS % OF TOTAL SURPLUS (1999-2021)

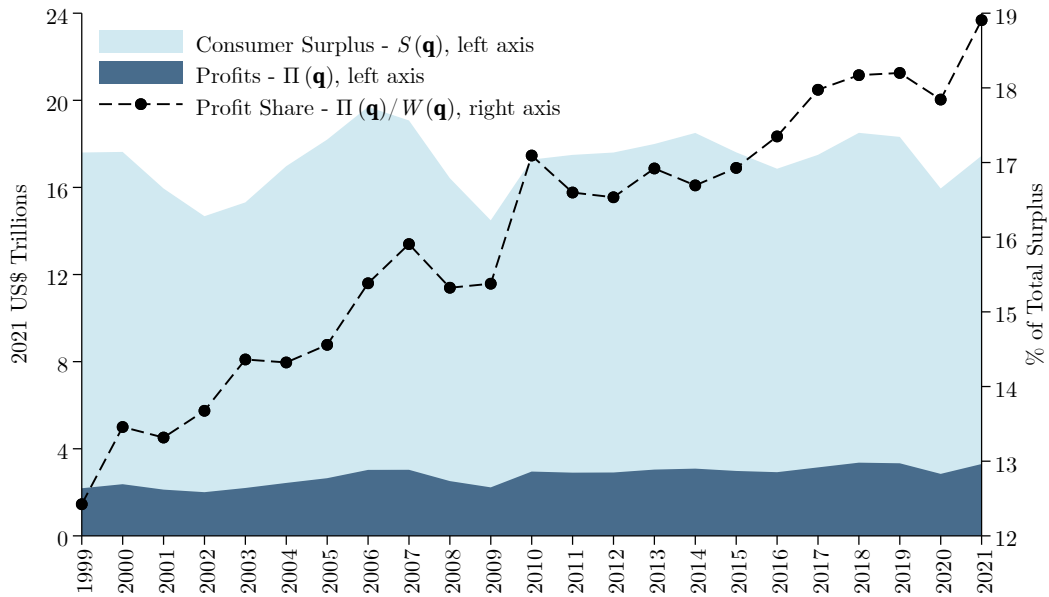
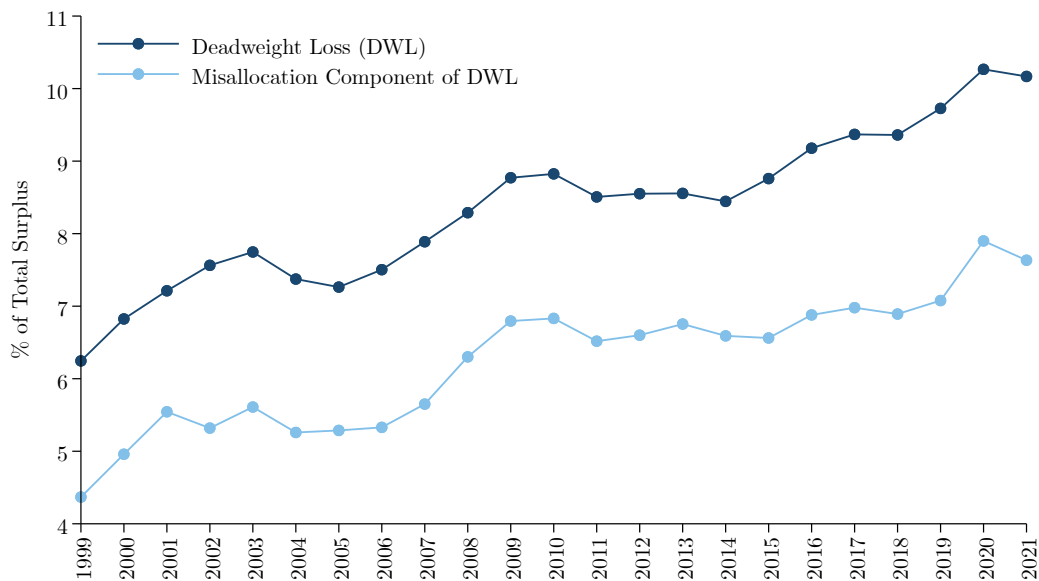


FIGURE 11: DEADWEIGHT LOSS FROM OLIGOPOLY (1999-2021)



G.5 Alternative Model with Input-Output Linkages

Figures 12 and 13 replicate, respectively, Figures 5 and 6 for the alternative model with input-output networks described in V.E (data ends in 2021).

FIGURE 12: PROFIT AS % OF TOTAL SURPLUS (1995-2021)

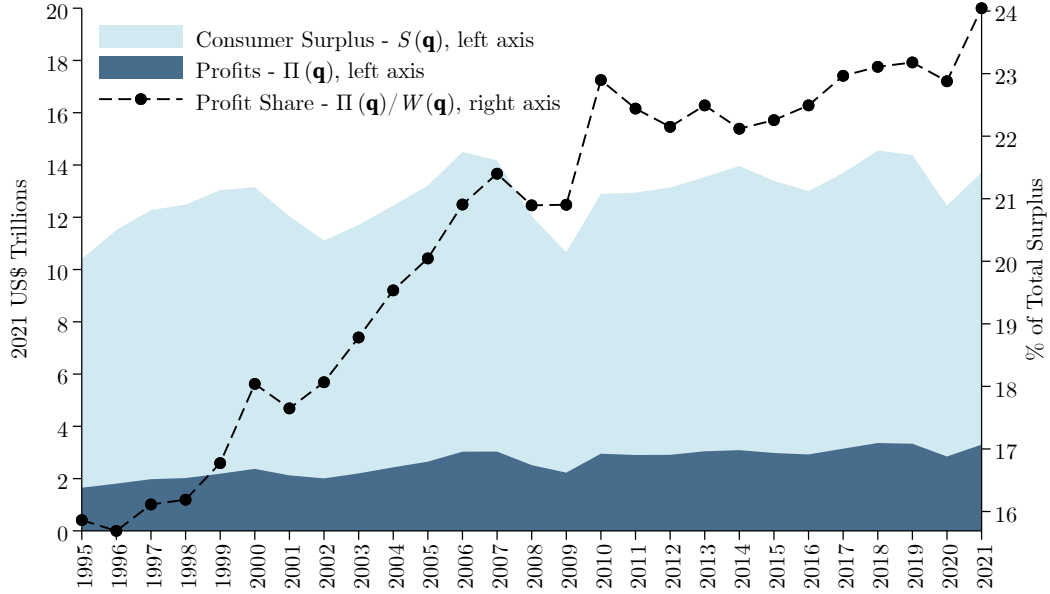
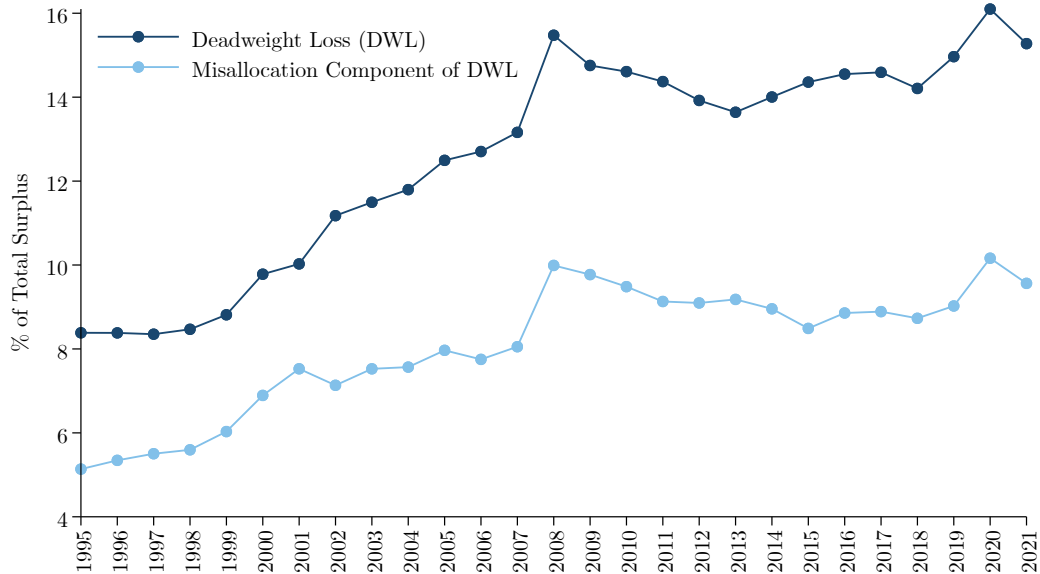


FIGURE 13: DEADWEIGHT LOSS FROM OLIGOPOLY (1995-2021)



G.6 Alternative Model with Operating Costs as Variable Costs

Figures 14 and 15 replicate, respectively, Figures 5 and 6 for the alternative model where Operating Costs are used to measure of total variable costs (fixed costs are assumed =0).

FIGURE 14: PROFIT AS % OF TOTAL SURPLUS (1995-2021)

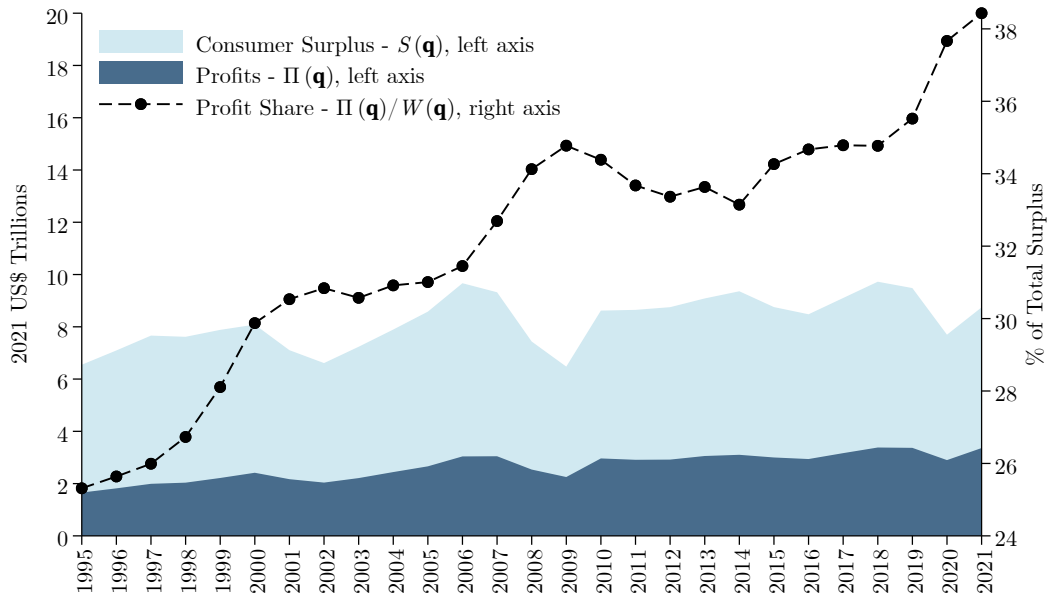
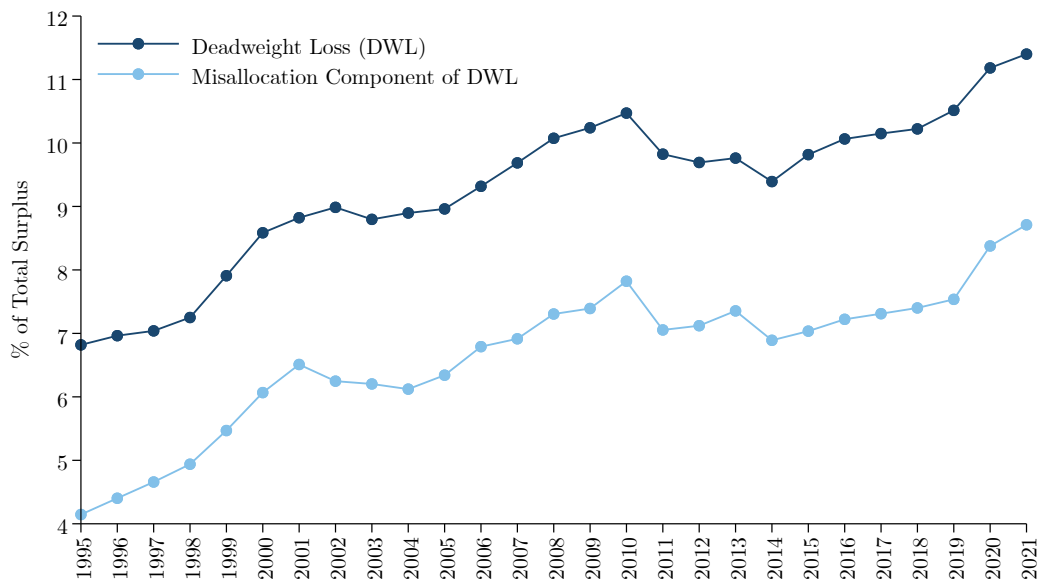


FIGURE 15: DEADWEIGHT LOSS FROM OLIGOPOLY (1995-2021)



G.7 Comparison of Cournot Oligopoly and Bertrand Oligopoly

Figures 16 and 17 compare, respectively, the profit share (of total surplus) and the deadweight loss of the baseline Cournot model against its Bertrand counterpart, presented in subsection V.G. Figure 16 replicates (for both) the dotted black line of Figure 5, while Figure 17 does the same with the blue line of Figure 6.

FIGURE 16: PROFIT AS % OF TOTAL SURPLUS (1995-2021)

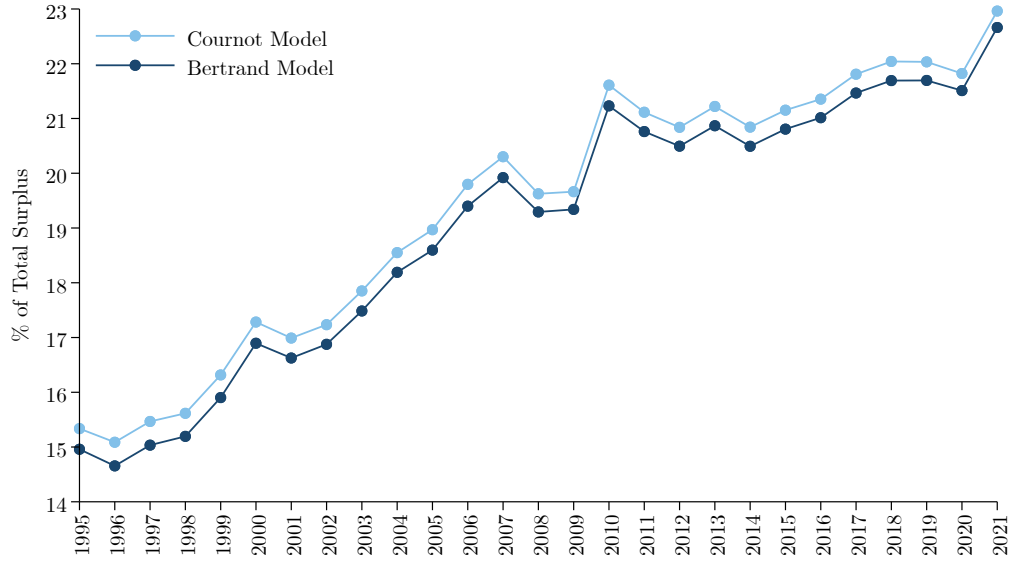
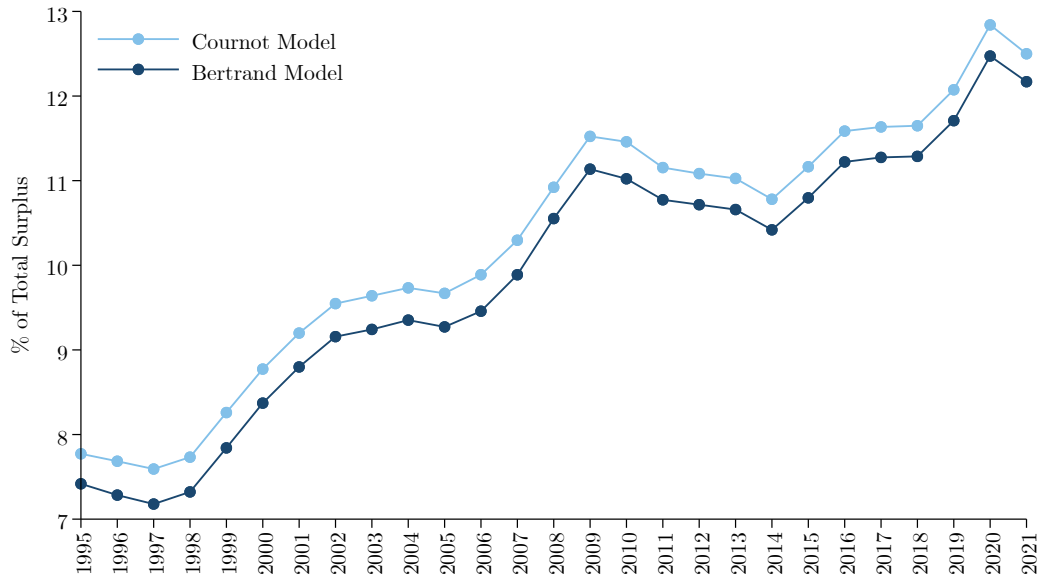


FIGURE 17: DEADWEIGHT LOSS FROM OLIGOPOLY (1995-2021)



H Comparison to Symmetric Benchmark 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:

$$\sigma_{ij} = \bar{\sigma} \forall (i, j) \quad (\text{H.1})$$

$\bar{\sigma}$ 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:

$$\bar{\sigma} = \frac{\mathbf{q}' \mathbf{\Sigma} \mathbf{q}}{\mathbf{q}' (\mathbf{1}\mathbf{1}' - \mathbf{I}) \mathbf{q}} \quad (\text{H.2})$$

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

$$b_i = p_i + q_i + d_i \quad (\text{H.3})$$

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 the same d_i (dashed vertical line).

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 be 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 I.D). 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 18: CROSS-SECTIONAL DISTRIBUTION OF DEGREE CENTRALITY AND DEMAND INTERCEPT (2021)

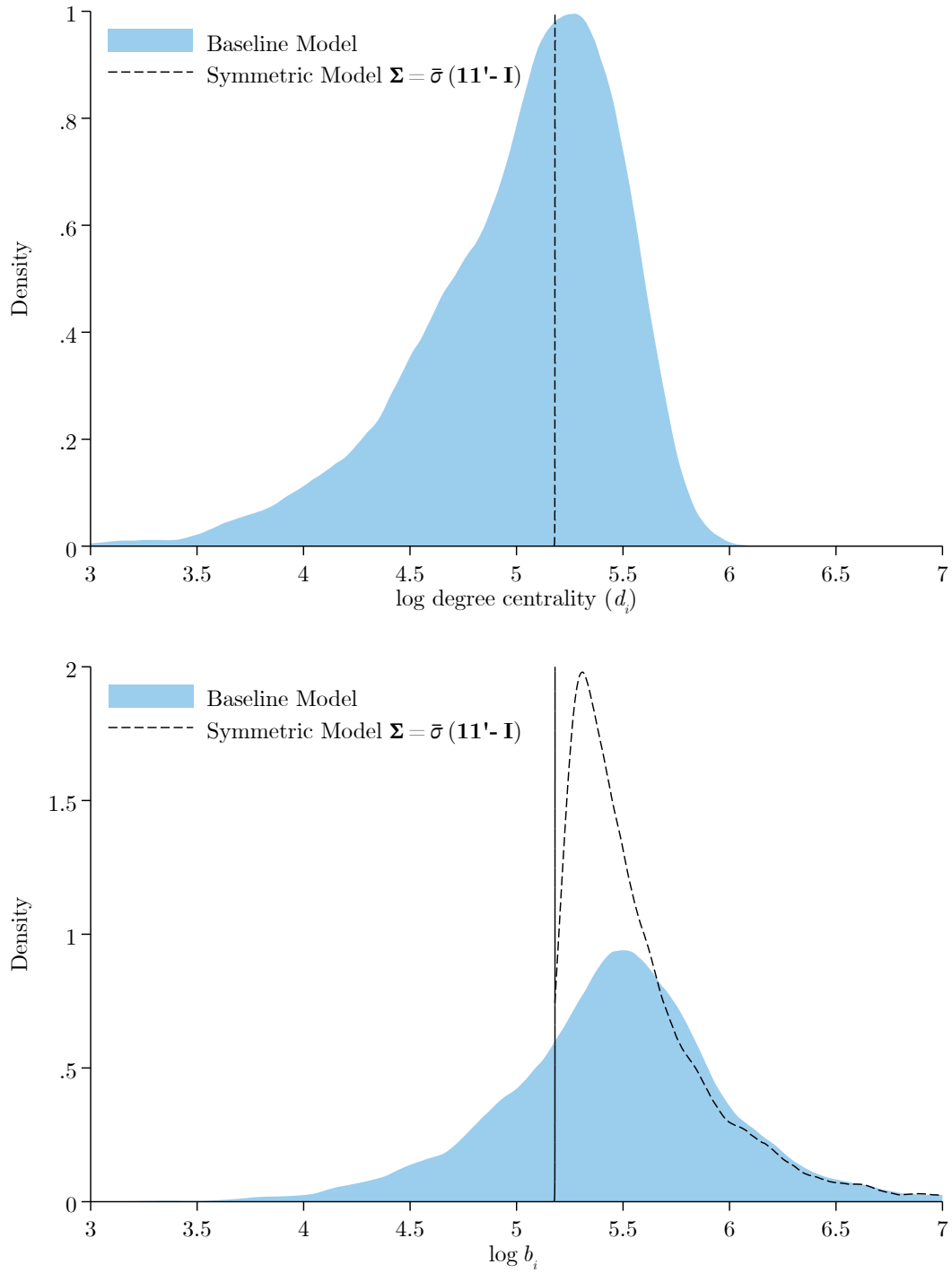


FIGURE NOTES: The figure above presents kernel density estimates of the cross-sectional distribution of the (log of) degree centrality (d_i - upper panel) and the demand intercept/quality (b_i - lower panel). Each observation is a firm, data is from the year 2021. The light blue area represents the distribution for the baseline model, while the dotted black line represents the distribution for an alternative model where $\Sigma = \bar{\sigma} (\mathbf{1}\mathbf{1}' - \mathbf{I})$.

I Proofs & Derivations

Proof to Lemma 1. Start from equation (19), evaluated at the Cournot equilibrium, and subtract $(\mathbf{I} + \mathbf{\Delta}) \mathbf{q}^\Phi$ from both sides to obtain:

$$\mathbf{p}^\Phi - \mathbf{c} - (\mathbf{I} + \mathbf{\Delta}) \mathbf{q}^\Phi = \mathbf{b} - \mathbf{c} - (2\mathbf{I} + \mathbf{\Sigma} + \mathbf{\Delta}) \mathbf{q}^\Phi \quad (\text{I.1})$$

next, notice that, by equation (29), the right hand side is zero. Move $-\mathbf{q}^\Phi$ to the right hand side with a positive sign:

$$\mathbf{p}^\Phi - \mathbf{c} - \mathbf{\Delta} \mathbf{q}^\Phi = \mathbf{q}^\Phi \quad (\text{I.2})$$

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 the equation (32). To prove the right hand side, re-write equation (29) as:

$$(2\mathbf{I} + \mathbf{\Sigma}) \mathbf{q} = \mathbf{b} - \mathbf{c} \quad (\text{I.3})$$

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

Proof to Lemma 2. Divide both sides of equation (35) by c_i and subtract one to obtain:

$$\frac{p_i}{c_i} = \frac{b_i + c_i - d_i}{2c_i} \quad (\text{I.4})$$

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. Using equation (32), we can rewrite equation (37) as

$$\mu_i - 1 = \frac{1}{2} (1 - \chi_i) (\omega_i - 1) \quad (\text{I.5})$$

plug the expression for $\omega_i - 1 = 2(\bar{\mu}_i - 1)$, from equation (35), to obtain:

$$\mu_i - 1 = (1 - \chi_i) (\bar{\mu}_i - 1) \quad (\text{I.6})$$

add 1 to both sides, take the (-1) out of the second parentheses and simplify to obtain equation (41). □

Proof to Proposition 3. Notice that the left hand side of equation (32) implies:

$$r_i = (p_i - c_i) q_i = q_i^2 \quad (\text{I.7})$$

$$\text{then} \quad \frac{r_i}{s_i} = \frac{2q_i^2}{q_i^2 + \sum_{j \neq i} \sigma_{ij} q_i q_j} = \frac{2q_i}{q_i + \sum_{j \neq i} \sigma_{ij} q_i} \equiv 2\mathcal{M}_i \quad (\text{I.8})$$

□

Proof to Proposition 6. We write the first order condition of firm z with respect to the output of product line i :

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

in vector form:

$$\mathbf{0} = \mathbf{b} - \mathbf{c} - (2\mathbf{I} + \mathbf{\Sigma} + \mathbf{K} \circ \mathbf{\Sigma}) \mathbf{q} \quad (\text{I.10})$$

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

Proof to Proposition 8. 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:

$$q(\zeta) = p_{n+1} \cdot \zeta \quad (\text{I.11})$$

hence labor supply and profits are given by:

$$h(\zeta) = \frac{p_{n+1}^2}{2} \cdot \zeta; \quad \pi(\zeta) = \frac{p_{n+1}^2}{2} \cdot \zeta \quad (\text{I.12})$$

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:

$$1 = \frac{p_{n+1}^2}{2} \cdot \zeta_{\min} \quad (\text{I.13})$$

Let us now compute aggregate output and aggregate labor

$$\begin{aligned} q_{n+1} &= \int_{\zeta_{\min}}^{\infty} q(\zeta) \, d\text{cdf}(\zeta) = p_{n+1} \int_{\zeta_{\min}}^{\infty} \zeta \cdot \text{pdf}(\zeta) \, d\zeta = \sqrt{\frac{2}{\zeta_{\min}}} \int_{\zeta_{\min}}^{\infty} \frac{\beta-1}{\zeta^{\beta}} \, dz = \sqrt{2} \zeta_{\min}^{\frac{1}{2}-\beta} \\ h_{n+1} &= \int_{\zeta_{\min}}^{\infty} h(\zeta) \, d\text{cdf}(\zeta) = \frac{p_{n+1}^2}{4} \int_{\zeta_{\min}}^{\infty} \zeta \cdot \text{pdf}(\zeta) \, d\zeta = \frac{1}{\zeta_{\min}} \int_{\zeta_{\min}}^{\infty} \frac{\beta-1}{\zeta^{\beta}} \, d\zeta = \zeta_{\min}^{-\beta} \end{aligned} \quad (\text{I.14})$$

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:

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

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

$$h_{n+1} = \frac{q_{n+1}^2}{2} \quad (\text{I.16})$$

$\square$

Proof to Proposition 9. The left hand side of equation (32) can be written as:

$$q_i^2 = p_i q_i - c_i q_i \quad (\text{I.17})$$

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:

$$(1 + \delta_i/2) q_i^2 = p_i q_i - \text{TVC}_i \quad (\text{I.18})$$

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

Proof to Proposition 11. We start by re-writing equation (20) as:

$$\alpha \cdot \cos_{\text{KQ}}^{\text{HP}} \cdot q_K q_Q = -\varepsilon_{\text{KQ}} p_K q_K \quad (\text{I.19})$$

Then, substituting q_K, q_Q using (91), with $(\delta_K = \delta_Q = 0 \forall i)$, we can write:

$$\alpha \cdot \cos_{\text{KQ}}^{\text{HP}} \cdot \sqrt{p_K q_K - \text{TVC}_K} \sqrt{p_Q q_Q - \text{TVC}_Q} = -\varepsilon_{\text{KQ}} p_K q_K \quad (\text{I.20})$$

We can then re-write the same equation for (j, i) instead of (i, j) , and sum the two to obtain:

$$2 \cdot \alpha \cdot \cos_{\text{KQ}}^{\text{HP}} \cdot \sqrt{p_K q_K - \text{TVC}_K} \sqrt{p_Q q_Q - \text{TVC}_Q} = -(\varepsilon_{\text{KQ}} p_K q_K + \varepsilon_{\text{KQ}} p_Q q_Q) \quad (\text{I.21})$$

Finally, we solve for α on the left hand side to obtain equation (96). $\square$

Proof to Proposition 12. The left bound is simply 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 :

$$c_i = \frac{\text{TVC}_i}{q_i} - \frac{\delta_i}{2} q_i = \frac{\text{TVC}_i}{\sqrt{\frac{p_i q_i - \text{TVC}_i}{1 + \delta_i/2}}} - \frac{\delta_i}{2} \sqrt{\frac{p_i q_i - \text{TVC}_i}{1 + \delta_i/2}} \geq 0 \quad (\text{I.22})$$

and rearrange the right hand side to obtain:

$$\delta_i \leq \frac{2 \cdot \text{TVC}_i}{p_i q_i - 2 \cdot \text{TVC}_i} \quad (\text{I.23})$$

$\square$

Proof to Proposition 13. Start by writing the vector of economic profits in terms of the price vector $\mathbf{p}$:

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

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

$$(\mathbf{I} + \boldsymbol{\Sigma})^{-1} = \mathbb{D} + \mathbb{O} \quad (\text{I.25})$$

Then we can write:

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

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

$$0 = (\mathbf{I} + \boldsymbol{\Sigma})^{-1} (\mathbf{b} - \mathbf{p}) - \mathbb{D}\mathbf{p} + \mathbb{D}\mathbf{c} \quad (\text{I.27})$$

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

$$0 = (\mathbf{I} + \mathbb{D}^{-1} + \boldsymbol{\Sigma}) \mathbf{q} - (\mathbf{b} - \mathbf{c}) \quad (\text{I.28})$$

the corresponding Bertrand potential is

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

and the Bertrand equilibrium is:

$$\mathbf{q}^\Psi = (\mathbf{I} + \mathbb{D}^{-1} + \boldsymbol{\Sigma})^{-1} (\mathbf{b} - \mathbf{c}) \quad (\text{I.30})$$

□

J Additional Figures

FIGURE 19: AVERAGE MARKUP VS. PROFIT RATE

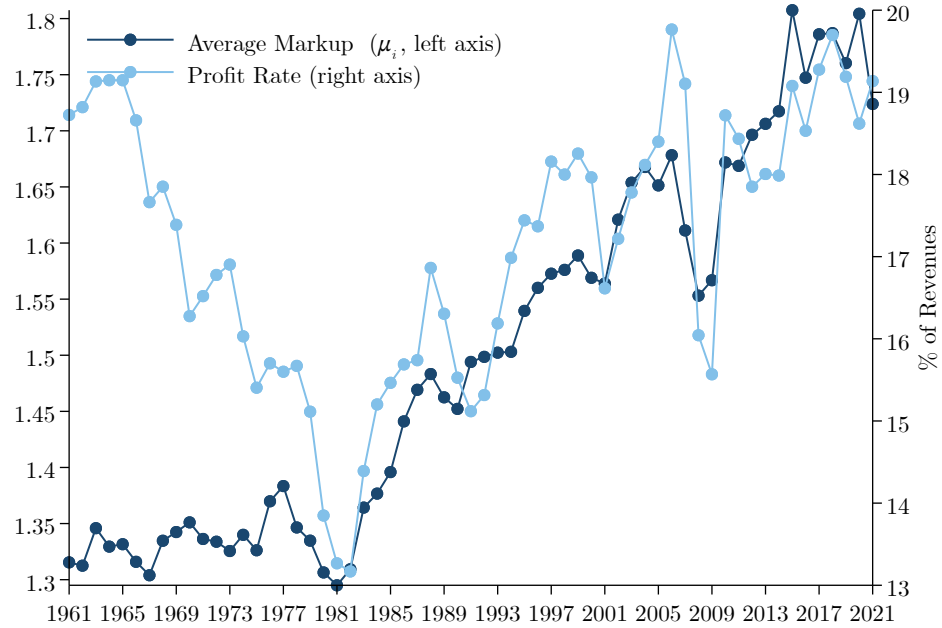


FIGURE 20: OPERATING VS. PRE-TAX PROFIT RATES

