Working Paper 24-033

Estimating Productivity in the Presence of Spillovers: Firm-level Evidence from the US Production Network

Ebehi Iyoha

Harvard Business School

Estimating Productivity in the Presence of Spillovers: Firm-level Evidence from the US Production Network

Ebehi Iyoha
Harvard Business School

Working Paper 24-033

This paper was previously circulated under the title "Production Function Estimation in the Presence of Spillovers." I am immensely grateful to Joel Rodrigue, Kamal Saggi, Yuya Sasaki, and Ben Munyan for their advice and support on this project. I have benefited greatly from feedback from Allan Collard-Wexler, Gautam Gowrisankaran, Atsushi Inoue, Tong Li, Yao Luo, Devesh Raval, David Rivers, Pedro Sant'Anna, Daniel Yi Xu, seminar participants at Vanderbilt University, Harvard Business School, Duke University, UC Berkeley, Virtual (IO)2 Seminar, NBER Productivity Lunch, Harvard-MIT Trade Lunch, as well as participants at the 2021 meetings of the European Association for Research in Industrial Economics, 2021 International Industrial Organization Conference and the Society for Economic Dynamics 2021 meetings. Many thanks to Enghin Atalay for sharing data with me. All errors remain my own.

Funding for this research was provided in part by Harvard Business School.

Estimating Productivity in the Presence of Spillovers: Firm-level Evidence from the US Production Network*

Ebehi Iyoha†

December 2023

Abstract

This paper examines the extent to which productivity gains are transmitted across US firms through buyer-supplier relationships. Many empirical studies measure firm-to-firm spillovers using firm-level productivity estimates derived from control function approaches. However, these methods implicitly rule out the interdependence of firms' outcomes and decisions through productivity spillovers. To address this limitation, I develop a framework to jointly estimate network effects and firm-level productivity, while accounting for common productivity shocks across firms and non-random buyer-supplier matching. Using this method, I characterize productivity spillovers over the US production network from 1977 to 2016. My results suggest that having 1% more productive trading partners on average leads to 0.076% higher productivity in the long run. Supplier spillovers, which are driven by both large and small firms, are 4 times greater than buyer effects, which are primarily generated by large firms. Heterogeneity in spillovers within and across sectors also has implications for overall productivity growth: aggregate spillovers tend to be much larger when manufacturers are central in the production network than when retailers and wholesalers are more central.

*This paper was previously circulated under the title "Production Function Estimation in the Presence of Spillovers." I am immensely grateful to Joel Rodrigue, Kamal Saggi, Yuya Sasaki, and Ben Munyan for their advice and support on this project. I have benefited greatly from feedback from Allan Collard-Wexler, Gautam Gowrisankaran, Atsushi Inoue, Tong Li, Yao Luo, Devesh Raval, David Rivers, Pedro Sant'Anna, Daniel Yi Xu, seminar participants at Vanderbilt University, Harvard Business School, Duke University, UC Berkeley, Virtual $(IO)^2$ Seminar, NBER Productivity Lunch, Harvard-MIT Trade Lunch, as well as participants at the 2021 meetings of the European Association for Research in Industrial Economics, 2021 International Industrial Organization Conference and the Society for Economic Dynamics 2021 meetings. Many thanks to Enghin Atalay for sharing data with me. All errors remain my own.

†Harvard Business School. E-mail: eiyoha@hbs.edu

1 Introduction

Production function estimation is at the heart of several important questions in economics. From examining changes to market power and assessing the impact of trade liberalization, to decomposing the sources of aggregate productivity growth, understanding firms' decisions and their implications on market outcomes often hinges on good estimates of firm-level total factor productivity (TFP).[1]

A significant finding of the literature on firm-level productivity is that businesses exhibit marked differences in TFP, even within narrowly-defined industries, and a vast body of work has sought to explain this dispersion.[2] One possible explanation is that firms may affect one another in ways that do not show up in the prices of intermediate goods and services; they may experience spillovers from knowledge transfers or agglomeration externalities. For example, in the trade literature, firms have been found to impact the productivity of counterparts through activities such as foreign direct investment (FDI) and exporting.[3] Javorcik (2004) found that FDI in Lithuania had a positive effect on the productivity of domestic firms through backward linkages, and Keller and Yeaple (2009) documented the existence of horizontal spillovers from multinationals to US firms. Likewise, Alvarez and López (2008) provided evidence from Chile of positive productivity spillovers from domestic and foreign-owned exporters on their suppliers, Alfaro-Urena et al. (2022) found TFP gains of $4\% - 9\%$ among Costa Rican firms after they began to supply to multinational corporations, and Amiti et al. (2023) documented significant productivity increases for Belgian firms supplying to superstar firms including multinationals, exporters and large firms.

The present paper quantifies the transmission of productivity gains through buyer-supplier relationships in the United States and examines how the existence of spillovers affects the measurement of TFP. I consider spillovers not just from firm activities, but directly from productivity as well. A firm's TFP could increase or decline due to the productivity of the firms with which it has a relationship. The expected direction of this effect is not immediately clear: firms may learn from their peers and become more productive or might free-ride on their trading partners' efficiency. Empirical investigations into direct efficiency spillovers are relatively new. Serpa and Krishnan

[1] See De Loecker and Syverson (2021) for a primer on firm-level productivity analysis.
[2] See Syverson (2011) for a review.
[3] See Keller (2010) for a review of the evidence on spillovers from FDI and exporting.

(2018) examined this question with data on firm-level buyer-supplier relationships in the US, whereas Bazzi et al. (2017) used input-output matrices to construct measures of the relationships between Indonesian firms. Both studies found that firms enjoy significant boosts to productivity from their relationships with more productive counterparts.

However, an important gap exists in the literature on productivity spillovers. Many studies assess the existence of spillovers using TFP estimates obtained from semi-parametric proxy variable/control function approaches. Introduced by Olley and Pakes (1996) and refined in Levinsohn and Petrin (2003), Wooldridge (2009), and Ackerberg et al. (2015) (hereafter OP, LP, Wooldridge and ACF respectively), these methods assume that a firm's future productivity depends only on its own past productivity and characteristics. Alternative methods such as Gandhi et al. (2020) that relies on first order conditions for identification also make the same assumption on the productivity evolution process. This implies that each firm's productivity evolves independently, and implicitly rules out the existence of anticipated spillovers.

The contributions of this paper are three-fold. First, I show that when productivity spillovers exist, failing to account for this interdependence could lead to biased estimates of production function elasticities and TFP. As De Loecker (2013), De Loecker et al. (2016), and Garcia-Marin and Voigtländer (2019) point out, our conclusions about what drives changes in productivity are sensitive to how it is measured. De Loecker (2013) showed that measuring TFP under standard assumptions can lead us to underestimate the impact of exporting on productivity. In Garcia-Marin and Voigtländer (2019), the downward bias in learning-by-exporting estimates arises from revenue-based productivity measures that cannot disentangle the lower prices firms charge upon entry into export markets from their increased efficiency. In this case, the direction of bias in spillover estimates is not as clear-cut. I find that, depending on the structure of the network and persistence of productivity over time, estimating spillovers on mismeasured TFP can lead us to *overestimate* network effects in some cases and *underestimate* them in others.[4]

Second, I propose a modification to standard control function and first order condition approaches that flexibly accounts for the presence of spillovers. To do so, I apply results from the peer effects and spatial econometrics literatures including Lee (2003); Bramoullé et al. (2009); Lee and Yu (2016), with an important distinction:

[4]See the online appendix in OA3 for Monte Carlo experiments demonstrating these biases.

these papers deal with observed outcomes, whereas I jointly estimate the outcome and spillovers. This comes at the cost of a few additional assumptions that are, nonetheless, compatible with both the standard production function and network effects frameworks. A few studies have explicitly allowed for cross-sectional dependence between firms through linkages in TFP estimation. Malikov and Zhao (2021) specified a framework in which a firm's productivity depends on the lagged productivity of nearby firms within the same industry to study inward FDI effects in China's electric machinery manufacturing sector. Merlevede and Theodorakopoulos (2023) examined the impact of intangibles between European firms in ownership networks by allowing affiliates to be influenced by the productivity of their parent companies and vice versa.

My proposed approach offers three advantages relative to prior approaches. First, I allow network effects to occur within the same period and do not impose that spillovers only occur with a lag. Contemporaneous spillovers have been found to matter empirically[5] and my proposed approach can plausibly capture these effects while also accommodating additional lagged effects. Second, I account for common shocks that are localized within the network, which typically confounds the estimation and interpretation of spillovers. Third, my approach can accommodate network formation that is endogenous to productivity shocks. The estimator is therefore useful in a broad range of contexts without overly strong assumption, and can be extended to examine heterogeneous spillovers in the manner of Dieye and Fortin (2017) and Patacchini et al. (2017), that vary by the nature of the relationship between firms and their characteristics.

Third, I apply this methodology to examine the transmission of efficiency gains through the production network of publicly listed firms in the United States from 1977 to 2016. I find evidence of positive productivity spillovers, with a stronger impact from suppliers to customers: a 1% increase in its average supplier's productivity raises a firm's long-run productivity by 0.083% whereas the customer effect would be 0.018%. Furthermore, while both large and small suppliers generate positive productivity spillovers on their customers, only large customers generate these gains for their trading partners. Decomposing these network effects by sector reveals substantial heterogeneity: retailers tend to be an important source of productivity gains to

[5]For example, (Keller, 2010) found that horizontal productivity spillovers from inward FDI in the US showed up both within the same period and with a one-year lag.

many other sectors, but do not benefit from their upstream relationships. Electronics manufacturers, on the other hand, both benefit from and generate spillovers both upstream and downstream. This sectoral variation in spillovers is positively associated with patent citation patterns and negatively associated with employee flows inter-sectoral employee flows.

My results highlight an additional channel for industrial policy to affect economic growth. Given that a substantial portion of these spillovers can be attributed to efficiency gains in distribution and information technology, policymakers could target high-growth sectors that can generate these second-order effects. Furthermore, there are significant differences in aggregate spillovers depending on which kinds of firms are most central in the US economy: during the decades when manufacturers were more central, aggregate spillovers from the 10 most central firms were often twice as large as after taking into account sector-specific effects, compared to a homogeneous network effect benchmark. By contrast, during the 2007-2016 period that featured greater centrality of retailers and wholesalers, aggregate spillovers were lower when accounting for sector-specific effects. This highlights another rationale for policymakers' concerns over the sectoral composition of the economy.

In the next section, I describe the data and features of the portion of the US production network that I observe. Section 3 presents my empirical framework and discusses the biases that arise from ignoring spillovers in the standard control function approach. In section 4, I propose a procedure for estimating production functions in the presence of various network effects and clarify the assumptions needed to obtain valid estimates. I introduce a model of network formation in section 5 to account for endogenous network selection. I consider extensions to the benchmark model including a gross output production function in section 6. Section 7 presents my empirical results and section 8 concludes.

2 Data: The US Production Network

I begin by describing the data with which I characterize the firm-level production network within the US, to highlight features that are important for my empirical methodology. To examine the magnitude and origins of productivity spillovers in the US, I rely on a panel of publicly-listed firms in the *Compustat* database from 1977 to 2016. *Compustat* collects companies' financial statements from form 10-K reports

Table 1: Firm Characteristics

	Mean	SD		Mean		Mean
Sales	6.08	21.03	Mining	0.063	Wholesale	0.041
Sales per 1000 employees	0.5	4.12	Utilities	0.044	Retail	0.044
Value Added	1.85	5.47	Construction	0.009	Transport and Warehousing	0.038
Capital stock	5.4	21.41	Durables Manufacturing	0.204	Information	0.090
Materials	4.37	17.56	Non-Durables Manufacturing	0.184	Finance, Insurance & Real Estate	0.035
Employees (thousands)	18.95	62.02	Electronics Manufacturing	0.176	Services	0.073
Large firm (employees $\geq$ 500)	0.68	-				
Observations	54557					

This table reports average characteristics of firms in the sample. All monetary values are in 2009 billion USD.

filed with the US Securities and Exchange Commission (SEC). This provides detailed information on firms' sales, capital stock, expenses and employees. I supplement this with industry-level deflators and wages from the US Bureau of Economic Analysis (BEA) to construct the necessary variables to estimate a production function.[6]

Information on buyer-supplier links also comes from 10-K reports. Statement no. 14 issued in December 1976 by the Financial Accounting Standards Board (FASB) requires each firm to report any customers that are responsible for 10% or more of its sales within a fiscal year. I conservatively match the reported customer names to company financial data. The resulting network contains 18,872 unique buyer-supplier pairs and 66,052 dyad-year observations.[7]

I restrict the firm-level sample to the businesses that either report or are reported as customers, and have positive values of sales, capital stock, labor, and materials. I discard firms in agriculture, forestry and fishing, because these industries have too few observations in both the firm- and dyad-level datasets. This yields an unbalanced panel of 8,353 firms and 55,557 firm-year observations.

Table 1 reports average firm characteristics by decade and over the full sample. Due to the nature of the firms in question, and the restriction to companies with customer or supplier data, firms in the sample tend to be large, averaging 19,000 employees and $6.08 billion in annual sales. Based on the BEA's classification of large enterprises as firms employing 500 or more workers, about two-thirds of the sample are large firms. As shown in table 1, manufacturers comprise more than half

[6] See section A in the appendix for further details on variable construction.

[7] Other studies that have used this dataset to study the US production network include Atalay et al. (2011), Lim et al. (2017) and Serpa and Krishnan (2018). I am grateful to the authors of Atalay et al. (2011) for graciously sharing their matched buyer-supplier data with me.

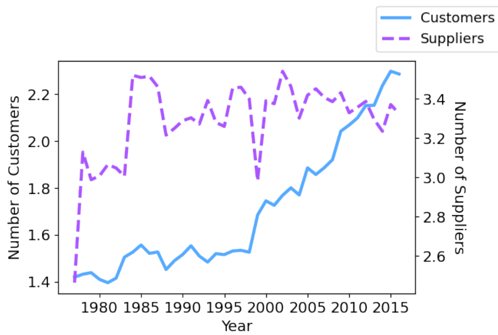


Figure 1: Average Firm Degree
Annual average out- and in-degrees (customers and suppliers) for firms in the sample.

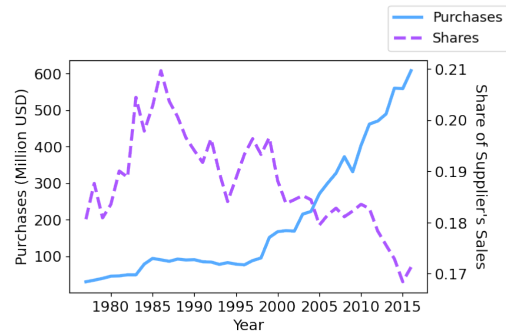


Figure 2: Value Traded in Relationships
Annual average value traded by each buyer-supplier pair and as a share of each seller's total sales.

of the firms in the sample. Information and Services are the next largest sectors represented in the sample.[8]

The observed sample of the production network is sparse; that is, the number of connections per firm is low. Figure 1 shows that firms report 1 or 2 customers on average, whereas the same customers are reported by about 3 or 4 suppliers. Consistent with the 10% sales reporting requirement, reported customers tend to be large; the average customer realizes about eight times as much in sales as the average supplier in the data (see figure 6). This may be due to two factors: relatively small firms are likely to have major customers and larger firms are likely to be major customers. However, although the value traded in the average reported relationship is sizable and increases over time, figure 2 indicates each individual relationship makes up a declining share of the suppliers' sales.

In figure 4, I examine features of the network that affect the identification of spillovers within my framework. Network density, measured by the number of observed links as a fraction of all possible links, does not exceed 0.28% in any year. The network gets sparser at the beginning of the sample and denser after the mid-90s. At the same time, network transitivity, the number of observed triads as a share of all possible triads, trends upward throughout the sample, but does not exceed 1.2%. In sections 3 and 4, I discuss the importance of density and transitivity for both the biases in input elasticities from standard approaches and the performance of my proposed estimator.

Each year, the production network is often dominated by a large cluster of firms

[8]See section A for a full list of industries in each sector.

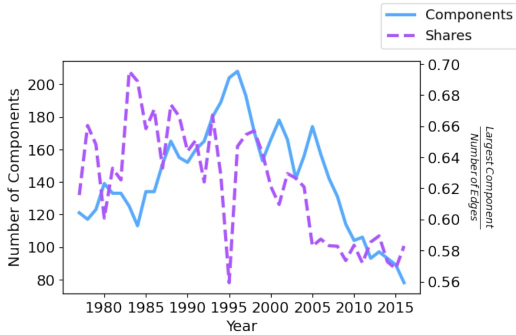


Figure 3: Clustering and Components
Number of connected components and largest component as a share of all edges in the network.

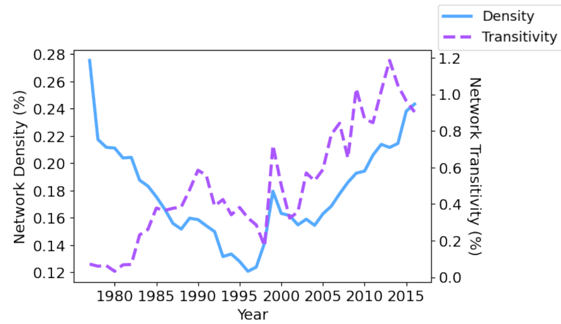


Figure 4: Density and Transitivity
Density and transitivity of the network sample over time.

connected to one another. Figure 3 shows that the number of edges in the largest connected component as a share of all edges in the network ranges from 56% to 70%. This is largely due to the presence of a few well-connected firms, whereas the remainder of the network consists of peripheral, small clusters.

Variations in clustering patterns over time reflect changes in the relative importance of each industry. Figure 5 reports the ten most central firms as measured by the number of links a firm has as a share of all observed links. In the first ten years of the sample, manufacturers of automotives and other durable goods dominated the list. In the next decade, AT&T rose to the top of the list, and electronics manufacturers such as IBM had begun to emerge. In the 1997-2006 period, Walmart rose to the top the list, and while automotive and electronics manufacturers still featured at the top of the centrality distribution, their centrality had declined relative to earlier decades. By the end of the sample, retailers and wholesalers had superseded most manufacturers, with Walmart continuing to top the list.

Figure 7 shows the relationship between a firm's labor productivity, as measured by the natural log of sales per employee and that of its average buyer or seller. The slope of the fitted regression line is 0.38, indicating a strong positive correlation between the two quantities. Interpreting this relationship requires distinguishing among several possible explanations. Foremost is the question of direction: does a firm become more efficient by learning from its neighbors, or does causation move in the opposite direction? And if a firm is simultaneously affected by and affecting its partners, how can one pin down the magnitude of the effect? On the other hand, this relationship may be driven by the sorting of firms; if more productive firms trade

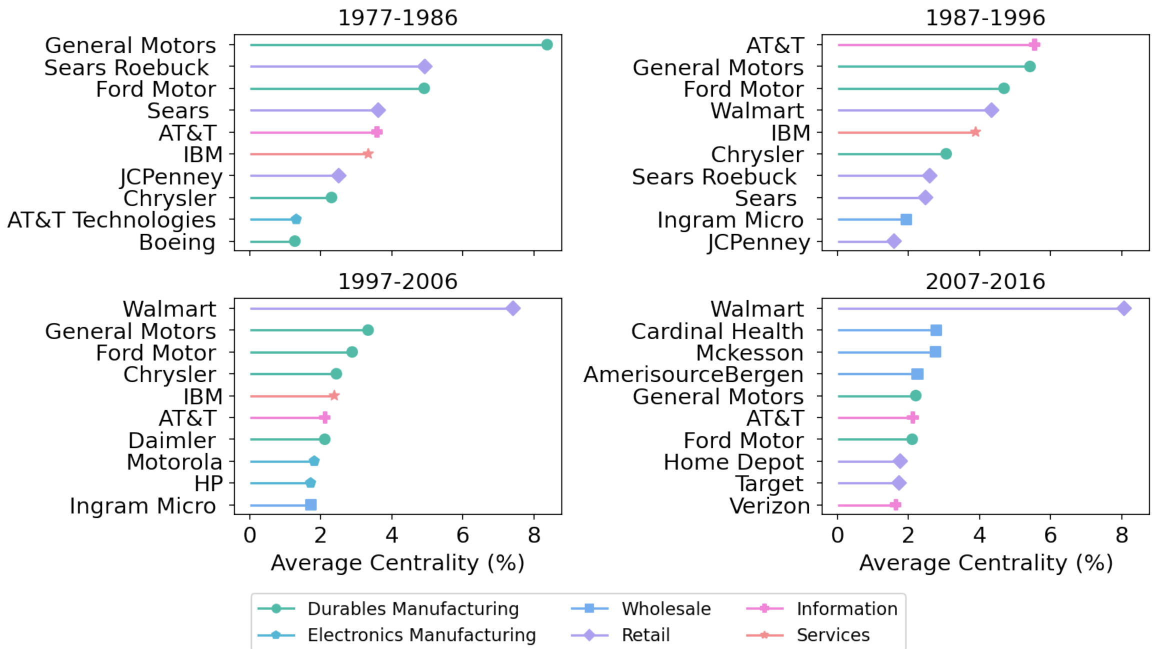


Figure 5: Firm Centrality

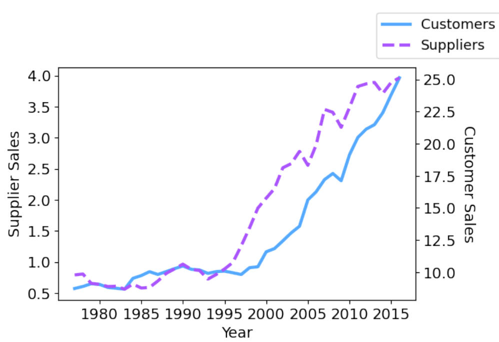


Figure 6: Customer and Supplier Sales

Annual average sales (in 2009 Billion USD) of customers and suppliers in the sample.

Figure 7: Firms' and Average Trading Partners' Labor Productivity

The slope of the fitted regression line is 0.38.

with one another, then this correlation is evidence of network formation rather than spillovers. Yet another possibility is that supply chains are a transmission channel for production and demand shocks, inducing the revenues of connected firms to move in the same direction.

Each of these explanations has different implications for how productivity is measured: if there are spillovers due to learning, then firms' input decisions will likely be influenced by the efficiency of their suppliers or buyers, whereas unanticipated common shocks are unlikely to affect input choices to the same degree. In the next section, I introduce an empirical framework with the goal of distinguishing among these channels, examining how they impact the measurement of TFP, and quantifying the direction and magnitude of productivity spillovers.

3 Empirical Framework

Consider a production technology for firm i in period t with Hicks-neutral productivity:

$$Y_{it} = F(L_{it}, K_{it})e^{\omega_{it}+\varepsilon_{it}} \tag{1}$$

where output, Y_{it} is a function of labor, L_{it} and capital, K_{it}. Output is shifted by an exogenous shock, $e^{\varepsilon_{it}}$ independent of all variables known to the firm by the end of the period, the information set, $\mathcal{I}_{it}$. $e^{\omega_{it}}$ is firm-specific TFP that is unobserved by researchers but known to the firm when making production decisions. $F(\cdot)$ is known up to some parameters. Taking the natural log of (1) yields:

$$y_{it} = f(l_{it}, k_{it}) + \omega_{it} + \varepsilon_{it} \tag{2}$$

The main limitation to estimating $f(\cdot)$ is a simultaneity problem: firms choose their inputs based on the realization of ω_{it}. Therefore, simply regressing a firm's output on its inputs would lead to a biased estimate of $f(\cdot)$.

To address this issue, the control function/proxy variable approach makes a set of assumptions on timing, a proxy variable and how productivity evolves over time. The existence of spillovers primarily poses a problem for the last set of assumptions.

Productivity is typically assumed to follow a first-order Markov process:

$$\omega_{it} = h(\omega_{it-1}) + \eta_{it} \tag{3}$$

where $h(\cdot)$ is unknown and η_{it} is mean independent of firm's information set at the beginning of the period $\mathcal{I}_{it-1}$. Suppose instead that ω_{it} is affected by some other firm j either through its past decisions $\boldsymbol{x_{jt-1}}$ and/or its current productivity, ω_{jt}:

$$\omega_{it} = h(\omega_{it-1}, \boldsymbol{x_{jt-1}}, \omega_{jt}) + \zeta_{it} \tag{4}$$

where $E[\zeta_{it} \,|\, \mathcal{I}_{it-1}] = 0$. The effect of $\boldsymbol{x_{jt-1}}$ represents spillovers from firm $j's$ activities such as research and development (R&D), FDI, exporting, etc. The inclusion of ω_{jt} indicates that j being more productive could contemporaneously influence i's productivity, and both firms' TFP may be jointly realized in the same period. Since firm j's TFP is also determined by its past productivity, ω_{jt-1}, this representation allows for productivity spillovers to have dynamic implications, while allowing for the possibility that firm i is also affected by random shocks, ζ_{jt} to j's productivity within the same period.

When researchers estimate TFP under the assumption in equation (3) whereas the true process is represented by equation (4), then the effect of firm j on i is attributed to η_{it}, which now violates the conditional independence assumption. In the following subsections, I examine the biases arising from standard control function approaches in greater detail.

Accounting for $\boldsymbol{x_{jt-1}}$ is fairly straightforward if we assume that it is known to i at the beginning of the period; that is, $\boldsymbol{x_{jt-1}} \in \mathcal{I}_{it-1}$. However, ω_{jt} poses a more serious challenge because it is jointly realized with ω_{it} and cannot therefore be assumed to be in $\mathcal{I}_{it-1}$. In section 4, I outline the assumptions needed to properly account for the effect of ω_{jt} on ω_{it} when estimating production functions.

3.1 Control Function Approach

Suppose $f(\cdot)$ takes the form of a simple structural value-added Cobb-Douglas production function as in Ackerberg et al. (2015):[9]

$$y_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it} \tag{5}$$

where y_{it}, k_{it}, and ℓ_{it} are the logs of value-added[10], capital, and labor respectively. Obtaining consistent estimates of $\boldsymbol{\alpha}$ and ω_{it} requires three sets of assumptions.

The first relates to the timing of firms' decisions. Capital is a state variable, determined in the preceding period as a deterministic function of the firm's previous capital stock and its investment decision: $k_{it} = \kappa(k_{it-1}, i_{it-1})$. Labor, on the other hand, may or may not have dynamic implications. It may be fully adjustable and chosen after productivity is realized, or partly (or wholly) determined in the previous period. It, however, needs to be chosen prior to the intermediate input decision. Based on its current capital stock, workforce and productivity, the firm chooses intermediate inputs according to the following function:

$$m_{it} = \mathbb{M}(k_{it}, \ell_{it}, \omega_{it})$$

Next, one needs to assume that the demand for materials, $g(\cdot)$ is strictly monotonic in productivity, and that productivity is the only unobservable component of the input demand function. This guarantees that TFP can be expressed solely as a function of observables $\omega_{it} = \mathbb{M}^{-1}(k_{it}, \ell_{it}, m_{it})$. Substituting into the production function yields:

$$y_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + \mathbb{M}^{-1}(k_{it}, \ell_{it}, m_{it}) + \varepsilon_{it} \tag{6}$$

Although α_k and α_ℓ are not identified in this equation, we can obtain consistent estimates of the firm's expected value-added:

$$E[y_{it}|\mathcal{I}_{it}] = \varphi_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} \tag{7}$$

[9] I choose ACF because it allows for relatively flexible assumptions on the data-generating process for output, capital, labor and materials. However, this critique applies more broadly to OP, LP, Wooldridge and first order condition approaches such as Gandhi et al. (2020) that rely on similar assumptions on the productivity evolution process.

[10] Output minus intermediate inputs.

This disentangles productivity from the idiosyncratic shock ε_{it}. In order to identify capital and labor elasticities, the evolution process for productivity must be specified. A standard assumption is that productivity follows a first-order Markov process given its information set $\mathcal{I}_{it-1}$ in the previous period:

$$\omega_{it} = h(\omega_{it-1}) + \eta_{it} \tag{8}$$

where $E[\omega_{it}|\mathcal{I}_{it-1}] = E[\omega_{it}|\omega_{it-1}] = h(\omega_{it-1})$. $h(\cdot)$ is known to the firm but unobserved by the researcher, while η_{it} is idiosyncratic. Given (7) I can write lagged productivity as:

$$\begin{aligned} \omega_{it-1} &= \varphi_{it-1} - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1} \\ \implies \omega_{it} &= h(\varphi_{it-1} - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1}) + \eta_{it} \end{aligned}$$

Substituting into the production function yields:

$$y_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + h(\varphi_{it-1} - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1}) + \eta_{it} + \varepsilon_{it}$$

Since $E[\varepsilon_{it}|\mathcal{I}_{it}] = 0$ and $E[\eta_{it}|\mathcal{I}_{it-1}] = 0$ by assumption, then we can identify α_k, α_ℓ based on the moment restriction:

$$\begin{aligned} &E[\varepsilon_{it} + \eta_{it}|\mathcal{I}_{it-1}] = 0 \\ &E[y_{it} - \alpha_k k_{it} - \alpha_\ell \ell_{it} - h(\varphi_{it-1} - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1})|\mathcal{I}_{it-1}] = 0 \end{aligned} \tag{9}$$

Using this equation, we can derive moment conditions to estimate the elasticities. Since, there are three unknowns, $(\alpha_k, \alpha_\ell, h(\cdot))$, a typical set of moments would be:

$$E[(\eta_{it} + \varepsilon_{it})k_{it}, \ell_{it-1}, \varphi_{it-1}] = 0 \tag{10}$$

3.2 Network Effects

To examine biases due to the existence of spillovers, we need to first understand how network effects are characterized. Within a given year, relationships between n_t firms result in a network. This can be represented by an $n_t \times n_t$ adjacency matrix A_t such that $A_{ij,t} = 1$ if firm i has a relationship with firm j in that year and zero otherwise. These relationships could be transactional (i sells inputs to j) or some other form of firm interdependence, such as i and j sharing a board member. The adjacency matrix need not be symmetric. As is standard in the peer-effects literature, I impose $A_{ii,t} = 0$ for all i so that a firm cannot have a spillover effect on itself.

In most examples, I focus on buyer-supplier networks, but this framework could apply to other types of inter-firm relationships.[11] Suppose we are interested in how upstream firms are affected by the productivity of their downstream network. Let N_{it} be the set of i's customers in period t and $n_{it} = |N_{it}|$.[12] We would like to estimate the following network effects equation:

$$\omega_{it} = \beta_1 + \rho\omega_{it-1} + \boldsymbol{x_{it-1}\beta_x} + \lambda\frac{1}{n_{it}}\sum_{j\in N_{it}}\omega_{jt} + \frac{1}{n_{it}}\sum_{j\in N_{it}}\boldsymbol{x_{jt-1}\beta_{\bar{x}}} + c_{\psi_t} + \zeta_{it} \tag{11}$$

where x_{it-1} is a $1 \times k$ vector of exogenous firm characteristics that could influence productivity, such as past R&D or exporting.

In this equation, there are three ways in which firm i's network could be related to its productivity. In the terminology of Manski (1993), the first channel is *endogenous network effects*: a firm's productivity is affected by the average productivity of its neighbors. This is measured by λ.

The second mechanism is *contextual effects* captured by $\boldsymbol{\beta_{\bar{x}}}$. Firms may be influenced by the characteristics or activities of their neighbors. For example, a firm's R&D could generate positive productivity spillovers on its business partners.

A firm's relationships could also result in *correlated effects*, productivity shocks common to all firms in a network cluster. Let ψ_t index the sub-components of a network in period t, that is firms who are at least indirectly connected to each other. Then c_{ψ_t} is a correlated effect for all firms in component ψ_t.

An underlying assumption here is that the network is exogenous; that is, firms do not select partners in ways that are systematically correlated with their productivity. For now, I abstract from network selection and address it in section 5.

For the rest of this discussion, it would be convenient to rewrite these equations in matrix notation. Define the interaction matrix G_t as the row-normalized form of

[11] Provided the network satisfies certain conditions for identification. See the rest of this section for details.

[12] Note that for some final goods producers and retailers, $n_{it} = 0$. These firms may not experience spillovers from others, but could still affect their suppliers.

A_t.[13] Equation (11) can be rewritten as:

$$\omega_t = \beta_1 \iota + \rho\omega_{t-1} + \boldsymbol{x_{t-1}\beta_x} + \lambda G_t \omega_t + G_t \boldsymbol{x_{t-1}\beta_{\bar{x}}} + c_{\psi_t} + \zeta_t \tag{12}$$

The reduced form is as follows:

$$\omega_t = (1 - \lambda G_t)^{-1} \left(\beta_1 \iota + \rho\omega_{t-1} + \boldsymbol{x_{t-1}\beta_x} + G_t \boldsymbol{x_{t-1}\beta_{\bar{x}}} + c_{\psi_t} + \zeta_t\right) \tag{13}$$

$|\lambda| < 1$ implies that we can represent $(I - \lambda G_t)^{-1}$ as a geometric series.

$$\omega_t = \sum_{s=0}^{\infty} \lambda^s G_t^s \left(\beta_1 \iota + \rho\omega_{t-1} + \boldsymbol{x_{t-1}\beta_x} + G_t \boldsymbol{x_{t-1}\beta_{\bar{x}}} + c_{\psi_t} + \zeta_t\right) \tag{14}$$

Bramoullé et al. (2009) proved that equation (12) is identified if the identity matrix I, G and G^2 are linearly independent. The presence of intransitive triads[14] guarantees that linear independence holds. Production networks naturally have this structure because supply-chains tend to be unidirectional. Therefore, if ω_t was observed, one could estimate equation (12) using 2SLS (Lee, 2003; Bramoullé et al., 2009), QMLE (Lee and Yu, 2016) or Bayesian methods in (Goldsmith-Pinkham and Imbens, 2013).

Measuring productivity adds a layer of complexity to the problem. A typical strategy, as in Javorcik (2004) and Serpa and Krishnan (2018), is to first obtain TFP values by estimating a production function such as using a method described above, and use these estimates in the network effects equation in (12). However, these approaches implicitly rule out the presence of spillovers, and the resulting TFP estimates are incompatible with the a wide set of network models nested in the peer effects model above.

3.3 Biases due to Network Effects

When productivity is affected by network effects, the independence assumption on the productivity shock is violated. However, the impact on the estimation of production function elasticities will differ by the type of effect.

Suppose TFP is estimated under the exogeneity assumption in equation (3) but

[13] $G_{ij,t} = 1/n_{it}$ if $A_{ij,t} = 1$ and zero otherwise.

[14] An intransitive triad in a graph is a set of nodes i, j, k, such that i is connected to j and j to k, but k is not connected to i.

the true process is given by equation (12). This implies:[15]

$$E[\eta_t|\mathcal{I}_{t-1}] = \boldsymbol{x_{t-1}}\boldsymbol{\beta_x} + \lambda G_t E[\omega_t \mid \mathcal{I}_{t-1}] + G_t\boldsymbol{x_{t-1}}\boldsymbol{\beta_{\bar{x}}} + E[c_{\psi_t} \mid \mathcal{I}_{t-1}]$$

In general, this expression is not equal to zero. $\boldsymbol{x_{t-1}}$ is a source of omitted variable bias but De Loecker (2013) and Gandhi et al. (2020) showed that the productivity process can be modified to account for its impact, provided $\boldsymbol{x_{t-1}}$ is in the firm's information set at the beginning of the period.[16] Contextual effects can be accounted for in the same way under similar assumptions. Assuming that network formation is exogenous, including $G_t\boldsymbol{x_{t-1}}$ in equation (3) would eliminate bias from this dimension.

$G_tE[\omega_t|\mathcal{I}_{t-1}]$ poses a serious challenge because in general, $E[\omega_t|\mathcal{I}_{t-1}] \neq 0$. Consider the correlation between neighbors' current productivity and current capital stock. Using the reduced form of $G_t\omega_t$:

$$E[G_t\omega_t \circ k_t] = E[G_t\,(1-\lambda G_t)^{-1}\,(\beta_1\iota + \rho\omega_{t-1} + \boldsymbol{x_{t-1}}\boldsymbol{\beta_x} + G_t\boldsymbol{x_{t-1}}\boldsymbol{\beta_{\bar{x}}} + \zeta_t) \circ k_t]$$

where $\circ$ is the Hadamard product.[17] Even though capital stock was determined in the previous period, it is still correlated with current productivity spillovers because productivity persists over time, and investment in the previous period was a function of productivity at the time. That is $k_(it) = \kappa(k_{t-1}, i_{t-1}(\omega_{t-1}))$ and therefore, $E[G_t\omega_t \times k_t] \neq 0$. The same argument can be made for labor which is a function of productivity in the same period: $l_{t-1}(\omega_{t-1}) \implies E[G_t\omega_t \circ \ell_{t-1}] \neq 0$.

The direction of bias will depend on the sign and size of λ and the relationship between capital, labor and productivity. For example, if networks generate positive productivity externalities and capital stock is increasing in productivity, then α_k will be biased upwards. If λ is small enough, then the size of bias will be minimal. TFP values will be underestimated but the direction of bias on λ is unclear.

On their own, correlated effects or network fixed effects do not introduce bias in the estimation of α_k and α_ℓ. Because the common component shocks are idiosyncratic each period, then k_t and ℓ_{t-1}, which were determined in the previous period are

[15] Here, I assume that G_t, $\{\boldsymbol{\omega_{jt-1}}\}_{j\in N_{it}}$ and $\{\boldsymbol{x_{jt-1}}\}_{j\in N_{it}}$ are in firm i's information set at the beginning of the period. I discuss this assumption explicitly in the next section.

[16] For example, as De Loecker (2013) noted, including a firm's current export status would not be valid because that is dependent on productivity in the same period; however, using its previous export status would satisfy this condition.

[17] Element-wise multiplication.

independent of c_{ψ_t}. However, to the extent that network components and links do not vary much over time, failing to account for $c_{\psi_t t}$ would bias the α_k and α_ℓ estimates.

To illustrate the bias from ignoring endogenous network effects, consider the following:

$$\omega_t = \rho(I - \lambda G_t)^{-1}\omega_{t-1} + (I - \lambda G_t)^{-1}\zeta_{it} = \rho \sum_{s=0}^{\infty} \lambda^s G_t^s \omega_{t-1} + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_{it} \tag{15}$$

Then the second stage of ACF is equivalent to estimating:[18]

$$\Delta^G y_t = \alpha_\ell \Delta^G \ell_t + \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{16}$$

Where $\Delta^G x_t = x_t - \rho \sum_{s=0}^{\infty} \lambda^s G_t^s x_{t-1}$, $\Delta_{x_t}^{err} = \rho \sum_{s=1}^{\infty} \lambda^s G_t^s x_{t-1}$ and $\Delta x_t = x_t - \rho x_{t-1} = \Delta^G x_t + \Delta_{x_t}^{err}$. This is equivalent to the dynamic panel approach in Blundell and Bond (2000). However, growth in output, labor and capital have been purged of the variation from network effects in the previous period. When we assume no spillovers, we estimate:

$$\Delta y_t = \alpha_\ell \Delta \ell_t + \alpha_k \Delta k_t + u_t \tag{17}$$

Therefore, in the linear AR1 case, ignoring spillovers is equivalent to introducing non-classical measurement error into both output and inputs. Bias from ignoring spillovers can also be characterized as an omitted variables problem. By estimating equation (17), where $u_t = \rho \sum_{s=1}^{\infty} \lambda^s G_t^s \omega_{t-1} + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \varepsilon_t$. That is, the standard ACF procedure succeeds in eliminating the endogeneity problem that arises from input decisions depending on the firm's own productivity, but is unable to account for the influence of its network's past productivity. In either case, an instrumental variable approach would help to eliminate the problem. The key would be to find variables that are correlated with changes to labor and capital but uncorrelated with output, particularly the input choices and output of other firms.

In the OP/LP case where the labor elasticity is consistently estimated in the first

[18] See section OA1 for derivation.

stage, the second stage is equivalent to estimating:

$$\Delta^G \widetilde{y}_t = \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{18}$$

where $\widetilde{y}_t = y_t - \widehat{\alpha}_\ell \ell_t$. Then by estimating $\Delta \widetilde{y}_t = \alpha_k \Delta k_t + u_t$ under the standard assumption of no-spillovers:

$$plim \ = \alpha_k \left(1 - \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s k_{t-1})}{var(\Delta k_t)}\right) + \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s \widetilde{y}_{t-1})}{var(\Delta k_t)} \tag{19}$$

On one hand, α_k is re-scaled by the covariance between the firm's capital growth and its network's previous capital. If this covariance is positive, then it would shrink $\widehat{\alpha}_k$ or even reverse its sign. Higher ρ will increase the attenuation factor, as will λ if it is positive. When λ is negative, it leads to an alternating series that dampens attenuation. The network structure also plays a role: when long chains exist, $G_t^s k_{t-1} > 0$ even for high values of s. By contrast, a network in which firms are paired off, so that the longest chain has a length of 1. Then $G_t^s k_{t-1} = 0$ for all $s > 1$ and attenuation would be lower under this scenario.

On the other hand, another source of bias exists that depends on the covariance between the firm's capital growth and its network's previous output purged of the variation from labor. When this covariance is positive, $\widehat{\alpha}_k$ overestimates α_k, and the effects of ρ, λ and G_t now work in the opposite direction. Depending on the signs and magnitudes of these covariances, it is possible to obtain estimates of α_k close to the true value if the two opposing effects cancel one another out.

Even in this simplified setting, the direction and magnitude of bias are not easily predictable *ex-ante*. This means that one cannot merely apply a bias correction to estimates obtained under standard assumptions. It motivates a modification to the estimation procedure that can flexibly account for a range of productivity processes and network effects. I propose a modification to the ACF procedure that achieves this with few additional assumptions.

4 Accounting for Spillovers

4.1 Endogenous and Contextual Effects

Assuming network exogeneity and no correlated effects, I write a more general form of the linear-in-means equation (12) above:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda G_t \omega_t + \zeta_t \tag{20}$$

Note that $h(\cdot)$ is unknown and can be estimated using a polynomial approximation. This allows for flexible interactions between ω_{t-1}, $\boldsymbol{x_{t-1}}$, and $G_t \boldsymbol{x_{t-1}}$. The key requirement is that the endogenous effect enters linearly. This leads to the reduced form:

$$\omega_t = (I - \lambda G_t)^{-1} h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + (I - \lambda G_t)^{-1} \zeta_t \tag{21}$$

$|\lambda| < 1$ implies that we can approximate $(I - \lambda G_t)^{-1}$ by a geometric series.

$$\omega_t = \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t \tag{22}$$

This yields a consistent estimate of the conditional expectation of TFP:

$$E[\omega_t | \mathcal{I}_{t-1}] = \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) \tag{23}$$

because the resulting error term satisfies the mean independence condition:

$$E\left[(I - \lambda G_t)^{-1} \zeta_t \middle| \mathcal{I}_{t-1}\right] = E\left[\sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t \middle| \mathcal{I}_{t-1}\right] = 0$$

Note that equation (22) also indicates how λ can be identified. Given the reduced-form equation, $G_t\omega_t$ can be written as:

$$\begin{aligned} G_t\omega_t =& G_t h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \sum_{s=1}^{\infty} \lambda^s G_t^{s+1} h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) \\ &+ \sum_{s=0}^{\infty} \lambda^s G_t^{s+1} \zeta_t \end{aligned} \tag{24}$$

Provided productivity is sufficiently persistent, we can use the current network's past productivity $G_t\omega_{t-1}$ as an instrument for the impact of the network's current productivity $G_t\omega_t$. This is because a firm is only affected by its current neighbors' past productivity through the neighbors' current productivity. Therefore, λ is identified from the variation in $G_t\omega_t$.

Equation (24) indicates that there are additional instruments are available to identify the endogenous network effect. These are more common in the network effects literature and rely on the existence of intransitive triads in the network (Lee, 2003; Bramoullé et al., 2009). For example $G_t^2\omega_t$ and $G_t^2\boldsymbol{x_{t-1}}$ is one set of possible instruments because G_t^2 captures the neighbors of a firm's neighbors, and these indirect connections affect the firm only through the firm's direct relationships.

Note however, that the relevance of these additional instruments relies on the strength of the endogenous effect. Whereas $G_t\omega_{t-1}$ is a good instrument as long as productivity is persistent, $G_t^2\omega_{t-1}$ requires both persistence and $|\lambda| > 0$ while $G_t^2\boldsymbol{x_{t-1}}$ requires that both endogenous and contextual network effects be nonzero.

Substituting the reduced form equation into the vectorized production function:

$$\begin{aligned} y_t =& \alpha_k k_t + \alpha_\ell \ell_t + \sum_{s=0}^{\infty} \lambda^s G_t^s h(\varphi_{t-1} - \alpha_k k_{t-1} - \alpha_\ell \ell_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) \\ &+ \varepsilon_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t \end{aligned} \tag{25}$$

Accounting for network effects in the estimation procedure comes at the cost of additional assumptions. First, as seen above, is that the endogenous effect enters the productivity process linearly. This would not hold if spillovers are non-monotonic. For example, if firms are likely to free-ride on very productive neighbors and are also negatively affected by very unproductive networks, but are able to learn from moderately productive firms, then the linearity assumption would not hold. However, there is reason to believe that linearity is, at the very least, a good approximation for understanding the network effect and it is a common assumption in the peer effects literature. Furthermore, one need not assume linearity if endogenous spillovers are not contemporaneous. For example, if we assume firms are affected by the past productivity of the previous network ($G_{t-1}\omega_{t-1}$), or the past productivity of their current network ($G_t\omega_{t-1}$), then either of these terms could enter $h(\cdot)$ non-linearly without posing a problem for identification.

Second, we need to assume that $\{G_{i,jt}\}_{j\in N_{it}}$ is in the firm's information set $\mathcal{I}_{it-1}$ at the beginning of the period. This is consistent with a network that is fixed over time: $G_t = G \; \forall \; t = 1...T$ or any network formation processes that takes place at the beginning of every period before productivity is realized. For example, in the context of production networks, if all firms choose their suppliers at the beginning of each year, this condition would be met. The key here is the timing: firms make production decisions based on their realized productivities inclusive of spillovers. In addition, $\omega_{jt-1}, \boldsymbol{x_{jt-1}} \in \mathcal{I}_{it-1} \forall j \in N_{it}$. That is, firms can observe the past productivity and decisions of their neighbors. This likely holds true for buyer-supplier relationships in which buyers often conduct due diligence on future suppliers, and would need to be examined in other contexts such as geographic proximity, family networks, affiliate relationships, interlocking boards, and so on.

Third, I assume that correlations between the TFPs of connected firms are generated by spillovers rather than common shocks. I relax this assumption in the next section.

Finally, this procedure requires that G_t is exogenous, that is, network formation and productivity are not driven by factors that firms observe but we do not. This assumption can also be relaxed but will require the network formation process to be specified. I do so in section 5.

4.2 Correlated Effects

Although network fixed effects alone do not bias the estimates of capital and labor elasticities, if endogenous or contextual spillovers are also present, failing to account for common shocks will lead to the mismeasurement of TFP. Therefore, given a productivity process with a component-year-specific fixed effect:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \lambda G_t\omega_t + c_{\psi_t} + \zeta_t \tag{26}$$

c_{ψ_t} can be eliminated by differencing using a matrix J_t such that $J_t)c_{\psi_t} = 0$. Bramoullé et al. (2009) suggest two ways to define J_t. The first is *within local differencing* by setting $J_t = I - G_t$. This subtracts the mean of a firm's neighbors' variables from the its own. An alternative would be *global differencing*, which subtracts not just the mean of a firm's neighbors, but all the firms in the component. That is, define J_t such that $H_{ij,t} = 1 - \frac{1}{n_{\psi_t}}$ if $i, j \in \psi_t$ and 1 otherwise.

Local differencing would suffice in an undirected network because if two firms are linked, then the link is reported in $G_{ij,t}$ and $G_{ji,t}$. However in directed networks, there may be some firms that are in the same sub-component and are therefore facing component-specific shocks but $\sum_{j\in N_{it}} G_{ij,t} = 0$, because the firm only has connections coming from one direction. For example, in a study of how customers affect the productivity of their suppliers, firm i may be a final goods producer whose productivity generates upstream spillovers but does not supply to any downstream firms. Yet it would be exposed to any shocks that affects the entire supply chain. If edges in G_t are classified as links from suppliers to customers, $G_{ij,t} = 0\,\forall j$ and $(I - G_t)c_{\psi_t} = c_{\psi_t}$. In this case, local differencing would not eliminate the correlated effect, but global differencing would.

When J_t is chosen appropriately, then transforming equation (26) yields the reduced form:

$$J_tG_t\omega_t = \sum_{s=0}^{\infty} \lambda^s J_t G_t^s h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \sum_{s=0}^{\infty} \lambda^s J_t G_t^s \zeta_t$$

Note that differencing the productivity process will require that the production function be transformed as well. That is:

$$\begin{aligned} J_ty_t =& \alpha_k J_t k_t + \alpha_\ell J_t \ell_t + \sum_{s=0}^{\infty} \lambda^s J_t G_t^s h(\varphi_{t-1} - \alpha_k k_{t-1} - \alpha_\ell \ell_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) \\ &+ \sum_{s=0}^{\infty} \lambda^s J_t G_t^s \zeta_t + J_t\varepsilon_t \end{aligned} \tag{27}$$

4.3 Estimation Procedure

I summarize my benchmark estimation procedure and outline modifications to deal with correlated effects. Estimation is a two-stage process. The first stage is the same as in Ackerberg et al. (2015). Estimate equation (6): $y_t = \alpha_k k_t + \alpha_\ell \ell_t + \mathbb{M}^{-1}(k_t, \ell_t, m_t) + \varepsilon_t$, using a polynomial approximation.[19] This yields estimates $\widehat{\varphi}_t = y_t - \widehat{\varepsilon}_t$.

In the second stage, estimate equation (25) by GMM with $k_t, \ell_t, \widehat{\varphi}_{t-1}, G_t\widehat{\varphi}_{t-1}$ as instruments. Alternatively, to reduce computational complexity, one can concentrate out the parameters in $h(\cdot)$ and proceed as follows. Start with guesses of the production function elasticities: $\alpha_k^*, \alpha_\ell^*$ and compute $\omega_t^* = \widehat{\varphi}_t - \alpha_k^* k_t - \alpha_\ell^* \ell_t$. Estimate the productivity process by 2SLS:

$$\omega_t^* = h(\omega_{t-1}^*, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \lambda G_t \omega_t^* + u_t \tag{28}$$

with a polynomial approximation of $h(.)$ and $[G_t\omega_{t-1}, G_t^2\omega_{t-1}, G_t^2\boldsymbol{x_{t-1}}]$ as instruments for $G_t\omega_t$. Using predicted values, $E[\omega_t^*|\mathcal{I}_{t-1}]$ from the regression, compute the residual in the productivity process:

$$u_t^* = \omega_t^* - h^*(\omega_{t-1}^*, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) - \lambda^* G_t \omega_t^*$$

Then solve for a new set of $(\alpha_k^*, \alpha_\ell^*)$ that satisfy the sample moment conditions:

$$E_{nt}[u_t^* \circ k_t, \ell_{t-1}] = 0 \tag{29}$$

Iterate through all steps of the second stage until the parameters converge to values $[\widehat{\alpha}_1, \widehat{\alpha}_k, \widehat{\alpha}_l]$. The corresponding second stage parameters, $\widehat{\lambda}$ and the parameters in $\widehat{h}(\cdot)$ are consistent estimates of network effects. Standard errors can be obtained by residual-based or vertex bootstrapping.[20]

To account for correlated effects, estimate the first stage as in the benchmark procedure, and apply the J_t transformation to all variables in the second stage.

[19]Like ACF, this estimation procedure can be used with other value-added production function specifications such as the translog.

[20]See section C in the appendix for details on bootstrapping network data.

5 Network Endogeneity

So far, I have assumed that the network is exogenous, but it is also possible that a firm's productivity may be correlated with how it forms relationships. This issue is reminiscent of the selection problem in Olley and Pakes (1996) – firms are only observed if their productivity is above some threshold. In this case, observed interfirm relationships may depend on TFP. To address this issue, I incorporate the network selection model in Arduini et al. (2015) and Qu et al. (2017) into the benchmark estimation procedure above.

5.1 Network Selection Model

Endogenous network formation as modeled by Qu et al. (2017) and Arduini et al. (2015) highlights a possible link between a firm's TFP and the nature of its network. Shocks to productivity are correlated with the chances of meeting potential partners. For example, firms that are better able to search for buyers or suppliers may also be more productive. In this case, a positive relationship between a firm's TFP and its networks' TFP or choices would be a result of the improved search rather than any spillovers.

At the beginning of each period, firms i and j consider the surplus of a link $V_i(A_{ij,t})$. Both firms want to form a link if $V_i(A_{ij,t} = 1) - V_i(A_{ij,t} = 0) > 0$.[21] I parametrize this difference in surplus as:

$$V_i(A_{ij,t} = 1) - V_i(A_{ij,t} = 0) = U_{ijt}(\gamma) + \xi_{ijt}$$

where ξ_{ijt} is i.i.d and follows a logistic distribution.

$$U_{ijt}(\gamma) = \gamma_1 + \boldsymbol{z_{it}\gamma_i} + \boldsymbol{z_{jt}\gamma_j} + \boldsymbol{z_{ijt}\gamma_{ij}} + \gamma_h H_{ijt} \tag{30}$$

Note that despite the slight abuse of notation, $\boldsymbol{\gamma_i, \gamma_j, \gamma_{ij}}$ are not random coefficients. They are parameters whose subscripts denote that they correspond to i, j or the dyad's characteristics.

[21] This model can apply to both directed and undirected networks. For example, in a buyer-supply network, the the surplus from i supplying j would be considered differently from the reverse direction.

$\boldsymbol{z_{it}}$ may include ω_{it-1}, x_{it-1} and other variables such as industry that influence a firm's relationship decision but may have no direct impact on productivity. $\boldsymbol{z_{ijt}}$ usually includes the distance between i and j's characteristic, $|\boldsymbol{z_{it}} - \boldsymbol{z_{jt}}|$ or some other dyad-specific measures, such as the physical distance between the firms, industry input-output shares, etc. A negative coefficient on $|\boldsymbol{z_{it}} - \boldsymbol{z_{jt}}|$ indicates that firm i wants to match with firms that are similar. H_{ijt} measures past linkages; a large and positive γ_h indicates that firm i prefers to stick with its previous partners. Past linkages can be specified broadly; for example, $H_{ijt} = A_{ij,t-1}$ would mean that firm i only considers linkages from the previous period, whereas $H_{ijt} = \mathbb{1}\left(\sum_{s=1}^{m} A_{ij,t-s} > 0,\ m \leq t\right)$ measures whether i and j were connected in any of the last m periods.[22]

The probability that a link $A_{ij,t}$ forms is given by:

$$P(A_{ij,t} = 1 \mid \mathcal{I}_{t-1}) = P(U_{ijt}(\gamma) + \xi_{ijt} > 0) = \frac{e^{U_{ijt}(\gamma)}}{1 + e^{U_{ijt}(\gamma)}}$$

The specified model, coupled with a logistic distribution implies that, conditional on firm and dyad characteristics, historical connectivity, and the unobserved ξ_t, the probability that i wants to form a link with j is independent of its decision to connect with some other firm k. While this may be restrictive, it is analytically and computationally tractable, and still manages to capture important features of real-world networks.

For example, this model allows for the possibility that a firm can choose multiple partners; firm i need not prefer j to all other firms, it just needs to prefer matching with j to not matching. This is useful for characterizing production networks, in which a non-negligible number of firms trade with more than one partner. As in Goldsmith-Pinkham and Imbens (2013), this model can also accommodate some interdependence in the linking decision through the choice of variables such as the number of links in the previous period, whether the firms had neighbors in common etc.

Network endogeneity arises from the relationship between ξ_{ijt} and the error term in the productivity process, ζ_{it}. Let $\xi_{it}^{'} = \{\xi_{ijt}\}_{j \neq i}^{n_t}$ be a row vector of the error terms from all the dyadic regressions with links originating from i. $(\zeta_{it}, \xi_{it}^{'}) \sim i.i.d.(0, \Sigma_{\zeta\xi})$ where $\Sigma_{\zeta\xi} = \begin{pmatrix} \sigma_{\zeta}^2 & \sigma_{\zeta\xi}\iota^{'} \\ \sigma_{\zeta\xi}\iota & \Sigma_{\xi} \end{pmatrix}$ is positive definite, σ_{ζ}^2 and $\sigma_{\zeta\xi}$ are variance and covariance

[22] There are alternative models such as Graham (2017) that include firm-year fixed effects in the dyadic regression model. Estimation of such models will depend on the sparsity of the network.

scalars, ι is an $n_t - 1$ column vector of ones, and $\Sigma_\xi = \sigma_\xi^2 I_{n_t-1}$. Stacking all the ξ_{it}'s in a matrix:

$$\Xi_t = \begin{bmatrix} \xi_{1t}' \\ \vdots \\ \xi_{n_t t}' \end{bmatrix}$$

then the error term in the productivity process can be written as:

$$\zeta_{it} = \Xi_t \boldsymbol{\delta} + \nu_t$$

where $\boldsymbol{\delta} = \Sigma_\xi^{-1} \sigma_{\zeta\xi}\iota$, ν_t is independent of ξ_{it} and $\sigma_\nu^2 = \sigma_\zeta^2 - \sigma_{\zeta\xi}\iota' \Sigma_\xi^{-1} \sigma_{\zeta\xi}\iota$. Therefore, the productivity process becomes:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda G_t \omega_t + \Xi_t \boldsymbol{\delta} + \nu_t \tag{31}$$

G_t is endogenous when $\sigma_{\zeta\xi} \neq 0$ and the selectivity bias is equal to $\Xi_t \boldsymbol{\delta}$.

5.2 Accounting for Selection

To the estimate model, assume ζ_{it} is normally distributed. Then Arduini et al. (2015) showed that the selectivity bias can be controlled for using a Heckman-style mills ratio:

$$\mu_{it} = \sum_{j \neq i}^{N_t} g_{ij,t} \frac{\phi(\Phi^{-1}(p))}{p} + (1 - g_{ij,t}) \frac{\phi(\Phi^{-1}(p))}{1-p} \tag{32}$$

where $p = P(A_{ij,t} = 1 \,|\, \mathcal{I}_{t-1})$, and ϕ and Φ are the probability and cumulative density functions for a standard normal variable. The i.i.d assumption on ξ_{ijt}'s dispenses with the need to estimate all $N_t - 1$ parameters in $\boldsymbol{\delta}$. Instead, due to the summation above, one only has to estimate a single parameter $\delta = \frac{\sigma_{\zeta\xi}}{\sigma_\xi^2}$.

5.3 Estimation Procedure

Incorporating the selection model is similar to the Olley and Pakes (1996) correction for attrition. The first stage of my benchmark procedure is unchanged with the estimation of $\widehat{\varphi}_{it}$ and $\widehat{\varepsilon}_{it}$ using the proxy variable. In the second stage, starting with the initial guesses of the labor and capital coefficients $(\alpha_k^*, \alpha_\ell^*)$, compute $\omega_{it-1}^* =$

$\widehat{\varphi}_{it-1} - \alpha_k^* k_{it-1} - \alpha_\ell^* \ell_{it-1}$.

Using ω_{it-1}^* and other variables that could determine the observed links between firms, estimate the selection model in equation (30) to obtain γ^*. Next, compute the predicted probabilities $p^* = \frac{e^{U_{ijt}(\gamma^*)}}{1+e^{U_{ijt}(\gamma^*)}}$ and the selection correction term $\mu_{it}^* = \sum_{j\neq i}^{N_t} g_{ij,t} \frac{\phi(\Phi^{-1}(p^*))}{p^*} + (1 - g_{ij,t}) \frac{\phi(\Phi^{-1}(p^*))}{1-p^*}$. Include this correction term as one of the explanatory variables in the productivity process equation:

$$\omega_t^* = \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}^*, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \delta \sum_{s=0}^{\infty} \lambda^s G_t^s \mu_t^* + u_t \tag{33}$$

The resulting residuals are now purged of the omitted variable bias arising from network selection and can be used to construct the sample moments in equation (28) for identification of the elasticities.[23]

6 Extensions

6.1 Gross Output Production Functions

So far, I have only considered a structural value-added production function, which often requires the assumption that the production function is Leontief with respect to intermediate inputs. In this section I consider a framework exploiting first order conditions on intermediate input choices as in Gandhi, Navarro, and Rivers (2020, GNR hereafter). Under similar assumptions as in the proxy variable approach above, the standard GNR procedure can be modified to jointly estimate network effects and productivity.

Like ACF, the GNR methodology assumes that TFP enters the production function in a Hicks-neutral fashion. However, intermediate inputs now enter directly into

[23] In principle, the selection model would be re-estimated for each value of ω_{it-1}^* as the values $(\alpha_k^*, \alpha_\ell^*)$ are updated in each iteration. However, this significantly increases the computational cost of the procedure. Provided the initial guesses of the elasticities, such as those obtained from an OLS regression, are reasonably close to their true values, measurement error in the lagged TFP variable should not have an outsized effect on the estimates of the selection correction term. In my Monte Carlo simulations, results were quite similar when selection was estimated only once and when it was re-estimated in each iteration.

the production function:

$$Y_t = F(L_t, K_t, M_t)e^{\omega_t+\varepsilon_t} \iff y_t = f(\ell_t, k_t, m_t) + \omega_t + \varepsilon_t \tag{34}$$

For simplicity, assume that materials are flexible while both labor and capital have dynamic implications.

The procedure consists of two stages. The first stage exploits first order conditions from profit maximization to estimate the elasticity of intermediate inputs with respect to output. Given the production technology above, the firm chooses materials to maximize profits:

$$\max_{M_t} P_t E[F(L_t, K_t, M_t)e^{\omega_t+\varepsilon_t}] - P_t^M M_t \tag{35}$$

where P_t and P_t^M are the prices of output and materials respectively. The static first order condition with respect to materials is:

$$P_t \frac{\partial}{\partial M_t} F(L_t, K_t, M_t)e^{\omega_t}\mathcal{E} = P_t^M \tag{36}$$

where $\mathcal{E} \equiv E[e^{\varepsilon_t} \mid \mathcal{I}_t] = E[e^{\varepsilon_t}]$ which relies on the assumption that the error terms are unconditionally independent.[24]

It is also pertinent to note that this first order condition makes an implicit assumption about market structure: that the firm is a price-taker in both input and output markets. Therefore, this framework cannot directly examine impacts of or effects on market power. I retain this assumption in my modified procedure.

$$\ln\left(\frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t)\right) - \varepsilon_t + \ln(\mathcal{E}) = s_t \tag{37}$$

where $s_t \equiv \ln(\frac{P_t^M M_t}{P_t Y_t})$ is the log of the intermediate input expenditure share of revenue.

$$E[\varepsilon_t \mid \mathcal{I}_t] = 0 \implies E[s_t \mid \mathcal{I}_t] = \ln\left(\frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t)\right) + \ln(\mathcal{E}) \tag{38}$$

Let $D^{\mathcal{E}}(\ell_t, k_t, m_t) \equiv \dfrac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) \times \mathcal{E}$. Then given the moment of ε_t in (38)

[24] See Gandhi et al. (2020) for details on estimation under a relaxed conditional independence assumption.

above, $\ln D^{\mathcal{E}}(\ell_t, k_t, m_t)$ can be estimated by non-linear least squares regression of the materials expenditure share on the log of a polynomial in labor, capital and materials. Furthermore:

$$\begin{aligned}\varepsilon_t =& \ln D^{\mathcal{E}}(\ell_t, k_t, m_t) - s_t \implies e^{\varepsilon_t} = D^{\mathcal{E}}(\ell_t, k_t, m_t)e^{-s_t} \\ \mathcal{E} =& E[e^{\varepsilon_t}] = E[D^{\mathcal{E}}(\ell_t, k_t, m_t)e^{-s_t}] \end{aligned} \tag{39}$$

Using the estimates of $D^{\mathcal{E}}$ from the share regression, we can replace the moment in (39) with its empirical equivalent and compute the constant $\mathcal{E}$. This enables us obtain an estimate of the materials elasticity:

$$D(\ell_t, k_t, m_t) = \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) = \frac{D^{\mathcal{E}}(\ell_t, k_t, m_t)}{\mathcal{E}} \tag{40}$$

The second stage of GNR relies further assumptions on the productivity process to estimate the rest of the production function. By the fundamental theorem of calculus:

$$\int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t = f(\ell_t, k_t, m_t) + \mathcal{C}(\ell_t, k_t) \tag{41}$$

The goal is to estimate $\mathcal{C}(\cdot)$ since we can compute $\int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t$ using $D(\ell_t, k_t, m_t)$ from the first stage. By substituting for $f(\ell_t, k_t, m_t)$ using equation (34):

$$\begin{aligned}\int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t &= y_t - \omega_t - \varepsilon_t + \mathcal{C}(\ell_t, k_t) \\ \mathcal{Y}_t \equiv y_t - \int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t - \varepsilon_t &= -\mathcal{C}(\ell_t, k_t) + \omega_t \end{aligned} \tag{42}$$

It is at this point that the assumption on the productivity evolution process comes into play. GNR maintains the same first-order Markov assumption as ACF:

$$\omega_t = h(\omega_{t-1}) + \eta_t, \qquad \text{where } E[\eta_t \mid \mathcal{I}_{t-1}] = 0 \tag{43}$$

$$\begin{aligned}\omega_{t-1} &= \mathcal{Y}_{t-1} + \mathcal{C}(\ell_{t-1}, k_{t-1}) \\ \implies \mathcal{Y}_t &= -\mathcal{C}(\ell_t, k_t) + h(\mathcal{Y}_{t-1} + \mathcal{C}(\ell_{t-1}, k_{t-1})) + \eta_t \end{aligned} \tag{44}$$

We can estimate $\mathcal{C}(\cdot)$ and $h(\cdot)$, normalizing the former to contain no constant, based on unconditional moments derived from $E[\eta_t \mid \mathcal{I}_t]$:

$$E[\eta_t \ell_t^{\tau_\ell} k_t^{\tau_k}] = 0 \;\text{ and }\; E[\eta_t \mathcal{Y}_{t-1}^{\tau_\mathcal{Y}}] = 0 \tag{45}$$

where τ_ℓ, τ_k and $\tau_\mathcal{Y}$ are determined by the degrees of the polynomial approximations for $\mathcal{C}(\cdot)$ and $h(\cdot)$ respectively.

6.1.1 Accounting for Network Effects

As with the modified ACF approach, I maintain the same assumptions and procedure in the first stage of GNR. Network effects come into play at the second stage when the law of motion on productivity is required for identification.

Note however, that by maintaining the same assumptions in the first stage, I do not account for ways in which the firm's network could potentially influence its intermediate input choices. For now, I focus specifically on network effects that operate through productivity spillovers and leave the implications for materials demand for future work.

I replace the productivity evolution process in (43) with one that allows for a linearly additive endogenous network effect:[25]

$$\begin{aligned} \omega_t &= h(\omega_{t-1}) + \lambda G_t \omega_t + \zeta_t \qquad \text{where } E[\zeta_t \mid \mathcal{I}_{t-1}] = 0 \\ \implies \omega_t &= \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t \end{aligned}$$

The equation (44) becomes:

$$\mathcal{Y}_t = -\mathcal{C}(\ell_t, k_t) + \sum_{s=0}^{\infty} \lambda^s G_t^s h(\mathcal{Y}_{t-1} + \mathcal{C}(\ell_{t-1}, k_{t-1})) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t \tag{46}$$

This yields an additional set of moments from which the endogenous effect λ can be identified:

$$E[\zeta_t G_t^s \mathcal{Y}_{t-1}^{\tau_\mathcal{Y}}] = 0 \qquad \text{where } s \geq 1 \tag{47}$$

[25] For clarity of exposition, I leave out contextual and correlated effects, but they can be included in much the same way as with ACF.

6.2 Alternative Network Effect Specifications

The modified ACF procedure introduced in section 4 can accommodate specifications of the productivity process that account for other ways in which spillovers may occur. In this section, I consider some of these specifications, and how they affect the estimator and what additional assumptions are needed, if any.

6.2.1 Local-Aggregate Endogenous Effect

The linear-in-means equation considered so far is also known as the local-average model because it assumes that the average productivity and characteristics of a firm's neighbors is the key source of spillovers. Another model is the local-aggregate model as in Liu and Lee (2010) and Liu et al. (2014), that considers the sum rather than the average. That is:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, A_t\boldsymbol{x_{t-1}}) + \lambda A_t \omega_t + \zeta_t \tag{48}$$

where A_t is the adjacency matrix. This model has different implications from the local-average model. There are also hybrid models that include local-average contextual effects and local-aggregate endogenous effects:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \lambda A_t \omega_t + \zeta_t \tag{49}$$

or both local-average and local-aggregate endogenous effects:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \lambda_A A_t \omega_t + \lambda_G G_t \omega_t + \zeta_t \tag{50}$$

See Liu and Lee (2010) and Liu et al. (2014) for further discussion of the conditions under which these network effects are identified. In general as long as the matrix inversion conditions to obtain a reduced form and the information set conditions hold, my benchmark procedure only needs to be modified by changing the network matrix where necessary.

6.2.2 Heterogeneous Network Effects

So far, my model of network effects has assumed homogeneous spillovers. However, the model can account for a finite set of heterogeneous network effects. If I partition the network into a finite set of B groups such as buyers and suppliers, industries, or based on firm size, then I can estimate:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, \{G_{b,t}\boldsymbol{x_{t-1}}\}_{b=1}^{B}) + \sum_{b=1}^{B} \lambda_b G_{b,t}\omega_t + \zeta_t \tag{51}$$

Note that $\boldsymbol{x_{t-1}}, \{G_{b,t}\boldsymbol{x_{t-1}}\}_{b=1}^{B}\omega_t = \lambda G_t \omega_t$ where λ is a weighted average of the heterogeneous effects. Therefore, my benchmark procedure can still be used to consistently estimate TFP without any modification. Afterwards, the heterogeneous network effect parameters can be obtained using the specification above. Dieye and Fortin (2017) and Patacchini et al. (2017) discuss the identification conditions and estimation procedures for this model in greater detail.

7 Results

In this section, I use my empirical framework to examine the magnitude of endogenous productivity spillovers through vertical relationships in the US production network. I explore how these spillovers vary over time, industry and firm size and document substantial heterogeneity in the sources and recipients of network effects.

I estimate a gross production function with a linear intermediate input share equation and a second-degree polynomial in capital and labor in the second stage.[26] I also estimate a value-added Cobb Douglas production function with materials as the proxy variable and a second-degree polynomial in the first stage. In both specifications, I assume a linear productivity process that includes an endogenous network effect and recover both production function elasticities and productivity spillovers from my modified approach. Because spillovers imply that TFP is jointly determined for linked firms across industries, the production function cannot be estimated separately for each industry. Therefore, I control for industry and year fixed effects in the productivity equation. In addition, due to the observed variation in the network structure over time, I estimate both specifications separately for each decade in the sample.

I compare my estimates with results from standard GNR and ACF approaches with industry and year fixed effects in the productivity equation for comparability.

[26]This specification implies a translog production function.

Because standard approaches do not yield estimates of productivity spillovers, I use TFP estimates from these procedures in a second stage. To obtain network effect coefficients, I apply the generalized 2SLS (G2SLS) approach in Lee (2003) and Bramoullé et al. (2009). In the first step, I estimate λ^* by 2SLS using $[G_t\omega_{t-1}, G_t^2\omega_{t-1}]$ as instruments for $G_t\omega_t$. I compute $E^*[G_t\omega_t|\mathcal{I}_{t-1}]$ using the reduced form equation in (13). This is the feasible estimate of the best instrumental variable (IV) for $G_t\omega_t$. Then I estimate 2SLS again, this time with $E^*[G_t\omega_t|\mathcal{I}_{t-1}]$ instrumenting for $G_t\omega_t$. To eliminate component-year fixed effects, I perform global differencing described in section 4.2 to both standard and modified procedures.

To control for network selection, I estimate the model in section 5, estimating the likelihood of a buyer-seller link as a function of the each firm's age, size, the absolute difference between these characteristics, past trading relationship, distance between their headquarters, and the compatibility of their industries as measured by inter-industry input and output shares. Given the sparsity of the network, I restrict a firm's choice set of buyers or sellers to those it could reasonably expect to trade with given industry-specific production technologies; that is, rather than include all dyads, I restrict estimation to dyads with non-zero entries from input-output tables for their respective industries. The resulting error terms are combined as described in section 5 and included in the productivity process.

As discussed in section 2, the buyer-supplier network is only partially observed because firms only need to report their major customers; only about 18% of links fall below the 10% sales threshold. To address this, I rely on information about link intensity: I weight each relationship by the value traded between the two firms in that year. This mitigates some of the bias from missing links, because links that fall below the 10% threshold are would have weights close to zero. There is also the added advantage of allowing more important trading partners to have a larger impact on a firm's productivity.[27]

7.1 Production Function Elasticities

Table 2 reports the average estimated elasticities of output with respect to inputs from gross output and value-added production functions, respectively. GNR and ACF refer

[27] As a robustness check, I estimate all specifications with an unweighted network in section B.1 of the appendix. The results are similar in magnitude, indicating that major trading partners are the more salient sources of spillovers.

to the standard procedures, GNR-N and ACF-N denote my modified approach that accounts only for endogenous productivity spillovers, GNR-ND and ACF-ND indicate specifications with both endogenous network effects and component-year fixed effects, and GNR-NDS and ACF-NDS include the correction for network selection in addition to all previous network effects. Because I assume that the network does not directly affect intermediate input demand in the gross output specification, the elasticity of output with respect to materials does not vary across specifications.

The input elasticities do not vary much with the inclusion of network effects, although the capital coefficient is slightly higher in with estimates obtained network-augmented GNR relative to standard GNR, and lower with the network-augmented ACF relative to standard ACF.[28]

Table 2: Gross Output Production Function Elasticities

Function Type	Estimator	Capital	Labor	Materials
Gross Output	GNR	0.210	0.321	0.500
	GNR-N	0.219	0.318	0.500
	GNR-ND	0.221	0.317	0.500
	GNR-NDS	0.218	0.318	0.500
Value Added	ACF	0.384	0.644	-
	ACF-N	0.383	0.646	-
	ACF-ND	0.375	0.657	-
	ACF-NDS	0.375	0.657	-

Average elasticities of output with respect to inputs from a gross output and value-added production function estimated on a subsample US firms in Compustat. All specifications include industry and year fixed effects in the productivity process.

7.2 Endogenous Productivity Spillovers

I now turn to estimates of productivity spillovers. First, I define the network as undirected: a firm i is affected by the average productivity of all identified buyers and sellers. Table 3 shows the endogenous network effects.

In the gross output specification, TFP measures from the standard GNR approach suggest that a firm's productivity rises by 0.0084 percent in the short run when its

[28]The latter pattern is consistent with the Monte Carlo simulations in table 2, as is the similarity across estimates given the sparsity of the observed network.

Table 3: Endogenous Productivity Spillovers

	Dependent Variable: $\ln\text{TFP}_t$						
	Gross Output				Value-added		
Estimator	$\ln\text{TFP}_{t-1}$	Neighbors' $\ln\text{TFP}_t$	Network Selection	Estimator	$\ln\text{TFP}_{t-1}$	Neighbors' $\ln\text{TFP}_t$	Network Selection
GNR	0.9035	0.0084	-	ACF	0.8687	0.007	-
	(0.0067)	(0.0023)	-		(0.0095)	(0.0026)	-
GNR-N	0.9025	0.009	-	ACF-N	0.8688	0.0064	-
	(0.0067)	(0.0023)	-		(0.0095)	(0.0026)	-
GNR-ND	0.8996	0.0076	-	ACF-ND	0.8663	0.001	-
	(0.0073)	(0.0024)	-		(0.0101)	(0.0026)	-
GNR-NDS	0.8997	0.0074	-0.0009	ACF-NDS	0.8672	0.0011	-0.00002
	(0.0073)	(0.0024)	(0.0005)		(0.0101)	(0.0026)	0.00045

Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 4: Network Formation Model Estimates

Dependent variable: Firm i sells to firm j	Coefficient	S.E.		Coefficient	S.E.
Age_{it}	-0.0177	0.001	Excluded: Distance in (100, 500) mi		
Size_{it}	-0.4094	0.019	Distance < 25 mi	0.6247	0.049
Age_{jt}	0.0052	0.001	Distance in (25, 100) mi	0.2405	0.030
Size_{jt}	0.7468	0.022	Distance in (500, 1000) mi	-0.1679	0.025
$\lvert\text{Age}_{it} - \text{Age}_{jt}\rvert$	0.0023	0.001	Distance in (1000, 1500) mi	-0.3124	0.031
$\lvert\text{Size}_{it} - \text{Size}_{jt}\rvert$	-0.3231	0.020	Distance in (1500, 2000) mi	-0.4492	0.037
Firm i sold to j in $t-1$	9.0060	0.084	Distance in (2000, 2500) mi	-0.4562	0.032
Share of industry_i's output sold to industry_j	1.1988	0.097	Distance > 2500 mi	-0.6308	0.030
Share of industry_j's inputs purchased from industry_i	2.5486	0.086	Distance measure does not exist	-0.7337	0.069

S.E. is for standard errors.

average buyer or seller gets 1 percent more productive. The persistence of TFP over time implies a magnified long-run effect of having more trading partners in one period: an estimated coefficient of 0.9 on ln TFP_{t-1} implies a 1 percent more efficient average trading partner in a single period results in a long-run efficiency gain of 0.084 percent.

Accounting for endogenous productivity spillovers in TFP estimation raises the long-run effect to 0.09 percent, but differencing out common shocks to productivity lowers the estimate to 0.076 percent. In the value-added specification, the impact of correlated effects is striking; the estimated long-run impact of a 1 percent rise in the average trading partner's TFP goes from 0.07 percent with standard ACF to 0.01 when I account for both endogenous and correlated effects in the production function estimation.

Across gross output and value-added specifications, controlling for network selection does not significantly change the estimates compared to just network differencing, and there does appear to be a negative but negligible relationship between productivity shocks and relationship formation. Estimates from the network formation model in table 4 suggest positive assortativity in size and negative assortavity in age between sellers and buyers. Consistent with gravity models of trade, distance is positively correlated with the likelihood of a trading relationship, and given that the reported links are to major customers, a past trading relationship appears to be strongly predictive of a current linkage, indicating stability of links over time. The persistence of observed links and the limited impact of the network formation model on the endogenous spillover estimates suggests that conditional on past productivity, firms do not appear to positively sort in response to unanticipated productivity shocks. Therefore, for the kinds of relationships that we observe, these contemporaneous network effects are not attributable to sorting.

Across all specifications, estimates from the standard approach and my modified procedure are often statistically indistinguishable, which is to be expected given the sparsity of the network as discussed in section OA1.

7.2.1 Relationship Direction

Using the approach outlined in section 6.2.2, I allow spillovers to vary by the direction of the relationship i.e. from suppliers to buyers and vice versa. Table 5 shows that productive suppliers have more than four times the impact on their customers as productive buyers have on their suppliers in the gross output specification:

having a 1 percent more productive supplier is associated with 0.083 percent higher productivity in the long-run, whereas an equivalent rise in the average customer's productivity would generate 0.018 percent long-run increase. The same pattern holds in the value-added specification but the difference is less pronounced and not statistically significant: a 0.016 percent long-run supplier effect compared to a 0.011 percent customer effect.

Table 5: Productivity Spillovers by Relationship Direction

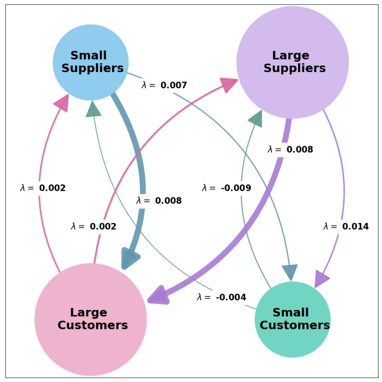
	Dependent Variable: $\ln TFP_t$				
	Gross Output			Value-added	
Estimator	Customers' $\ln TFP_t$	Suppliers' $\ln TFP_t$	Estimator	Customers' $\ln TFP_t$	Suppliers' $\ln TFP_t$
GNR	0.0026 (0.0008)	0.0053 (0.0009)	ACF	0.0013 (0.0003)	0.0016 (0.0004)
GNR-N	0.0032 (0.001)	0.0102 (0.0013)	ACF-N	0.0012 (0.0003)	0.0018 (0.0004)
GNR-ND	0.002 (0.0009)	0.0095 (0.0012)	ACF-ND	0.0011 (0.0003)	0.0016 (0.0003)
GNR-NDS	0.0018 (0.0008)	0.0083 (0.001)	ACF-NDS	0.0011 (0.0003)	0.0016 (0.0003)

Standard errors are in parentheses. All specifications include industry and year fixed effects.

Estimates from value-added specifications do not show a significant difference between spillovers from buyers to sellers or vice versa: a 10 percent more efficient supplier is associated with a 0.016 percent rise in productivity while the effect of buyers is 0.011 percent. These effects are statistically indistinguishable from each other. For the rest of this discussion, I focus on estimates from gross output production functions, but additional results from value-added specifications are in section B.2 in the appendix.

Larger supplier spillovers could be to the fact that supplier efficiencies can be passed downstream more passively than customer efficiencies flowing upstream. Improved logistics and customer relationship management (CRM) practices on the part of the supplier could generate operational gains for buyers by enabling them to better streamline production or eliminate bottlenecks. To the extent that suppliers cannot fully price these changes into their long-standing contracts, these would show up as

Figure 8: Spillovers by Firm Size and Relationship Direction



Arrows indicate direction of spillovers. See table 6 for standard errors.

productivity spillovers even in the absence of any direct collaboration. By contrast, spillovers from customers to suppliers often necessitate some form of information or knowledge sharing because the customer's production is not a direct input into the supplier's production. Therefore, customers may be affected by both large and small suppliers but suppliers are likely to be mainly affected by large customers who are more likely to exert a meaningful influence on the supplier's operations.

An alternative explanation could be that these differences arise from the estimation of revenue-based TFP (TFPR) rather than quantity-based TFP (TFPQ). If large customers exert downward pressure on their suppliers' prices, this would have a dampening effect on the suppliers' TFPR even if TFPQ is rising.

Table 6: Productivity Spillovers by Varying Firm Size Cutoffs (Gross Output)

		Dependent Variable: $\ln TFP_t$ Size Cutoff							
		500		1000		5000		Median	
Partner Size	Relationship	Large Firm	Small Firm	Large Firm	Small Firm	Large Firm	Small Firm	Large Firm	Small Firm
Large	Customers	0.002	0.0021	0.0017	0.0036	0.0015	0.0019	0.0011	0.0015
		(0.0008)	(0.001)	(0.0008)	(0.0011)	(0.0008)	(0.001)	(0.0008)	(0.001)
	Suppliers	0.0083	0.0143	0.008	0.0094	0.006	0.0086	0.0088	0.0093
		(0.001)	(0.0077)	(0.001)	(0.0036)	(0.0008)	(0.0027)	(0.001)	(0.0016)
Small	Customers	-0.0091	-0.0045	-0.001	-0.0051	-0.0038	0.0007	0.0019	0.0026
		(0.0076)	(0.0052)	(0.0061)	(0.0036)	(0.0039)	(0.0012)	(0.0015)	(0.0012)
	Suppliers	0.0081	0.0075	0.0086	0.0074	0.008	0.0091	0.0074	0.0088
		(0.0009)	(0.0053)	(0.001)	(0.0033)	(0.0012)	(0.0013)	(0.0011)	(0.0013)

Large firms are defined by having at least as many employees as the cutoffs indicated above. The median cutoff is determined by industry and year. Estimates are from a gross output production function with endogenous and correlated effects. Standard errors are in parentheses. All specifications include industry and year fixed effects.

To distinguish between these two explanations, I estimate productivity spillovers that vary by firm size. I classify firms according to the BEA's definition: a firm is large if it has 500 or more employees. If the difference between customer and supplier spillovers is driven by the first explanation, then we should expect small customers to have a negligible effect on their suppliers compared to large customers. On the other hand, price-driven explanation for the spillover differential would imply that spillovers from customers should be greater for large firms compared to small firms, assuming size is a good proxy for relative bargaining power. The results reported in table 6 and depicted in figure 8 are more consistent with the former explanation than with the latter. Large customers have a similar impact on large and small firms, small customers have a negative and not statistically significant effect on firms of both sizes. Supplier-to-buyer spillovers, however, are driven by both small and large suppliers with equally-sized effects on large firms.

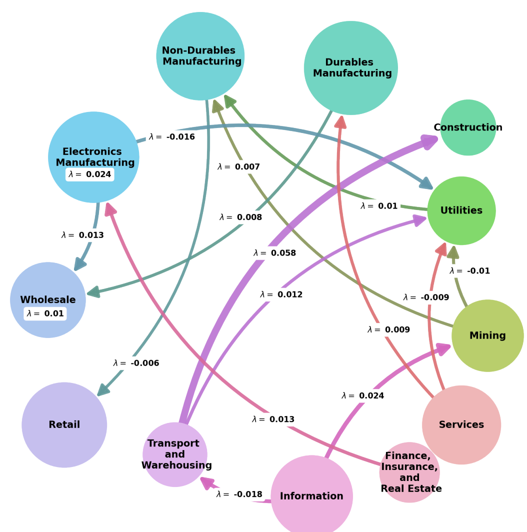
Given that the average firm in my sample is larger than the average firm in the US, at least 60 percent or more the sample can be classified as large based on this definition. In table 6, I check how sensitive these results are to different classifications of firm size. I consider three alternative definitions based on the number of employees: greater than or equal to 1000, 5000 or an industry-year specific median. The results are similar across definitions except that, as expected, the impact of large firms diminishes while that of small firms rises as the threshold is raised.

7.2.2 Inter- and Intra-sectoral spillovers

In this section, I investigate the transmission of productivity gains within and across sectors. I estimate a gross output production function with endogenous and correlated effects, allowing spillovers to vary by the sector of the firm and its trading partners. For ease of illustration, I only depict spillovers that are significant at the 5 percent level, separating forward spillovers in figure 9 and backward spillovers in figure 10. To do so, I classify inter-sectoral spillovers as based on the share of observed links that are from sellers to buyers, or from buyers to sellers respectively. If 50 percent or more of links between sector u and sector v are from suppliers in u to customers in v, then the spillovers from u to v are classified as forward or upstream to downstream, while the impact of sector v on firms in u is considered backward or downstream to upstream. Table 7 reports the full set of estimates.

These results highlight the important role of electronics manufacturers, information technology (IT) firms, and retailers and wholesalers in US productivity growth.

Figure 9: Upstream to Downstream Productivity Spillovers by Sector
Arrows indicate the direction of productivity spillovers (λ). See table 7 for the full set of coefficients.



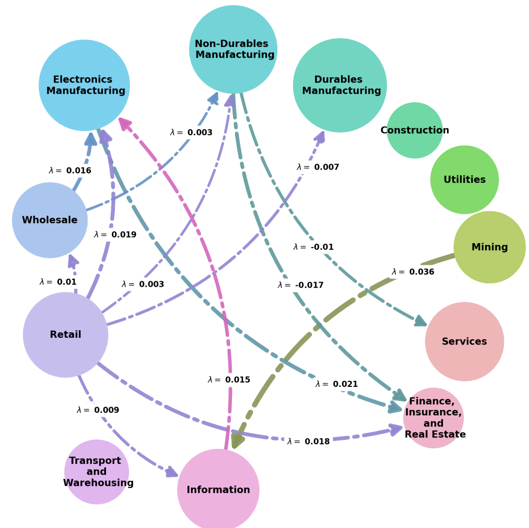


Figure 10: Downstream to Upstream Productivity Spillovers by Sector
Arrows indicate the direction of productivity spillovers (λ). See table 7 for the full set of coefficients.

Table 7: Productivity Spillovers by Sector (Gross Output)

	Dependent Variable: $\ln\text{TFP}_t$ Firm's Sector u											
Partners' Sector v	Mining	Utilities	Constr	Durables	Non-Durables	Electronics	Wholesale	Retail	Trans & WH	Info	FIRE	Services
Mining	-0.0084	-0.0096	-0.0064	0.0017	0.0075	-0.0161	0.0172	0.0094	0.0102	0.0357	0.0042	0.0028
	(0.0064)	(0.0038)	(0.0331)	(0.005)	(0.0027)	(0.0326)	(0.0104)	(0.008)	(0.0092)	(0.0111)	(0.0085)	(0.008)
Utilities	-0.0032	-0.0007	0.0061	0.003	0.01	0.0134	0.0041	-0.0028	0.0018	-0.0024	-0.0038	0.0105
	(0.0059)	(0.0028)	(0.0104)	(0.0027)	(0.0021)	(0.0082)	(0.0083)	(0.0084)	(0.0065)	(0.0125)	(0.0178)	(0.006)
Construction	-0.0192	0.0014	0.0147	0.0055	-0.0058	-0.0079	0.0113	-0.0075	0.017	0.0043	-0.0005	0.0282
	(0.0312)	(0.0041)	(0.0115)	(0.0059)	(0.0036)	(0.0133)	(0.0076)	(0.009)	(0.0228)	(0.0035)	(0.0061)	(0.0218)
Durables Mfg	0.0077	0.0002	0.0265	0.0025	0.0012	-0.0036	0.0079	-0.0016	0.0029	0.0039	0.0052	0.0032
	(0.0065)	(0.003)	(0.0231)	(0.0022)	(0.0016)	(0.0024)	(0.0027)	(0.0016)	(0.0035)	(0.0034)	(0.0063)	(0.0034)
Non-Durables Mfg	-0.0061	-0.0014	-0.0028	0.0001	-0.001	-0.0067	0.0034	-0.0062	-0.0065	-0.0026	-0.0175	-0.0102
	(0.0052)	(0.0028)	(0.0132)	(0.0018)	(0.0013)	(0.0048)	(0.0027)	(0.0016)	(0.0038)	(0.0044)	(0.0085)	(0.0038)
Electronics Mfg	-0.0433	-0.0157	-0.0275	-0.0003	-0.0043	0.0244	0.013	0.0021	0.0079	0.0007	0.0207	0.0001
	(0.0501)	(0.0055)	(0.0355)	(0.0046)	(0.0038)	(0.0029)	(0.0031)	(0.0032)	(0.0058)	(0.0031)	(0.0091)	(0.0043)
Wholesale	-0.0052	0.0126	0.0121	0.0022	0.0032	0.0165	0.01	0.0017	0.0033	0.0022	-0.0132	0.0027
	(0.0118)	(0.0075)	(0.0164)	(0.0021)	(0.0012)	(0.0021)	(0.0038)	(0.0014)	(0.0104)	(0.0024)	(0.0118)	(0.004)
Retail	0.0113	-0.0037	-0.0075	0.0074	0.003	0.0194	0.0104	0.0012	-0.0013	0.009	0.018	0.0017
	(0.0118)	(0.0104)	(0.016)	(0.0028)	(0.0013)	(0.0036)	(0.0019)	(0.0026)	(0.004)	(0.0041)	(0.0046)	(0.0044)
Transport and Warehousing	0.0079	0.0116	0.0575	0.0037	0.0033	0.0067	0.0078	0.0022	-0.0031	0.0007	-0.0088	0.0
	(0.0094)	(0.0034)	(0.0184)	(0.0041)	(0.002)	(0.0059)	(0.0109)	(0.0039)	(0.0042)	(0.0081)	(0.0073)	(0.0103)
Information	0.0244	-0.0023	0.0085	-0.0035	-0.001	0.0151	0.0064	-0.001	-0.018	0.0041	0.0066	-0.0074
	(0.0118)	(0.0069)	(0.017)	(0.0039)	(0.0028)	(0.0025)	(0.0049)	(0.0026)	(0.0048)	(0.0028)	(0.0047)	(0.004)
Finance, Insur & Real Estate	0.0004	0.005	-0.0134	0.0	-0.0044	0.0134	-0.0026	0.0017	0.0045	0.0042	-0.0021	0.0048
	(0.0112)	(0.0143)	(0.0128)	(0.0044)	(0.0025)	(0.0034)	(0.0149)	(0.002)	(0.0043)	(0.0031)	(0.0043)	(0.0036)
Services	0.0035	-0.0088	0.0083	0.0088	0.0042	0.0022	0.0084	0.0042	0.0025	0.0052	0.009	0.0025
	(0.0117)	(0.0036)	(0.0253)	(0.0023)	(0.0023)	(0.0028)	(0.0072)	(0.0026)	(0.005)	(0.0031)	(0.0048)	(0.0037)

Sectors are determined according to the BEA industry classification.Estimates are from a gross-output production function with endogenous and correlated effects. Standard errors are in parentheses. All specifications include industry and year fixed effects.

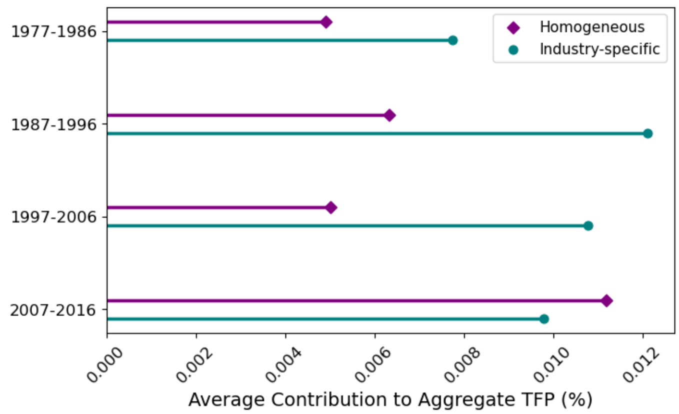
Substantial forward spillovers occur within the electronics manufacturing sector. Furthermore, the synergies between electronics manufacturing and the finance, insurance and real estate sector is primarily driven by technology patent holders (SIC 6794 and NAICS 533110) such as InterDigital Inc. which provides mobile technology research services to mobile phone manufacturers such as Apple. Manufacturers also tend to amplify the impact of productivity growth in other sectors because they enjoy efficiency boosts from both directions: electronics manufacturers from mainly customers and manufacturers of non-durables from their suppliers. Although retailers and transport and warehousing firms generate substantial positive backward spillovers, they do not benefit from efficiency boosts from other sectors, and in fact, experience negative network effects which suggests free-riding on efficiency improvements by trading partners.

The variation in backward and forward spillovers within and across sectors suggests that aggregate productivity growth can shaped by the sectoral composition of the country's production network. To examine this, I compute a back-of-the-envelope estimate of the long-run impact of a 1 percent rise in the productivity the most 10 central firms in each decade as shown in figure 5. Let J_t be the set of 10 most central firms at time t. Then for each firm $j \in J_t$, I sum j's contribution to the network average for each of connected firm i, weighting by firm i's share of revenues in that period, and multiply that by the spillover estimate. That is: $\text{Contribution}_{jt} = \widehat{\lambda} \sum_{it} \frac{Revenue_{it}}{Total\ Revenue_t} G_{ijt}$.

Then I compute $\sum_{j \in J_t} \text{Contribution}_{jt}$, divide by $(1 - \widehat{\rho})$ to get the long-run effect, and average by decade under two scenarios: $\widehat{\lambda} = 0.0074$, constant for all firms, and $\widehat{\lambda}$ varies by sector of the firms and their trading partners as in table 7. By comparing both estimates within each decade, we can understand how differences in sectoral composition change spillovers in the aggregate.

Figure 11 shows that the sectoral composition does make a significant difference in the magnitude of aggregate spillovers. In the earlier decades in which manufacturing firms were more central, sector-specific spillovers often twice as large as under the homogeneous spillover scenario. For example, in the 1987-1996 period, a 1 percent rise in the productivity of the 10 most central firms would contribute an estimated 0.006 percent to long term aggregate TFP growth assuming equal spillovers across firms, compared to 0.012 percent after taking into account sector-specific spillovers. By contrast, in the more recent decade with retailers and wholesalers most central in

Figure 11: Estimated Impact of 10 Most Central Firms on Aggregate TFP



Estimated long-run impact of a 1% increase in the TFP of the 10 most central firms in each year on sales-weighted aggregate productivity through spillovers alone.

Table 8: Spillovers and Intersectoral Characteristics

Dependent Variable: Estimated Spillovers from Partners' Sector v to Firm's Sector u ($\widehat{\lambda}$)	
Characteristic	Coefficient
Log(1 + Backward patent citations from u to v)	0.0019
	(0.0008)
Log(Employee flows from v to u)	-0.0033
	(0.0014)

Includes Partners' Sector (v) and Firm's Sector (u) fixed effects. Patent citations and employee flows are quarterly averages from 2011 to 2016. Standard errors in parentheses.

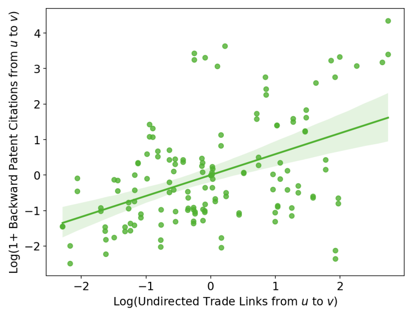
the network, overall spillovers are lower after taking into account sectoral variation. This lends some credence to idea that policymakers may prefer an economy in which manufacturing is central relative to other sectors.

Next, I examine two possible sources of the sectoral variation on productivity spillovers. The first is knowledge transmission as reflected in patent citations. As Fadeev (2023) points out, patent citations tend to be concentrated among business partners, suggesting the sharing of trade secrets across buyer-supplier links. Using the USPTO-Compustat crosswalk from the Wharton Research Data Service, I match US patents issued between 2011 to 2016 to firms in my sample, and obtain quarterly averages of sector-by-sector backward patent citations (including citations of patents issued prior to 2011). As shown in figure 12, the number of inter- and intra-sectoral buyer-seller links is positively correlated with backward citations.

The second correlate I consider is employee flows. Prior work has provided evi-

Figure 12: Trade Linkages and Patent Citations

Sector-level correlations between patent citations and buyer-supplier links controlling for citing and cited sector fixed effects. Patent citations are quarterly averages from 2011 to 2016.



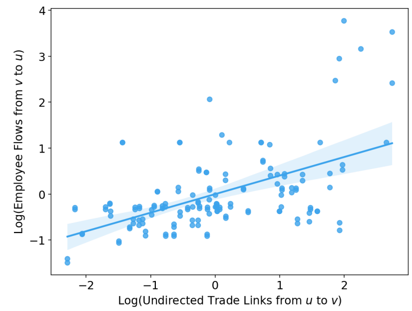


Figure 13: Trade Linkages and Employee Flows

Sector-level correlations between employee flows and buyer-supplier links controlling for employee origin and destination sector fixed effects. Employee flows are quarterly averages from 2011 to 2016.

dence that workers tend to move across firms that have trading relationships (Cardoza et al., 2023) and that knowledge transmitted through employee flows is a mechanism for inter-firm productivity spillovers (Stoyanov and Zubanov, 2012). I obtain national quarterly averages of sector-by-sector job-to-job flows from the Longitudinal Employer-Household Dynamics published by the US Census Bureau from 2011-2016. In figure 13, I show that, conditional on employee origin and destination sector fixed effects, there is a positive relationship between employee flows and trade relationships.

I explore how these two measures relate to the estimated intra- and inter-sectoral spillovers. As shown in table 8, estimated spillovers from sector v to u were higher if firms in sector u tend to cite patents from v, but lower if workers from v to tend to move u. Although these associations are by no means causal, the difference in directions is striking, suggesting that the estimated spillovers may reflect the kinds of efficiency gains through innovation than through worker knowledge. Further investigation of these differences is left for future work.

8 Concluding Remarks

The findings of this paper demonstrate the importance of considering firm interdependence through networks on the measurement of productivity and spillovers across firms. By incorporating network effects directly into the production function estimator, my modified approach can allow for the recovery of consistent estimates of productivity and contemporaneous spillovers in a unified approach. Although I have focused primarily on vertical relationships, the framework can be extended to other settings in which one may be interested in productivity spillovers over other types of linkages such as interlocking boards, ownership networks, geographic proximity etc. The key consideration is to understand in which settings the identifying assumptions such as network intransitivity hold and to adjust the framework accordingly.

Empirically, my estimates provide suggestive evidence of the importance of productivity spillovers across US publicly-listed companies through domestic firm-to-firm trading relationships. One limitation of this study is the relatively large size of the firms in this sample compared to the median US firm. Given that the backward spillovers are mainly driven by large customers, and both small suppliers and customers benefit from these spillovers even within this sample, we might expect larger estimates on a more representative sample that includes much smaller, younger, privately-held firms.

Finally, the difference in sector-specific spillovers raises questions about the role of upstreamness and downstreamness in diffusing productivity gains through the economy. Manufacturing firms that tend to both generate and benefit from spillovers from firms above and below them in the supply chain appear to amplify productivity gains more than downstream retailers which are an important source but not beneficiary. Unpacking the precise mechanism of these transmissions, particularly as it relates to innovation or tacit knowledge embodied in human capital is an important area of exploration for both decision-makers within firms and policy makers.

A Variable Construction

- Sales: Net sales deflated by an industry deflator for GDP.
- Labor: Number of employees
- Capital: Total property, plant and equipment (gross) before depreciation. Following the method in İmrohoroğlu and Tüzel (2014), I deflate using the yearly implicit price deflator for fixed investment at the calculated age of capital. Capital age is computed as the ratio of accumulated depreciation to current depreciation, smoothed by taking a 3-year moving average. The year at which the deflator is applied is current year − average capital age. All years before 1929 are bottom-coded because that is the earliest year in the deflator data.
- Materials: Estimated as Cost of goods sold plus Selling, General and Administrative Expenses minus labor costs. Salaries and wage costs are missing for most firms, so I estimate labor costs by multiplying number of employees by 2-digit industry wages per full-time equivalent employee. These estimates strongly correlate with wage costs that were reported in the data. Estimated materials are deflated by the 2-digit industry price indices for intermediate inputs.
- Value-added: Sales minus materials, deflated by industry price indices for value-added.
- Industry: Industry classifications are based on those used in input-output tables from the Bureau of Economic Analysis (BEA). There are 65 industries from before 1997 and 71 industries from 1997 onwards. These roughly correspond to 3-digit NAICS and 2-digit SIC codes. *Compustat*'s annual financials only reports the latest industry classification, therefore, I obtain historical NAICS codes from the primary business segment. I also replace SIC codes for companies that are incorrectly coded as "99" (unclassifiable) from annual reports in the EDGAR database and business segment data. These are then converted to BEA industry codes using the concordances provided by the bureau. All deflators, price indices and input-output tables are based on these BEA industry codes. However, in regressions I combine industries with too few observations. These include: transit and ground transportation with general transportation and warehousing, and other transportation and support activities; Funds, trusts

and other financial vehicles combined with securities, commodity contracts and investments; Legal services with miscellaneous professional services; Ambulatory health, hospitals, nursing and residential care with social assistance. This results in 41 industry groups.

B Additional Results and Robustness Checks

B.1 Unweighted Estimates

Table 9: Endogenous Productivity Spillovers (Unweighted)

	Dependent Variable: $\ln\text{TFP}_t$						
	Gross Output				Value-added		
Estimator	$\ln\text{TFP}_{t-1}$	Neighbors' $\ln\text{TFP}_t$	Network Selection	Estimator	$\ln\text{TFP}_{t-1}$	Neighbors' $\ln\text{TFP}_t$	Network Selection
GNR	0.9038 (0.0066)	0.0076 (0.0023)	- -	ACF	0.8688 (0.0095)	0.0061 (0.0026)	- -
GNR-N	0.9027 (0.0067)	0.0086 (0.0024)	- -	ACF-N	0.8689 (0.0094)	0.0056 (0.0026)	- -
GNR-ND	0.8998 (0.0073)	0.0069 (0.0025)	- -	ACF-ND	0.8662 (0.0101)	0.0002 (0.0028)	- -
GNR-NDS	0.9003 (0.0073)	0.0070 (0.0025)	-0.0009 (0.0005)	ACF-NDS	0.8671 (0.0101)	0.0002 (0.0028)	-0.00002 (0.00045)

Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 10: Productivity Spillovers by Relationship Direction (Unweighted)

	Dependent Variable: $\ln\text{TFP}_t$				
	Gross Output			Value-added	
Estimator	Customers' $\ln\text{TFP}_t$	Suppliers' $\ln\text{TFP}_t$	Estimator	Customers' $\ln\text{TFP}_t$	Suppliers' $\ln\text{TFP}_t$
GNR	0.0025 (0.0008)	0.0053 (0.0009)	ACF	0.0013 (0.0003)	0.0016 (0.0004)
GNR-N	0.0024 (0.0008)	0.0079 (0.0011)	ACF-N	0.0012 (0.0003)	0.0018 (0.0004)
GNR-ND	0.0015 (0.0008)	0.008 (0.001)	ACF-ND	0.0011 (0.0003)	0.0016 (0.0003)
GNR-NDS	0.0016 (0.0008)	0.008 (0.001)	ACF-NDS	0.0011 (0.0003)	0.0015 (0.0003)

Standard errors are in parentheses. All specifications include industry and year fixed effects.

B.2 Value-Added Estimates

Table 11: Productivity Spillovers by Varying Firm Size Cutoffs (Value-Added)

		Dependent Variable: $\ln TFP_t$ Size Cutoff							
		500		1000		5000		Median	
Partner Size	Relationship	Large Firm	Small Firm	Large Firm	Small Firm	Large Firm	Small Firm	Large Firm	Small Firm
Large	Customers	0.0008	0.002	0.0008	0.003	-0.0	0.0014	0.0002	0.0022
		(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
	Suppliers	0.0004	0.0028	0.0002	0.0049	-0.0006	0.0037	-0.0	0.0048
		(0.0002)	(0.0027)	(0.0002)	(0.0015)	(0.0002)	(0.001)	(0.0002)	(0.0006)
Small	Customers	-0.0054	-0.0003	-0.002	-0.0014	-0.0009	0.0004	-0.0003	0.0019
		(0.0023)	(0.0028)	(0.0015)	(0.0018)	(0.0009)	(0.0006)	(0.0005)	(0.0007)
	Suppliers	0.0007	0.0039	0.0006	0.0038	-0.0001	0.0039	0.0	0.0032
		(0.0002)	(0.0025)	(0.0002)	(0.0015)	(0.0003)	(0.0006)	(0.0003)	(0.0006)

Large firms are defined by having at least as many employees as the cutoffs indicated above. The median cutoff is determined by industry and year. Estimates are from a value-added production function with endogenous and correlated effects. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 12: Productivity Spillovers by Sector (Value-Added)

	Dependent Variable: $\ln TFP_t$ Firm's Sector											
Partners' Sector	Mining	Utilities	Constr	Durables	Non-Durables	Electronics	Wholesale	Retail	Trans & WH	Info	FIRE	Services
Mining	0.0037	-0.0004	0.009	0.0027	0.0014	-0.0057	-0.0011	-0.0026	0.0121	0.0135	-0.001	-0.0035
	(0.0018)	(0.0011)	(0.0092)	(0.0013)	(0.0009)	(0.0081)	(0.0034)	(0.0025)	(0.0055)	(0.0068)	(0.0036)	(0.0025)
Utilities	0.0029	0.0025	0.0034	-0.0001	0.0022	0.0025	0.0052	-0.0023	0.0024	-0.0013	-0.003	0.0033
	(0.0018)	(0.001)	(0.004)	(0.001)	(0.0011)	(0.0021)	(0.0033)	(0.0028)	(0.0021)	(0.0054)	(0.0107)	(0.0034)
Construction	0.0151	0.0014	0.009	0.0013	-0.0014	-0.0018	-0.0107	-0.0057	0.0047	0.0004	0.001	0.0096
	(0.0172)	(0.0013)	(0.0049)	(0.0023)	(0.0009)	(0.0032)	(0.0068)	(0.0037)	(0.0091)	(0.001)	(0.0033)	(0.0107)
Durables Mfg	0.0008	-0.0007	0.0171	-0.0002	-0.0019	-0.0017	0.0013	-0.0014	-0.0002	-0.0011	0.0027	0.003
	(0.0018)	(0.0008)	(0.0084)	(0.0006)	(0.0006)	(0.0006)	(0.0012)	(0.0007)	(0.0013)	(0.0013)	(0.0035)	(0.0014)
Non-Durables Mfg	0.0023	0.0007	0.0048	-0.0011	-0.0011	-0.0006	0.003	-0.0043	0.0035	-0.0012	-0.0073	-0.0059
	(0.0019)	(0.0011)	(0.0061)	(0.0006)	(0.0005)	(0.0011)	(0.0012)	(0.0007)	(0.0019)	(0.0016)	(0.0035)	(0.0017)
Electronics Mfg	-0.0087	-0.0011	-0.008	0.0004	-0.0014	0.0012	0.0042	-0.0013	0.0008	-0.0006	0.0049	-0.0
	(0.0081)	(0.0014)	(0.0084)	(0.0006)	(0.0012)	(0.0006)	(0.0011)	(0.0013)	(0.0013)	(0.0008)	(0.004)	(0.0014)
Wholesale	0.0007	-0.0005	0.0214	0.0007	0.001	0.0024	0.0061	-0.002	-0.0022	-0.0004	-0.0068	-0.0012
	(0.0042)	(0.0013)	(0.0216)	(0.0007)	(0.0006)	(0.0008)	(0.0025)	(0.0009)	(0.0026)	(0.0012)	(0.0058)	(0.0015)
Retail	-0.0052	-0.0082	0.0023	0.0018	0.0006	0.005	0.0033	0.0001	-0.0012	0.0044	0.0086	0.0002
	(0.0039)	(0.0032)	(0.0169)	(0.0007)	(0.0006)	(0.0018)	(0.0016)	(0.0012)	(0.0013)	(0.0022)	(0.0032)	(0.0022)
Transport and Warehousing	0.0031	-0.0003	0.013	0.0006	0.0005	-0.0016	0.0162	0.0025	0.0004	-0.0	0.0161	-0.0034
	(0.0029)	(0.0008)	(0.0084)	(0.0013)	(0.0006)	(0.0029)	(0.0059)	(0.0019)	(0.0016)	(0.0019)	(0.0044)	(0.0024)
Information	0.0071	0.0002	0.0036	-0.0026	-0.0017	0.0011	0.0055	-0.0003	-0.0042	0.002	0.0029	-0.0022
	(0.0052)	(0.0013)	(0.0064)	(0.0012)	(0.0011)	(0.0007)	(0.0021)	(0.0014)	(0.0014)	(0.0009)	(0.0019)	(0.0015)
Finance, Insur & Real Estate	0.002	-0.0049	-0.0054	-0.0007	-0.0008	0.0031	0.0038	-0.0022	0.0027	-0.0013	-0.0008	0.0009
	(0.0034)	(0.0045)	(0.003)	(0.0012)	(0.0006)	(0.0009)	(0.0066)	(0.001)	(0.0012)	(0.0012)	(0.0017)	(0.0015)
Services	-0.0006	-0.0004	0.0055	0.0004	-0.001	0.0003	-0.0004	0.0003	-0.0014	0.0004	0.0011	0.0007
	(0.0027)	(0.001)	(0.015)	(0.0008)	(0.0006)	(0.0007)	(0.002)	(0.0016)	(0.0015)	(0.0012)	(0.0017)	(0.0015)

Sectors are determined according to the BEA industry classification. Standard errors are in parentheses. Estimates are from a value-added production function with endogenous and correlated effects. All specifications include industry and year fixed effects.

C Bootstrap for Network Data

Resampling network data needs to preserve the dependence structure between firms and across time. In my empirical application, I use the residual-based bootstrap whose asymptotic properties have been studied in the context of cross-sectional spatially correlated data by Jin and Lee (2012). I modify the procedure by treating my unbalanced panel as repeated cross-sections. I estimate the model, and obtain my first stage estimates $\widehat{\varphi}$ and residuals $\widehat{\varepsilon}_t$. If the residuals do not have zero mean, I subtract the empirical mean from each residual and obtain $\widetilde{\varepsilon}_t$. Then, for each $t = \{1, \cdots, T\}$ I draw samples of size n_t from $\widetilde{\epsilon}_{nt}$. Sampling R times, I obtain $\{\varepsilon_t^{*r}\}_{r=1}^{R}$ and use these to generate psuedosamples:

$$y_t^{*r} = \widehat{\varphi}_t + \varepsilon_t^{*r}$$

I re-estimate both the production function and productivity process on these pseudosamples, obtaining a set of elasticities $\{(\alpha_\ell^{*r}, \alpha_\ell^{*r})\}$ and productivity process parameters $\{(\rho^{*r}, \lambda^{*r}, \boldsymbol{\beta}^{*r})\}$ that I use to construct standard errors and confidence intervals.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451. (Cited on pages 3, 12, 23, and iii.)

Alfaro-Urena, A., I. Manelici, and J. P. Vasquez (2022). The effects of joining multinational supply chains: New evidence from firm-to-firm linkages. *The Quarterly Journal of Economics 137*(3), 1495–1552. (Cited on page 2.)

Alvarez, R. and R. A. López (2008). Is exporting a source of productivity spillovers? *Review of World Economics 144*(4), 723–749. (Cited on page 2.)

Amiti, M., C. Duprez, J. Konings, and J. Van Reenen (2023). Fdi and superstar spillovers: Evidence from firm-to-firm transactions. Technical report, National Bureau of Economic Research. (Cited on page 2.)

Arduini, T., E. Patacchini, E. Rainone, et al. (2015). Parametric and semiparametric IV estimation of network models with selectivity. Technical report, Einaudi Institute for Economics and Finance (EIEF). (Cited on pages 24 and 26.)

Atalay, E., A. Hortaçsu, J. Roberts, and C. Syverson (2011). Network structure of production. *Proceedings of the National Academy of Sciences 108*(13), 5199–5202. (Cited on page 6.)

Bazzi, S., A. V. Chari, S. Nataraj, and A. D. Rothenberg (2017). Identifying productivity spillovers using the structure of production networks. (Cited on page 3.)

Bernard, A. B. and A. Moxnes (2018). Networks and trade. *Annual Review of Economics* (0). (Cited on page xii.)

Blundell, R. and S. Bond (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric reviews 19*(3), 321–340. (Cited on pages 17 and i.)

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*, 41–55. (Cited on pages 3, 15, 20, 22, and 33.)

Cardoza, M., F. Grigoli, N. Pierri, and C. Ruane (2023). Worker mobility in production networks. (Cited on page 43.)

Collard-Wexler, A. and J. De Loecker (2016). Production function estimation with measurement error in inputs. Technical report, National Bureau of Economic Research. (Cited on pages iii and iv.)

De Loecker, J. (2013). Detecting learning by exporting. *American Economic Journal: Microeconomics 5*(3), 1–21. (Cited on pages 3 and 16.)

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, markups, and trade reform. *Econometrica 84*(2), 445–510. (Cited on page 3.)

De Loecker, J. and C. Syverson (2021). An industrial organization perspective on productivity. In *Handbook of Industrial Organization*, Volume 4, pp. 141–223. Elsevier. (Cited on page 2.)

Dieye, R. and B. Fortin (2017). Gender peer effects heterogeneity in obesity. (Cited on pages 4 and 32.)

Erdős, P. and A. Rényi (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5*(1), 17–60. (Cited on pages iii and xi.)

Fadeev, E. (2023). Creative construction: Knowledge sharing and cooperation between firms.". Technical report, Working Paper. (Cited on page 42.)

Gandhi, A., S. Navarro, and D. A. Rivers (2020). On the identification of gross output production functions. *Journal of Political Economy 128*(8), 2973–3016. (Cited on pages 3, 12, 16, 27, and 28.)

Garcia-Marin, A. and N. Voigtländer (2019). Exporting and plant-level efficiency gains: It's in the measure. *Journal of Political Economy 127*(4), 1777–1825. (Cited on page 3.)

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264. (Cited on pages 15 and 25.)

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica 85*(4), 1033–1063. (Cited on page 25.)

İmrohoroğlu, A. and Ş. Tüzel (2014). Firm-level productivity, risk, and return. *Management Science 60*(8), 2073–2090. (Cited on page i.)

Javorcik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages. *American Economic Review 94*(3), 605–627. (Cited on pages 2 and 15.)

Jin, F. and L. F. Lee (2012). On the bootstrap for spatial econometric models. *Working Paper*. (Cited on page iv.)

Keller, W. (2010). International trade, foreign direct investment, and technology spillovers. *Handbook of the Economics of Innovation 2*, 793–829. (Cited on pages 2 and 4.)

Keller, W. and S. R. Yeaple (2009). Multinational enterprises, international trade, and productivity growth: firm-level evidence from the united states. *The Review of Economics and Statistics 91*(4), 821–831. (Cited on page 2.)

Lee, L. F. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews 22*(4), 307–335. (Cited on pages 3, 15, 20, and 33.)

Lee, L. F. and J. Yu (2016). Identification of spatial Durbin panel models. *Journal of Applied Econometrics 31*(1), 133–162. (Cited on pages 3 and 15.)

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317–341. (Cited on page 3.)

Lim, K. et al. (2017). Firm-to-firm trade in sticky production networks. In *2017 Meeting Papers*, Number 280. Society for Economic Dynamics. (Cited on page 6.)

Liu, X. and L. F. Lee (2010). GMM estimation of social interaction models with centrality. *Journal of Econometrics 159*(1), 99–115. (Cited on page 31.)

Liu, X., E. Patacchini, and Y. Zenou (2014). Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior & Organization 103*, 39–59. (Cited on page 31.)

Malikov, E. and S. Zhao (2021, 08). On the Estimation of Cross-Firm Productivity Spillovers with an Application to FDI. *The Review of Economics and Statistics*, 1–46. (Cited on page 4.)

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542. (Cited on page 14.)

Merlevede, B. and A. Theodorakopoulos (2023). Intangibles within firm boundaries. (Cited on page 4.)

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment. *Econometrica 64*(6), 1263–1297. (Cited on pages 3, 24, and 26.)

Patacchini, E., E. Rainone, and Y. Zenou (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior & Organization 134*, 190–227. (Cited on pages 4 and 32.)

Qu, X., L. F. Lee, and J. Yu (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics 197*(2), 173–201. (Cited on page 24.)

Serpa, J. C. and H. Krishnan (2018). The impact of supply chains on firm-level productivity. *Management Science 64*(2), 511–532. (Cited on pages 2, 6, and 15.)

Stoyanov, A. and N. Zubanov (2012). Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics 4*(2), 168–98. (Cited on page 43.)

Syverson, C. (2011). What determines productivity? *Journal of Economic literature 49*(2), 326–65. (Cited on page 2.)

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *The Journal of Industrial Economics 55*(3), 529–569. (Cited on pages iii and iv.)

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters 104*(3), 112–114. (Cited on page 3.)

Online Appendix

OA1 Derivation of Bias Terms

In this section, I derive expressions for the bias in production function elasticities shown in section 3.3.

$$
\begin{aligned}
y_t &= \alpha_\ell \ell_t + \alpha_k k_t + \omega_t + \varepsilon_t \\
\omega_t &= \rho(I-\lambda G_t)^{-1}\omega_{t-1} + (I-\lambda G_t)^{-1}\zeta_{it} = \rho \sum_{s=0} \lambda^s G_t^s \omega_{t-1} + \sum_{s=0} \lambda^s G_t^s \zeta_t \\
\implies y_t &= \alpha_\ell \ell_t + \alpha_k k_t + \rho \sum_{s=0} \lambda^s G_t^s \omega_{t-1} + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \varepsilon_t \\
\omega_{t-1} &= \varphi_{t-1} - \alpha_\ell \ell_{t-1} - \alpha_k k_{t-1} \\
\implies y_t &= \alpha_\ell \ell_t + \alpha_k k_t + \rho \sum_{s=0} \lambda^s G_t^s \left(\varphi_{t-1} - \alpha_\ell \ell_{t-1} - \alpha_k k_{t-1}\right) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \varepsilon_t \\
y_{t-1} &= \varphi_{t-1} + \varepsilon_t \\
\implies y_t &= \alpha_\ell \ell_t + \alpha_k k_t + \rho \sum_{s=0}^{\infty} \lambda^s G_t^s \left(y_{t-1} - \alpha_\ell \ell_{t-1} - \alpha_k k_{t-1} - u_{t-1}\right) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \varepsilon_t
\end{aligned}
$$

Let $\Delta^G x_t = x_t - \rho \sum_{s=0}^{\infty} \lambda^s G_t^s x_{t-1}$, $\Delta_{x_t}^{err} = \rho \sum_{s=1}^{\infty} \lambda^s G_t^s x_{t-1}$ and $\Delta x_t = x_t - \rho x_{t-1} = \Delta^G x_t + \Delta_{x_t}^{err}$. This implies:

$$
\Delta^G y_t = \alpha_\ell \Delta^G \ell_t + \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{52}
$$

This is equivalent to the dynamic panel approach in Blundell and Bond (2000). However, growth in output, labor and capital have been purged of the variation from network effects in the previous period. When we assume no spillovers, we estimate:

$$
\Delta y_t = \alpha_\ell \Delta \ell_t + \alpha_k \Delta k_t + u_t \tag{53}
$$

Therefore, in the linear AR1 case, ignoring spillovers is equivalent to introducing non-classical measurement error into both output and inputs.

Bias from ignoring spillovers can also be characterized as an omitted variables problem. By estimating equation (53), where $u_t = \rho \sum_{s=1}^{\infty} \lambda^s G_t^s \omega_{t-1} + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \varepsilon_t$. That is, the standard ACF procedure succeeds in eliminating the endogeneity problem that arises from input decisions depending on its own productivity, but is unable to account for the influence of its network's past productivity. In the OP case where the labor elasticity is estimated in the first stage, the second stage is equivalent to estimating:

$$\Delta^G \widetilde{y}_t = \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{54}$$

where $\widetilde{y}_t = y_t - \widehat{\alpha}_\ell \ell_t$. Then by estimating $\Delta \widetilde{y}_t = \alpha_k \Delta k_t + u_t$ under the standard assumption of no-spillovers:

$$plim\ \widehat{\alpha}_k = \frac{cov(\Delta k_t, \Delta \widetilde{y}_t)}{var(\Delta k_t)} \tag{55}$$

$$plim\ = \alpha_k \left(1 - \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s k_{t-1})}{var(\Delta k_t)}\right) + \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s \widetilde{y}_{t-1})}{var(\Delta k_t)} \tag{56}$$

When productivity is mismeasured by ignoring spillovers, the resulting estimates also result in incorrect conclusions about spillover effects. When (α_ℓ, α_k) are consistently estimated,

$$plim\ \widehat{\omega}_t = \varphi_t - \alpha_\ell \ell_t - \alpha_k k_t = \omega_t \tag{57}$$

However, when we estimate $(\widetilde{\alpha}_\ell, \widetilde{\alpha}_k) = (\widehat{\alpha}_\ell + \alpha_\ell^{err}, \widehat{\alpha}_k + \alpha_k^{err})$, to obtain $\widetilde{\omega}_t = \widehat{\varphi}_t - \widetilde{\alpha}_\ell \ell_t - \widetilde{\alpha}_k k_t$. Then

$$\widetilde{\omega}_t = \widehat{\varphi}_t - \widetilde{\alpha}_\ell \ell_t - \widetilde{\alpha}_k k_t = \widehat{\omega}_t - \omega_t^{err} \tag{58}$$

where $\omega_t^{err} = \alpha_\ell^{err} \ell_t + \alpha_k^{err} k_t$. In the generalized 2SLS procedure for estimating network effects, we estimate $\widetilde{\lambda}$ in the first stage by using $G_t \widetilde{\omega}_{t-1}$ as an instrument for $G_t \widetilde{\omega}_t$ in this equation:[29] And so we estimate the following instead of the true model:

$$\widetilde{\omega}_t = \rho \widetilde{\omega}_{t-1} + \lambda G_t \widetilde{\omega}_t + v_t \tag{59}$$

[29] Further lags of the network effect can be used ($G_t^2 \widetilde{\omega}_t, G_t^3 \widetilde{\omega}_t$ and so on). However, for ease of exposition, I focus on the just-identified case.

OA2 Monte Carlo Setup

The Monte Carlo setup closely follows Collard-Wexler and De Loecker (2016), Van Biesebroeck (2007) and Ackerberg et al. (2015) with modifications for network generation and the inclusion of spillovers in the productivity process. I generate a balanced panel of 1000 firms over 10 time periods.

OA2.1 Production Function

I use a structural value-added production function that is Leontief in materials.

$$Y_{it} = \min\{L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}}, \alpha_m M_{it}\} e^{\varepsilon_{it}} \tag{60}$$

$$\text{In logs,} \quad y_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it}, \quad \text{where } \varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2) \tag{61}$$

I set $\alpha_\ell = 0.6, \alpha_k = 0.4$ and $\sigma_\varepsilon^2 = 1$

OA2.2 Productivity Process and Network

Productivity evolves according to an AR1 process that allows for contemporaneous endogenous productivity spillovers. In vectorized form:

$$\omega_t = \beta_1 \iota + \rho \omega_{t-1} + \lambda G_t \omega_t + \zeta_t, \quad \text{where } \zeta_{it} \sim \mathcal{N}(0, \sigma_\zeta^2 = 5) \tag{62}$$

I generate productivity using the reduced form of the above equation:

$$\omega_t = (\mathrm{I} - \lambda G_t)^{-1} (\beta_1 \iota + \rho \omega_{t-1} + \zeta_t) \tag{63}$$

G_t is the interaction matrix defined as in section 3.2 derived from the network. I generate exogenous networks using Erdős and Rényi (1960) graphs, also known as binomial graphs. Firms are edges are formed $A_{ijt} \overset{i.i.d.}{\sim} \text{Bern}(p)$.

OA2.3 Intermediate Input Demand

Materials demand is given by:

$$M_{it} = \frac{1}{\alpha_m} K_{it}^{\alpha_k} L_{it}^{\alpha_\ell} e^{\omega_{it}} \iff m_{it} = \alpha_k k_{it} + \alpha_\ell \ell_{it} + \omega_{it} - \ln(\alpha_m) \tag{64}$$

OA2.4 Labor Demand

Wages, W_{it} are firm-year specific and distributed log-normally: $\ln(W_t) \sim \mathcal{N}(0, \sigma_w^2)$. Then each firm chooses optimal labor according to:

$$L_{it} = \left(\alpha_\ell \frac{K_{it}^{\alpha_k}}{W_{it}} e^{\omega_{it}}\right)^{\frac{1}{1-\alpha_\ell}} \iff \ell_{it} = \frac{1}{1-\alpha_\ell}\left(\ln(\alpha_\ell) + \alpha_k k_{it} + \omega_{it} - \ln(W_{it})\right) \quad (65)$$

OA2.5 Capital and Optimal Investment

Capital is accumulated according to $K_{it} = (1-\delta)K_{it-1} + I_{t-1}$ where $\delta = 0.2$.

Investment is subject to convex adjustment costs $c(I_{it}) = \frac{b}{2} I_{it}^2$ with $b = 0.3$. Optimal investment can be derived by setting up the profit maximization problem:[30]

$$\Pi_{it} = L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}} - W_{it} L_{it} - \frac{b}{2} I_{it}^2 \quad (66)$$

Here, I assume perfect competition and normalize the price of output to 1. The firm's value function is :

$$\begin{aligned} V(L_{it}, K_{it}, W_{it}, \omega_{it}) = \max_{L_{it}, K_{it}} & L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}} - W_{it} L_{it} - \frac{b}{2} I_{it}^2 \\ & + \beta\, \mathbb{E}_{it}\, V(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \end{aligned} \quad (67)$$

β is the discount factor set to 0.95. Optimal investment solves the Euler equation $\frac{\partial V}{\partial I} = 0$:

$$b I_{it} = \beta\, \mathbb{E}_{it}\, V_K(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \quad (68)$$

The envelope condition yields:

$$V_K(L_{it}, K_{it}, W_{it}, \omega_{it}) = \alpha_k L_{it}^{\alpha_\ell} K_{it}^{\alpha_k - 1} e^{\omega_{it}} + \beta(1-\delta)\, \mathbb{E}_{it}\, V_K(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \quad (69)$$

Substituting in (65) and (68):

$$V_K(L_{it}, K_{it}, W_{it}, \omega_{it}) = \alpha_k \alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} K_{it}^{\frac{\alpha_k + \alpha_\ell - 1}{1-\alpha_\ell}} W_{it}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it}}{1-\alpha_\ell}} + b(1-\delta) I_{it} \quad (70)$$

[30]This derivation follows Collard-Wexler and De Loecker (2016) and Van Biesebroeck (2007).

Given a constant returns to scale technology ($\alpha_\ell + \alpha_k = 1$), the Euler equation becomes:

$$I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}}\,\mathbb{E}_{it}\left[W_{it+1}^{\frac{-\alpha_\ell}{1-\alpha_\ell}}e^{\frac{\omega_{it+1}}{1-\alpha_\ell}}\right] + \beta(1-\delta)\,\mathbb{E}_{it}\,I_{it+1} \tag{71}$$

$$\implies I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}}\sum_{\tau=0}^{\infty}\beta^\tau(1-\delta)^\tau\,\mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}}e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] \tag{72}$$

Since wages and productivity are drawn independently,

$$\mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}}e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] = \mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}}\right]\mathbb{E}_t\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right]$$

for all $\tau \geq 0$. Furthermore, $\ln(W_{it}) \sim \mathcal{N}(0,\sigma_w)^2 \implies \mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}}\right] = \exp\left(\frac{\alpha_\ell^2\sigma_w^2}{2(1-\alpha_\ell)^2}\right)$. The value of $\mathbb{E}_t\left[e^{\frac{\omega_{t+1+\tau}}{1-\alpha_\ell}}\right]$ depends on the productivity process:

$$\omega_{t+1+\tau} = \rho^{\tau+1}\prod_{r=0}^{\tau}(I-\lambda G_{t+\tau+1-r})^{-1}\omega_t + \sum_{r=0}^{\tau}\rho^r\prod_{s=0}^{r}(I-\lambda G_{t+\tau+1-s})^{-1}\varepsilon_{t+\tau+1-r} \tag{73}$$

$\mathbb{E}_t\left[e^{\frac{\omega_{t+1+\tau}}{1-\alpha_\ell}}\right]$ depends on the whether spillovers exist, and if they do, how firms form expectations about future links. When there are no spillovers $\lambda = 0$:

$$\mathbb{E}_t = \exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell}\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2}\right) \tag{74}$$

Let $\overline{G}$ represent the result of firms' beliefs about their future network. For example, if networks are non-stochastic or firms naively believe that $G_{t+\tau} = G_t\ \forall\tau > 0$, then we can set $\overline{G} = G_{t+1}$, which is deterministic given our previous assumption that $G_{t+1} \in \mathcal{I}_t$:

$$\mathbb{E}_{it}\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] = \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2(1-\lambda)^{2(r+1)}}\right)$$

Therefore, optimal investment choice reduces to a function of parameters and current productivity:

$$I_t = \frac{\beta \alpha_k}{b} \alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{\alpha_\ell^2 \sigma_w^2}{2(1-\alpha_\ell)^2}\right) \sum_{\tau=0}^{\infty} \beta^\tau (1-\delta)^\tau \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I - \lambda \overline{G})^{-(\tau+1)} \omega_t \right. \\ \left. + \frac{\sigma_\zeta^2}{2(1-\alpha_\ell)^2(1-\lambda)^2} \sum_{r=0}^{\tau} \left(\frac{\rho}{1-\lambda}\right)^{2r}\right) \tag{75}$$

When there are no spillovers, this reduces to:

$$I_t = \frac{\beta \alpha_k}{b} \alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{\alpha_\ell^2 \sigma_w^2}{2(1-\alpha_\ell)^2}\right) \sum_{\tau=0}^{\infty} \beta^\tau (1-\delta)^\tau \exp\left(\frac{\rho^{\tau+1}\omega_t}{1-\alpha_\ell} + \frac{\sigma_\zeta^2 \sum_{r=0}^{\tau} \rho^{2r}}{2(1-\alpha_\ell)^2}\right) \tag{76}$$

For alternative assumptions on the productivity process, such as a quadratic AR1 process, and endogenous network formation, it is not feasible to derive an closed-form solution as above. However, as long technology exhibits constant returns to scale, I approximate optimal investment as follows. Firstly, given $|\beta(1-\delta)| < 1$, then for some tolerance level close to zero, $\beta^\tau(1-\delta)^\tau <$ tolerance. Therefore, I can choose M sufficiently high such that $\sum_{\tau=0}^{M} \beta^\tau (1-\delta)^\tau \, \mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right]$ is a good approximation for $\sum_{\tau=0}^{\infty} \beta^\tau (1-\delta)^\tau \, \mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right]$. I set a tolerance level of e^{-4}, and given $\beta(1-\delta) = 0.95(1-0.2)$, then $M = 34$.

Next, at each time t, I draw 100 realizations of the sequence $\{\omega_{it+1+\tau}\}_{\tau=0}^{M}$ for each firm i and approximate $\mathbb{E}_{it}\left[exp\left(\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}\right)\right] = \frac{1}{100}\sum_{s=0}^{100} exp\left(\frac{\omega_{it+1+\tau,s}}{1-\alpha_\ell}\right)$.

OA3 Monte Carlo Experiments

I conduct two sets of experiments to assess the performance of the standard ACF estimator and my modified procedure when various types of network effects are present. In the first set of experiments, I sequentially add in each type of network effect and demonstrate their impact on the bias and efficiency of parameters estimated with the standard and modified approaches. Second, I vary the size of the endogenous effect, network density and productivity persistence to these factors affect spillover identification.

For all three experiments, I draw a balanced sample of 1000 firms over 10 years. I use a Cobb-Douglas production function in logs:

$$y_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it}$$

where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. I set $\alpha_\ell = 0.6, \alpha_k = 0.4$ and $\sigma_\varepsilon^2 = 1$.[31] The productivity process varies depending on the experiment. To avoid the impact of arbitrary initial values, I simulate 20 periods and discard the first 10.

To induce variation in cluster (component) size and the length of supply chains, I split the firms into four industries with 400, 300, 200, and 100 firms in the first, second, third and fourth industries respectively and construct an inter-industry trade structure as follows: Industry 1 sells 17%, 33% and 44% percent of its output to industries 2, 3 and 4 respectively. Industry 2 sells to 50% each to 3 and 4, while industry 3 sells all its output to industry 4, which sells nothing to other firms. This structure is fixed over time, and does not represent the actual network but is a measure of industry compatibility that I use to generate both exogenous and endogenous networks as described below.[32]

[31] See section OA2 in the appendix for further details on the Monte Carlo setup.

[32] I have set up data-generating processes for my Monte Carlo experiments to be as simple as possible while allowing for network effects. However, it is worth noting that these DGPs do not reflect important features of firm-level empirical data, particularly fat-tailed productivity and network degree distributions. Exploring how these features would affect bias and precision on my estimator is left for future work.

OA3.1 Experiment 1: Comparison of Estimates from Standard and Modified ACF Procedures

I simulate five data generating processes (DGPs) in which productivity evolves as follows:

$$\omega_t = \beta_1 \iota + \rho \omega_{t-1} + \beta_x x_t + \lambda G_t \omega_t + \beta_{\bar{x}} G_t x_t + c_{\psi_t} + \zeta_t \tag{77}$$

where $\zeta_{it} \sim \mathcal{N}(0, \sigma_\zeta^2)$. To induce a non-linear relationship between x and capital, I generate it according to $x = 0.5 \ln(\sqrt{K}_{t-1}) + \widetilde{x}$, where $\widetilde{x} \sim \mathcal{N}(-2, \sigma_{\widetilde{x}}^2)$. Since it depends on K_{t-1}, it is not correlated with ζ_t. I set $\beta_1 = 0.5, \rho = 0.6, \beta_x = 0.4, \sigma_\zeta^2 = 1.25$, and $\sigma_{\widetilde{x}}^2 = 5$.

For DGPs 1 to 4, I generate an exogenous directed network in each period by randomly assigning links with probability $P(A_{ijt} = 1) = \frac{indshare_{ij}}{indsize_j}$ where $indshare_{ij}$ is the compatibility of i and j's industries obtained from the industry compatibility matrix described above, while $indsize_j$ is the number of firms in j's industry. DGP 1 has no network effects ($\lambda = 0, \beta_{\bar{x}} = 0, c_{\psi_t} = 0$) and exogenous network formation, and the ACF estimates should be consistent. DGP 2 features only the endogenous effect ($\lambda = 0.3$) and DGP 3 adds in the contextual effect ($\beta_{\bar{x}} = 0.3$). In DGP 4, I draw component fixed effects in each period from a normal distribution with a mean of 1 and a standard deviation of 1 ($c_{\psi_t} \mathcal{N}(1, 1)$). For DGP 5, I start with an exogenous network in the first period, then simulate future networks using the model in section 5 with the coefficient of the selection term $\delta = \frac{\sigma_{\zeta\xi}}{\sigma_\xi^2} = 0.003$.

I consider 4 estimators. The first is a standard ACF that assumes no network effects with a second-degree polynomial approximation in the first and second stages.. Using the TFP measure obtained from ACF, I estimate network effects with the generalized 2SLS procedure described in section 4.3. This is the approach typically used in empirical studies of productivity spillovers. ACF-N is my modified procedure that jointly estimates productivity and network effects. ACF-ND uses global differencing to eliminate correlated effects, and ACF-NDS accounts for selection using the network formation model in section 5.3. All estimators use a second-degree polynomial in capital, labor and materials in the first stage, and a linear productivity process in the second. The results are shown in table 1.

Under DGP 1, all estimators perform well when estimating both the production

Table 1: Comparison of Estimates from Standard ACF and Modified ACF Procedures

DGP	Estimator		Elasticities		Productivity Process Coefficients				
			α_ℓ	α_k	ρ	β_x	$\beta_{\bar{x}}$	λ	$\frac{\sigma_{\zeta\xi}}{\sigma_\xi^2}$
DGP 1		True values	0.6	0.4	0.6	0.4	0.0	0.0	0.0
	ACF	Mean	0.599	0.4	0.6	0.401	0.	-0.001	
		Std. Dev.	(0.025)	(0.061)	(0.015)	(0.026)	(0.009)	(0.01)	
	ACF-N	Mean	0.602	0.392	0.601	0.398	0.	-0.001	
		Std. Dev.	(0.018)	(0.061)	(0.016)	(0.019)	(0.009)	(0.01)	
	ACF-ND	Mean	0.603	0.389	0.601	0.397	-0.	-0.	
		Std. Dev.	(0.024)	(0.064)	(0.016)	(0.024)	(0.01)	(0.011)	
	ACF-NDS	Mean	0.603	0.39	0.601	0.397	-0.	0.	-0.
		Std. Dev.	(0.024)	(0.064)	(0.016)	(0.025)	(0.01)	(0.012)	(0.002)
DGP 2		True values	0.6	0.4	0.6	0.4	0.1	0.3	0.0
	ACF	Mean	0.595	0.516	0.556	0.402	0.092	0.332	
		Std. Dev.	(0.035)	(0.07)	(0.017)	(0.035)	(0.016)	(0.042)	
	ACF-N	Mean	0.601	0.401	0.596	0.399	0.121	0.242	
		Std. Dev.	(0.018)	(0.046)	(0.016)	(0.018)	(0.013)	(0.026)	
	ACF-ND	Mean	0.602	0.398	0.595	0.397	0.118	0.249	
		Std. Dev.	(0.028)	(0.055)	(0.016)	(0.028)	(0.014)	(0.026)	
	ACF-NDS	Mean	0.602	0.396	0.596	0.397	0.115	0.257	-0.004
		Std. Dev.	(0.027)	(0.055)	(0.016)	(0.028)	(0.014)	(0.026)	(0.002)
DGP 3		True values	0.6	0.4	0.6	0.4	0.1	0.3	0.0
	ACF	Mean	0.616	0.496	0.479	0.362	0.121	0.357	
		Std. Dev.	(0.169)	(0.417)	(0.171)	(0.161)	(0.102)	(0.496)	
	ACF-N	Mean	0.741	0.162	0.514	0.257	0.082	0.222	
		Std. Dev.	(0.154)	(0.215)	(0.269)	(0.154)	(0.072)	(0.62)	
	ACF-ND	Mean	0.614	0.368	0.605	0.385	0.109	0.266	
		Std. Dev.	(0.032)	(0.052)	(0.017)	(0.032)	(0.012)	(0.018)	
	ACF-NDS	Mean	0.614	0.368	0.605	0.385	0.108	0.269	-0.002
		Std. Dev.	(0.032)	(0.052)	(0.018)	(0.032)	(0.012)	(0.018)	(0.002)
DGP 4		True values	0.6	0.4	0.6	0.4	0.1	0.3	0.003
	ACF	Mean	0.607	0.35	0.603	0.374	0.128	0.255	
		Std. Dev.	(0.138)	(0.239)	(0.147)	(0.166)	(0.109)	(0.122)	
	ACF-N	Mean	0.705	0.183	0.637	0.291	0.067	0.236	
		Std. Dev.	(0.137)	(0.213)	(0.184)	(0.142)	(0.056)	(0.2)	
	ACF-ND	Mean	0.619	0.368	0.61	0.383	0.091	0.281	
		Std. Dev.	(0.073)	(0.116)	(0.056)	(0.07)	(0.023)	(0.037)	
	ACF-NDS	Mean	0.621	0.362	0.612	0.38	0.09	0.28	0.001
		Std. Dev.	(0.078)	(0.129)	(0.064)	(0.076)	(0.026)	(0.037)	(0.002)

Based on 1000 replications.

function and the productivity process. Furthermore, precision is not diminished. It is important to note that allowing for spillovers under the modified procedure does not produce spurious estimates of network effects. With the combined impact of endogenous and contextual effects in DGP 2, ACF significantly overestimates the capital coefficient but still gives reasonable estimates of network effects in the productivity process, although the endogenous effect is slightly overestimated. All three modified procedures yield estimates of the input elasticities that are close to the truth but slightly underestimate λ.

When there are network fixed effects, my benchmark procedure, ACF-N overestimates the labor coefficient and underestimates capital elasticity, the persistence parameter, and the endogenous effect. This signals the need for caution when introducing network terms without accounting for correlated effects: indeed, the standard ACF performs better because all network terms containing G_t introduce bias due to their corrrelation with the error term. Differencing improves both consistency and precision, with standard deviations up to 60 times smaller than under ACF and ACF-N. Bias due to endogenous network formation is negligible, presumably because the coefficient $\frac{\sigma_{\zeta\xi}}{\sigma_{\xi}^2}$ on the omitted variable is small. Other than reduced precision when compared to ACF-ND, estimates of the productivity process and elasticities are not different from when selection is accounted for with ACF-NDS.

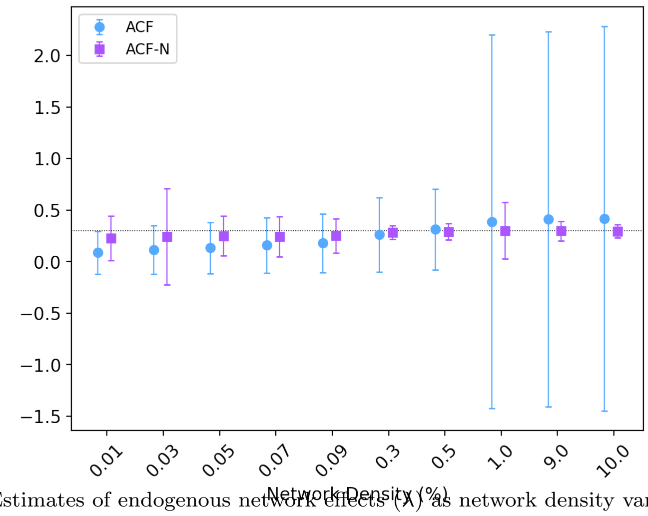
OA3.2 Experiment 2: Effect of Network Density on Bias and Precision

In this experiment, I further explore how precision and consistency vary with network density in the presence of an endogenous spillover. I employ a quadratic AR1 process for productivity:

$$\omega_t = \beta_1 + \rho_1 \omega_{t-1} + \rho_2 \omega_{t-1}^2 + \lambda G_t \omega_t + \zeta_t \tag{78}$$

where $\zeta_{it} \sim \mathcal{N}(0, \sigma_\zeta^2)$. I set $\beta_1 = 0.5, \rho_2 = -0.01$, and $\sigma_\zeta^2 = 5$. The quadratic term is necessary to explore high values of λ and ρ_1. If productivity is persistent and the endogenous spillover is also large, then the simulated values of productivity grow quite large for some firms, and the resulting investment series soon tends to infinity

Figure 1: Effect of Network Density on Spillover Estimates



Estimates of endogenous network effects (λ) as network density varies. See table 2 for the full set of parameter estimates.

for highly productive firms[33]. The quadratic term serves as a dampener to control the size ω_t in the simulation.[34] Additionally, it allows for the comparison of ACF and my modified procedure when the productivity process is not linear.

To vary network density, I draw random exogenous networks using Erdős and Rényi (1960) graphs, also known as binomial graphs. Edges are formed $A_{ijt} \overset{i.i.d.}{\sim} \text{Bern}(p)$ and the density of the graph is equal to the probability of a link forming between two firms, p. This class of graphs has several features worth noting. First, intransitivity rises as the density falls. This is an advantage because intransitivity helps with identification of the endogenous network effect, so we can expect more precise estimates as the network gets more sparse. Secondly, when $p > \frac{1}{N_t}$, a giant component emerges that contains more vertices than any other component of the network. In my Monte Carlo experiments, this means that for graphs with density > 0.001 the infinite series of terms G_t^s will go to zero much more slowly than with density ≤ 0.001. Therefore, one would expect the potential bias to be greater as density increases, particularly once it crosses the 0.001 threshold. However, it is worth noting that the resulting degree distribution is binomial $B(N_t - 1, p)$, which is approximately normal whereas buyer-supplier networks have empirically been found

[33] See details on optimal investment in section OA2.5 in the appendix

[34] It is also worth mentioning that in empirical applications, estimating flexible forms of the productivity process may be necessary. Otherwise, the linearity of the Markov process may force estimates of λ to be small or negative.

to follow a Pareto (power-law) degree distribution (Bernard and Moxnes, 2018).

Table 2: Effect of Sparsity on Bias and Precision (Quadratic AR1)

Density (%)	Estimator	Elasticities		Productivity Process Coefficients				Density (%)	Estimator	Elasticities		Productivity Process Coefficients			
		α_ℓ	α_k	β_1	ρ_1	ρ_2	λ			α_ℓ	α_k	β_1	ρ_1	ρ_2	λ
		0.6	0.4	0.5	0.8	-0.01	0.3			0.6	0.4	0.5	0.8	-0.01	0.3
0.01	ACF	0.603	0.358	-0.125	0.809	-0.01	0.087	0.5	ACF	0.605	0.637	-0.109	0.462	-0.019	0.312
		(0.024)	(0.239)	(2.369)	(0.216)	(0.003)	(0.106)			(0.067)	(0.138)	(0.987)	(0.161)	(0.017)	(0.201)
	ACF-N	0.617	0.413	-0.23	0.76	-0.01	0.226		ACF-N	0.608	0.388	0.509	0.818	-0.01	0.291
		(0.057)	(0.165)	(2.845)	(0.196)	(0.022)	(0.109)			(0.03)	(0.057)	(0.298)	(0.05)	(0.007)	.042)
0.03	ACF	0.604	0.359	0.122	0.81	-0.01	0.113	0.7	ACF	0.606	0.639	-0.405	0.47	-0.026	0.545
		(0.024)	(0.216)	(1.975)	(0.19)	(0.003)	(0.12)			(0.073)	(0.149)	(13.879)	(0.694)	(0.207)	(0.696)
	ACF-N	0.632	0.381	0.379	0.764	-0.011	0.241		ACF-N	0.607	0.385	0.437	0.818	-0.01	0.294
		(0.093)	(0.113)	(1.456)	(0.195)	(0.038)	(0.238)			(0.027)	(0.068)	(0.339)	(0.053)	(0.002)	(0.027)
0.05	ACF	0.605	0.377	0.209	0.798	-0.01	0.132	1	ACF	0.606	0.639	0.011	0.452	-0.019	0.388
		(0.024)	(0.195)	(1.691)	(0.169)	(0.003)	(0.126)			(0.072)	(0.153)	(1.8)	(0.2)	(0.046)	(0.925)
	ACF-N	0.641	0.371	0.412	0.753	-0.009	0.25		ACF-N	0.606	0.386	0.404	0.815	-0.01	0.301
		(0.106)	(0.113)	(1.226)	(0.217)	(0.034)	(0.097)			(0.031)	(0.078)	(0.404)	(0.061)	(0.005)	0.14)
0.07	ACF	0.606	0.387	0.271	0.791	-0.01	0.159	5	ACF	0.606	0.643	0.054	0.414	-0.024	0.446
		(0.027)	(0.182)	(1.509)	(0.158)	(0.003)	(0.137)			(0.072)	(0.152)	(3.357)	(0.782)	(0.151)	(0.833)
	ACF-N	0.646	0.362	0.51	0.745	-0.007	0.243		ACF-N	0.605	0.388	0.417	0.813	-0.01	0.299
		(0.116)	(0.116)	(0.506)	(0.237)	(0.046)	(0.1)			(0.032)	(0.084)	(0.399)	(0.065)	(0.004)	.062)
0.09	ACF	0.606	0.411	0.266	0.771	-0.01	0.18	9	ACF	0.606	0.643	0.001	0.426	-0.021	0.413
		(0.03)	(0.168)	(1.359)	(0.147)	(0.004)	(0.145)			(0.071)	(0.154)	(1.774)	(0.406)	(0.074)	(0.928)
	ACF-N	0.635	0.371	0.532	0.767	-0.011	0.252		ACF-N	0.604	0.388	0.42	0.813	-0.01	0.297
		(0.101)	(0.098)	(0.308)	(0.2)	(0.037)	(0.085)			(0.03)	(0.084)	(0.401)	(0.063)	(0.003)	(0.049)
0.3	ACF	0.602	0.617	-0.343	0.523	-0.017	0.261	10	ACF	0.606	0.642	-0.003	0.425	-0.021	0.417
		(0.053)	(0.114)	(1.261)	(0.137)	(0.011)	(0.184)			(0.073)	(0.157)	(1.812)	(0.417)	(0.077)	(0.952)
	ACF-N	0.611	0.389	0.585	0.815	-0.01	0.283		ACF-N	0.604	0.388	0.42	0.814	-0.01	0.296
		(0.038)	(0.049)	(0.256)	(0.06)	(0.009)	(0.035)			(0.028)	(0.083)	(0.403)	(0.061)	(0.003)	(0.032)

Based on 1000 replications. Standard deviations in parentheses.

Figure 1 and table 2 shows the results. ACF estimates of the capital elasticity appear unbiased for densities ≤ 0.001 and increases to over 50% of the true value for densities above 0.001. Estimates of λ increase with density while ρ_1 moves in the opposite direction. In comparison, my benchmark procedure ACF-N provides stable and consistent estimates of both the elasticities and productivity process at most densities. When the network is very sparse, however, my procedure underestimates λ and does so with less precision because the instrument $G_t^2\omega_{t-1}$ is weaker when there are few triads in the network.