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

Click here for the latest version

Ebehi Iyoha*

This version: October 2020

Abstract

This paper examines the extent to which efficiency gains diffuse over a network of firms. Many empirical studies investigating firm-to-firm spillovers measure productivity by exploiting proxy variables or first-order conditions to estimate a production function. However, these methods implicitly rule out the interdependence of firms' outcomes and decisions through productivity spillovers. I show that ignoring network effects when estimating production functions may lead to substantial biases in both measured productivity and estimated spillovers. Furthermore, the direction of bias cannot be generally be predicted *a priori*: depending on the structure of the network and persistence of productivity over time, estimates of network effects may be biased upwards or downwards. To address this limitation of existing methods, I develop a framework to jointly estimate network effects and productivity in value-added and gross output production functions. I demonstrate that my approach is robust to endogenous network formation and can accommodate network effects that are heterogeneous in direction and firm characteristics. Using this method, I characterize productivity spillovers over the US production network from 1977 to 2016 and find substantial heterogeneity by direction, industry, firm size and over time. My results suggest that, due to spillovers, the average firm in 1978 would be 20 percent more productive by 2016. In addition, a 10 percent increase in the productivity of the most central firm in each year would result in a 2 to 4 percent rise in aggregate TFP through spillovers alone.

1 Introduction

Production function estimation is at the heart of a number of literatures in economics. A vast body of work in macroeconomics seeks to understand productivity differences between countries, and decompose the sources of aggregate productivity growth. From understanding the impact of trade liberalization on the distribution of firm-level productivity to the estimation of markups in industrial organization, many important questions hinge on the accurate measurement of total factor productivity (TFP) and input elasticities.

A significant finding of the TFP estimation literature is that firms exhibit marked differences in productivity, even within narrowly-defined industries, and a vast body of work seeks to explain this dispersion.[1] One possible explanation is that firms may affect each other in ways that do

*Vanderbilt University, Department of Economics. Contact: e.iyoha@vanderbilt.edu

[1] See Syverson (2011) for a review.

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 their counterparts' productivities through activities such as foreign direct investment (FDI) and exporting.[2] Javorcik (2004) finds that FDI in Lithuania has a positive effect on the productivity of domestic firms through backward linkages, while Keller and Yeaple (2009) document the existence of horizontal spillovers from multinationals to US firms. Likewise, Alvarez and López (2008) provide evidence from Chile of positive productivity spillovers from domestic and foreign-owned exporters on their suppliers, while Alfaro-Urena et al. (2019) finds TFP gains of $6-9\%$ among Costa Rican firms after they begin to supply to multinational corporations.

More recent work considers 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: a firm may improve its productivity by learning from others or might free-ride on the productivity-enhancing practices of its partners. Serpa and Krishnan (2018) examine this question with data on firm-level buyer-supplier relationships in the US, while Bazzi et al. (2017) use input-output matrices to construct measures of the relationships between Indonesian firms. Both studies find 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 rely on an assumption that a firm's future productivity depends only on its own past productivity and characteristics. Alternative methods like Gandhi et al. (2020) rely on first order conditions for identification, but still rely on 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. Using Monte Carlo experiments, I demonstrate that input elasticities are generally not consistent when the law of motion for productivity precludes spillovers. 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) shows 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 comes from revenue-based productivity measures that cannot disentangle the lower prices firms charge upon entry into export markets from their increased efficiency. Unfortunately, the direction of bias in spillover estimates is not so clear-cut. I find that, depending on the structure of the network and the persistence of productivity over time, estimat-

[2] See Keller (2010) for a review of the evidence on spillovers from FDI and exporting.

ing spillovers on mismeasured TFP can lead us to *overestimate* network effects in some cases and *underestimate* them in others.

Secondly, 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 literature Lee (2003); Bramoullé et al. (2009); Lee and Yu (2016), with an important distinction: these papers deal with outcomes that are observed, 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. An advantage of the proposed method is that, even in the absence of spillovers, the estimator does not generate spurious network effects and provides consistent, albeit less precise, estimates of the input elasticities. It can also accommodate confounders such as common shocks to firms in the same network and the endogeneity of network formation. I extend the framework to examine heterogeneous spillovers in the manner of Dieye and Fortin (2017), that may vary by the nature of the relationship between firms.

Third, I apply this methodology to examine the transmission of productivity 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, and substantial heterogeneity by firm size and industry. Estimates suggest that if the most connected firm in a given year was 10 percent more productive, spillovers would lead to an increase in aggregate TFP of 2 to 4 percent. Furthermore, the cumulative impact of spillovers over time implies that the average firm in 1978 would be 20 percent more productive by 2016 due to spillovers alone, even if it was only connected to a single, less productive firm. This is due to the mutually reinforcing nature of endogeneous network effects. Decomposing the spillovers by sector, shows that manufacturing firms have benefited from almost all other sectors and wholesalers are an important source of efficiency gains. In fact, estimates of positive spillovers are strongest from 1997 to 2006, when retailers and wholesalers were becoming more central to the US production network than manufacturing firms.

In the next section, I describe the data and features of the sample 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. Section 6 demonstrates the advantages of my approach over existing methods using Monte Carlo experiments. I consider extensions to the benchmark model including a gross output production function in section 7. Section 8 presents my empirical results and section 9 concludes.

2 Data: The US Production Network

I first describe the data with which I characterize the firm-level production network within the United States, to highlight features that will be 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 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 for TFP estimation.[3]

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 20,687 unique buyer-supplier pairs and 70,259 dyad-year observations. I restrict the firm-level sample to the businesses that either report a customer or are reported as a customer. This results in an unbalanced panel of 9,196 firms and 58,653 firm-year observations. The dataset contains rich panel data on the US production network that enables me examine spillovers in a dynamic setting.[4]

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 18,000 employees and $5.8 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 2, manufacturers comprise more than half of the firms in the sample. Transportation and Utilities, and Services are the next largest sectors represented in the sample.

The observed network in the sample is sparse; that is, the number of connections per firm is low. Figure 1 shows that firms report 1 or 2 customers on average, while the same customers are reported by about 3 suppliers. Consistent with the 10% sales reporting requirement, reported customers tend to be quite large; the average customer realizes about eight times as much in sales as the average supplier in the data (see figure 5). This is may due to two factors: relatively small firms are likely to have major customers and larger firms are more 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 suppliers' sales.

In figure 4, I examine features of the network that affect the identification of spillovers within my

[3] See section D in the appendix for further details on variable construction.

[4] 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 Atalay et al. (2011) for graciously sharing their matched buyer-supplier data with me.

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. Interestingly, the network gets sparser at the beginning of the sample and gets denser after the mid-90's. At the same time, the network transitivity, the number of observed triads as a share of all possible triads trends upwards throughout the sample, but does not exceed 1.2%. In sections 3 and 4, I will discuss the importance of density and transitivity for both the biases in input elasticities from standard approaches, and the performance of my proposed estimator.

The network is primarily dominated by a large cluster of firms connected to each other. 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 while 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. Table 3 reports the 10 most central firms as measured by the number of links a firm has as a share of all observed links. In 1977, large manufacturers dominated the list and continued to do so in 1997 although their centrality declined. By 2016, however, retailers and wholesalers had become the most connected firms, with Walmart topping the list.

Figure 6 shows the relationship between a firm's labor productivity, as measured by log(sales per employee) and that of its average buyer or seller. The slope of the fitted regression line is 0.44, indicating a strong positive correlation between the two quantities. Interpreting this relationship would require distinguishing 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 affecting and being affected by its partners, how can we pin down the magnitude of the effect? On the other hand, this relationship may driven by the sorting of firms; if more productive firms trade with each other, then this correlation is evidence of network formation rather than spillovers. Yet another possibility is that supply chains are a channel for the transmission of 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 between these channels, examining how they impact the measurement of TFP, and quantifying the direction and magnitude of productivity spillovers.

Table 1: Firm Characteristics

	1977-1986	1987-1996	1997-2006	2007-2016	Full Sample
Sales	3.2	3.49	6.08	10.12	5.84
	(12.01)	(14.25)	(21.65)	(28.56)	(20.72)
Sales per 1000 employees	0.15	0.23	0.42	0.87	0.43
	(0.62)	(0.45)	(3.62)	(7.75)	(4.35)
Value Added	0.9	0.98	1.78	3.3	1.77
	(2.79)	(3.02)	(5.39)	(8.25)	(5.5)
Capital stock	3.15	3.37	4.6	9.31	5.16
	(11.83)	(14.13)	(17.93)	(32.36)	(20.98)
Materials	2.46	2.73	4.44	6.84	4.2
	(10.87)	(13.07)	(18.22)	(22.51)	(17.16)
Employees (thousands)	15.23	13.39	17.93	25.86	18.14
	(49.12)	(42.73)	(59.96)	(82.8)	(60.98)
Large firm (employees $\geq$ 500)	0.64	0.6	0.63	0.72	0.65
Observations	10731	16226	17180	14516	58653

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

Table 2: Industry Composition

	1977-1986	1987-1996	1997-2006	2007-2016	Full Sample
Agriculture	0.8	0.6	0.5	1.1	0.7
Mining	9.6	5.8	4.2	7.1	6.3
Construction	2.8	3.0	3.2	4.2	3.3
Manufacturing	54.1	53.4	53.8	51.3	53.1
Transportation & Utilities	16.1	15.9	12.5	12.7	14.1
Wholesale Trade	1.6	1.7	1.5	1.8	1.6
Retail Trade	4.2	4.9	5.2	5.4	5.0
Finance, Insurance, & Real Estate	2.2	2.5	2.8	3.7	2.8
Services	8.6	12.3	16.3	12.6	12.9
Total	100	100	100	100	100

This table reports the distribution of firms in the sample by Primary SIC.

Table 3: Top 10 firms by degree centrality

	1977			1997			2016		
Rank	Company	Centrality	Industry	Company	Centrality	Industry	Company	Centrality	Industry
1	General Motors	0.08	Mfg	Walmart	0.057	Retail	Walmart	0.076	Retail
2	Sears Roebuck	0.065	Retail	General Motors	0.035	Mfg	Cardinal Health	0.031	Wholesale
3	Ford Motor	0.04	Mfg	Ford Motor	0.032	Mfg	Mckesson	0.03	Wholesale
4	JCPenney	0.037	Retail	IBM	0.029	Svcs	AmerisourceBergen	0.026	Wholesale
5	Sears	0.035	Retail	AT&T	0.029	Trans & Util	Ford Motor	0.023	Mfg
6	Chrysler	0.024	Mfg	Chrysler	0.023	Mfg	General Motors	0.021	Mfg
7	IBM	0.024	Svcs	Ingram Micro	0.021	Trans & Util	Akamai Technologies	0.02	Svcs
8	AT&T Technologies	0.022	Mfg	Motorola Solutions	0.017	Mfg	Home Depot	0.02	Wholesale
9	General Electric	0.018	Svcs	Sears	0.017	Retail	AT&T	0.017	Trans & Util
10	McDonnell Douglas	0.013	Mfg	HP	0.016	Mfg	DHI Group	0.016	Svcs

This table reports the top 10 firms ranked by centrality in the indicated years.

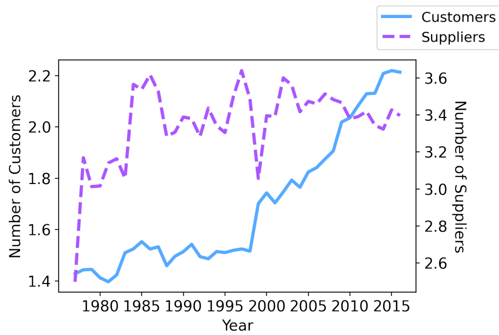


Figure 1: Average Firm Degree
This figure shows annual average out- and in-degrees (customers and suppliers) for firms in the sample.

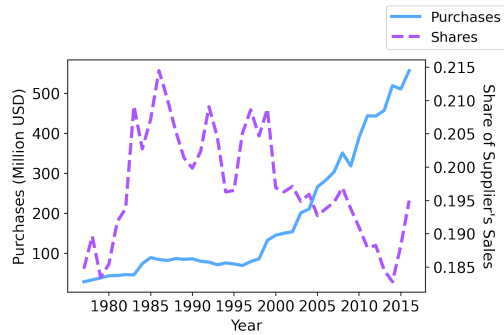


Figure 2: Value Traded in Relationships
This figure shows the annual average value traded by each buyer-supplier pair in nominal Million USD and as share of the each seller's total sales.

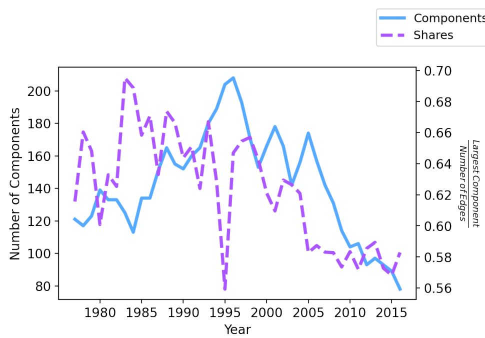


Figure 3: Network Edges and Components
This figure shows the number of buyer-supplier links (edges) and connected components in the network sample over time.

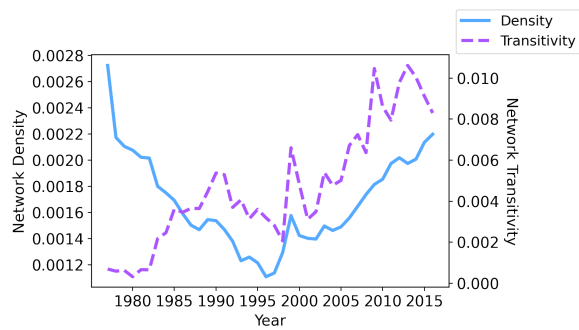


Figure 4: Network Density and Transitivity
This figure shows the density and transitivity of the network sample over time.

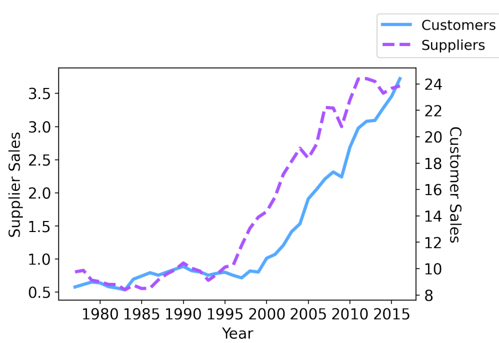


Figure 5: Customer and Supplier Sales
This figure shows annual average sales (in 2009 Billion USD) of firms reporting and reported as customers in the sample.

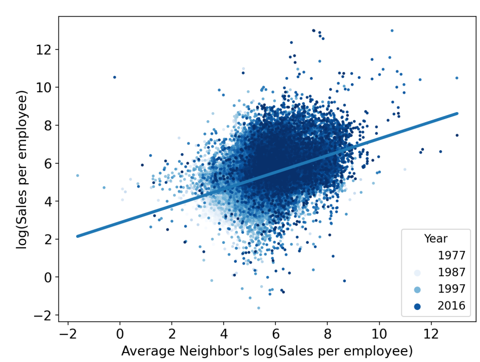


Figure 6: Assortativity of Labor Productivity
This figure shows the relationship between the labor productivities of a firm and its buyers and suppliers. The slope of the fitted regression line is 0.44.

3 Empirical Framework

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

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

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 $\mathcal{I}_t$. $e^{\omega_{it}}$ 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 i.i.d conditional on the firm's information set at the beginning of the period $\mathcal{I}_{t-1}$. 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 $E[\zeta_{it} \mid \mathcal{I}_{t-1}] = 0$. 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. Since firm j's TFP is also determined by its past productivity ω_{jt}, this representation indirectly allows for spillovers from productivity to occur with a one-period lag, but also accommodates the possibility that firm i is also affected by random shocks to j's productivity, ζ_{jt} within the same period. In addition, it enables researchers to differentiate between direct effects of firm activities

When researchers estimate TFP under the assumption in (3) whereas the true process is represented by (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, $\boldsymbol{x_{jt-1}} \in \mathcal{I}_{t-1}$. However, ω_{jt} poses a more serious problem because it is jointly realized with ω_{it} and cannot therefore be assumed to be in $\mathcal{I}_{t-1}$. 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 Cobb-Douglas production function as in Ackerberg et al. (2015):[5]

$$y_{it} = \alpha_1 + \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it} \tag{5}$$

where y_{it}, k_{it}, and ℓ_{it} are the logs of value-added[6], 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:

$$y_{it} = \alpha_1 + \alpha_\ell \ell_{it} + \alpha_k k_{it} + \mathbb{M}^{-1}(k_{it}, \ell_{it}, m_{it}) + \varepsilon_{it} \tag{6}$$

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

$$E[y_{it}|\mathcal{I}_{it}] = \varphi_{it} = \alpha_1 + \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} \tag{7}$$

This disentangles productivity from the idiosyncratic shock ε_{it}. In order to identify capital and

[5] 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.

[6] Output minus intermediate inputs.

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:

$$\begin{aligned} \omega_{it-1} &= \varphi_{it-1} - \alpha_1 - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1} \\ \implies \omega_{it} &= h(\varphi_{it-1} - \alpha_1 - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1}) + \eta_{it} \end{aligned}$$

Substituting into the production function yields:

$$y_{it} = \alpha_1 + \alpha_\ell \ell_{it} + \alpha_k k_{it} + h(\varphi_{it-1} - \alpha_1 - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1}) + \eta_{it} + \varepsilon_{it}$$

Since $E[\varepsilon_{it}|\mathcal{I}_{it}] = 0$ and $E[\eta_{it}|\mathcal{I}_{it-1}] = 0$ by assumption, then we can identify $\alpha_1, \alpha_k, \alpha_\ell$ based on the moment restriction:

$$E[\varepsilon_{it} + \eta_{it}|\mathcal{I}_{it-1}] = E[y_{it} - \alpha_1 - \alpha_k k_{it} - \alpha_\ell \ell_{it} - h(\varphi_{it-1} - \alpha_1 - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1})|\mathcal{I}_{it-1}] = 0 \tag{9}$$

Using this equation, we can derive moment conditions to estimate the elasticities. Since, there are four unknowns, $(\alpha_1, \alpha_k, \alpha_\ell, h(\cdot))$, a typical set of moments would be:

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

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.[7] Suppose we are interested in how upstream firms are affected by

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

the productivity of their downstream network. Let N_{it} be the set of i's customers in period t and $n_{it} = |N_{it}|$.[8] We would like to estimate the following network effects equation:

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

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 G_t as the row-normalized form of A_t.[9] 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) prove that (12) is identified if the identity matrix I, G and G^2 are linearly independent. The presence of intransitive triads[10] guarantees that linear independence holds. Production networks naturally have this structure because supply-chains tend to be unidirectional.

[8] 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.

[9] $G_{ij,t} = 1/n_i$ if $A_{ij,t} = 1$ and zero otherwise.

[10] 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.

Therefore, if ω_t was observed, one could estimate (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 i.i.d. 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 (3) but the true process is given by equation (12). This implies:[11]

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

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) show that the productivity process can be modified to account for its impact, as long as $\boldsymbol{x_{t-1}}$ is in the firm's information set at the beginning of the period.[12] Contextual effects can be accounted for in the same way under similar assumptions. Provided that network formation is exogenous, including $G_t\boldsymbol{x_{t-1}}$ in equation (3) would eliminate bias from this dimension.

$G_t E[\omega_t|\mathcal{I}_{t-1}]$ 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 $G_t\omega_t$:

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

where $\circ$ is the Hadamard product.[13] 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,

[11] 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.

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

[13] Element-wise multiplication.

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. Since the common component shocks are idiosyncratic each period, then k_t and ℓ_{t-1}, which were determined in the previous period are 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 α_k and α_ℓ estimates.

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

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

Then the second stage of ACF is equivalent to estimating:[14]

$$\implies y_t = \alpha_\ell \ell_t + \alpha_k k_t + \rho \sum_{s=0}^{\infty} \lambda^s G_t^s \left(y_{t-1} - \alpha_\ell \ell_{t-1} - \alpha_k k_{t-1} - u_{t-1}\right) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \varepsilon_t$$

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:

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

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.

[14] See section A for derivation.

In the OP/LP case where the labor elasticity is consistently estimated in the first stage, the second stage is equivalent to estimating:

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

$$\tag{19}$$

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:

$$\begin{aligned} plim\ \widehat{\alpha}_k &= \frac{cov(\Delta k_t, \Delta \widetilde{y}_t)}{var(\Delta k_t)} && (20) \\ &= \alpha_k \left(1 - \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s k_{t-1})}{var(\Delta k_t)}\right) + \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s \widetilde{y}_{t-1})}{var(\Delta k_t)} && (21) \end{aligned}$$

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 attentuation. The network structure also plays a role: when long chains exist, $G_t^s k_{t-1} > 0$ even for high values of s. On the other extreme, consider a network in which firms are paired off, so that the longest chain has length 1. Then $G_t^s k_{t-1} = 0$ for all $s > 1$ and attentuation would be lower under this scenario.

On the other hand, there is another source of bias 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 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 variety of productivity processes and network effects. I propose a modification to the ACF procedure that achieves this without many 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:

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

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:

$$\omega_t = (I - \lambda G_t)^{-1}h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + (I - \lambda G_t)^{-1}\zeta_t \tag{23}$$

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

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

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

$$E[\omega_t|\mathcal{I}_{t-1}] = \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) \tag{25}$$

since 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 (24) also indicates how λ can be identified. Given the reduced-form equation, $G_t\omega_t$ can be written as:

$$G_t\omega_t = G_t h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \sum_{s=1}^{\infty} \lambda^s G_t^{s+1} h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \sum_{s=0}^{\infty} \lambda^s G_t^{s+1}\zeta_t \tag{26}$$

As long as 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 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 (26) indicates that there are additional instruments 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:

$$y_t = \alpha_1\iota + \alpha_k k_t + \alpha_\ell \ell_t + (1-\lambda G_t)^{-1}\left[h(\varphi_{t-1} - \alpha_1 - \alpha_k k_{t-1} - \alpha_\ell \ell_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \zeta_t\right] + \varepsilon_t \tag{27}$$

which leads to the polynomial expansion:

$$y_t = \alpha_1\iota + \alpha_k k_t + \alpha_\ell \ell_t + \sum_{s=0}^{\infty}\lambda^s G_t^s h(\varphi_{t-1} - \alpha_1 - \alpha_k k_{t-1} - \alpha_\ell \ell_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \sum_{s=0}^{\infty}\lambda^s G_t^s \zeta_t + \varepsilon_t \tag{28}$$

Accounting for network effects in the estimation procedure comes at the cost of additional assumptions. The first, as seen above, is that the endogenous effect enters the productivity process linearly. This would not hold if there were non-monoticities in spillovers. For instance, 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.

Secondly, 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 do 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:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda G_t \omega_t + c_{\psi_t} + \zeta_t \tag{29}$$

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 (29) yields:

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

with the corresponding reduced form:

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

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

$$J_t y_t = \alpha_1 J_t \iota + \alpha_k J_t k_t + \alpha_\ell J_t \ell_t + \sum_{s=0}^{\infty} \lambda^s J_t G_t^s h(\varphi_{it-1} - \alpha_1 - \alpha_k k_{it-1} - \alpha_\ell \ell_{it-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) \quad (30)$$

$$+ \sum_{s=0}^{\infty} \lambda^s J_t G_t^s \zeta_t + J_t \varepsilon_{it} \quad (31)$$

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): $y_t = \alpha_1 + \alpha_k k_t + \alpha_\ell \ell_t + \mathbb{M}^{-1}(k_t, \ell_t, m_t) + \varepsilon_t$, using a polynomial approximation.[15] This yields estimates $\widehat{\varphi}_t = y_t - \widehat{\varepsilon}_t$.

In the second stage, estimate equation (28) 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: $\alpha_1^*, \alpha_k^*, \alpha_\ell^*$ and compute $\omega_t^* = \widehat{\varphi}_t - \alpha_1^* - \alpha_k^* k_t - \alpha_\ell^* \ell_t$.[16] Estimate the productivity process by 2SLS:

$$\omega_t^* = h(\omega_{t-1}^*, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda G_t \omega_t^* + u_t \quad (32)$$

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 $(\alpha_1^*, \alpha_k^*, \alpha_\ell^*)$ that satisfy the sample moment conditions:

$$\begin{aligned} E_{nt}[u_{it}^* \iota] &= 0 \\ E_{nt}[u_{it}^* k_t] &= 0 \\ E_{nt}[u_{it}^* \ell_{t-1}] &= 0 \end{aligned} \quad (33)$$

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. See section E in the appendix for details on bootstrapping network data.

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.

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

[16] It is also possible to also concentrate out α_1 and instead compute $\omega_t^* + \alpha_1^* = \widehat{\varphi}_t - \alpha_k^* k_t - \alpha_\ell^* \ell_t$.

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.[17]

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 $V_i(A_{ij,t} = 1) - V_i(A_{ij,t} = 0) > 0$.[18] 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.

$$U_{ijt}(\gamma) = \gamma_1 + \boldsymbol{z_{it}\gamma_i} + \boldsymbol{z_{jt}\gamma_j} + \boldsymbol{z_{ijt}\gamma_{ij}} + \gamma_h H_{ijt} \tag{34}$$

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.

$\boldsymbol{z_{it}}$ 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 characteristics: $|\boldsymbol{z_{it}} - \boldsymbol{z_{jt}}|$ or some other dyad-specific measures, such as the physical distance between the firms, industry input-output shares, etc. A negative coefficient on

[17] Other studies such as Goldsmith-Pinkham and Imbens (2013) and Hsieh and Lee (2016) model network endogeneity as a correlation between unobserved variables in the network selection model and the error term of the outcome equation. The interpretation differs; in this setting selection would be driven by unobserved synergies such as common business philosophies. If these factors are also correlated with productivity, then estimated spillovers would capture the effect of assortativity in these unobserved characteristics (see Serpa and Krishnan (2018)) for an application to productivity spillovers. I choose the Arduini et al. (2015) model for two reasons. First, it is consistent with the notion of search ability being related to productivity, something that I think is more important to capture than unobserved synergies. Secondly, the reduction of the problem to a selection correction term preserves the usual structure of the estimator, while the Goldsmith-Pinkham and Imbens (2013) and Hsieh and Lee (2016) relies on Bayesian estimation of a full likelihood model.

[18] 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.

$|\boldsymbol{z_{it}} - \boldsymbol{z_{jt}}|$ 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 instance, $H_{ijt} = A_{ij,t-1}$ would mean that firm i only considers linkages from the previous period, while $H_{ijt} = \mathbb{1}\left(\sum_{s=1}^{m} A_{ij,t-s} > 0,\ m \leq t\right)$ measures whether i and j were connected in any of the last m periods.[19]

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; 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 $\Sigma_{\zeta\xi} = \begin{pmatrix} \sigma_\zeta^2 & \sigma_{\zeta\xi}^{'} \\ \sigma_{\zeta\xi} & \Sigma_\xi \end{pmatrix}$ is positive definite, σ_ζ^2 is a scalar, $\sigma_{\zeta\xi}$ is an $n_t - 1$ column vector of covariances, 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 $\boldsymbol{\delta} = \Sigma_\xi^{-1} \sigma_{\zeta\xi}$, ν_t is independent of ξ_{it} and $\sigma_\nu^2 = \sigma_\zeta^2 - \sigma_{\zeta\xi}^{'} \Sigma_\xi^{-1} \sigma_{\zeta\xi}$. Therefore, the productivity process becomes:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda G_t \omega_t + \Xi_t \boldsymbol{\delta} + \nu_t \tag{35}$$

[19]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.

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) shows that the selectivity bias can be controlled for using a Heckman-type mills ratio:

$$\begin{aligned}\mu_{it} &= \sum_{j\neq i}^{N_t} g_{ij,t} \frac{\phi(\Phi^{-1}(p))}{\Phi(\Phi^{-1}(p))} + (1 - g_{ij,t}) \frac{\phi(\Phi^{-1}(p))}{1 - \Phi(\Phi^{-1}(p))} \\ &= \sum_{j\neq i}^{N_t} g_{ij,t} \frac{\phi(\Phi^{-1}(p))}{p} + (1 - g_{ij,t}) \frac{\phi(\Phi^{-1}(p))}{1-p}\end{aligned} \tag{36}$$

where $p = P(A_{ij,t} = 1 \mid \mathcal{I}_{t-1})$, 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 $(\alpha_1^*, \alpha_k^*, \alpha_\ell^*)$, compute $\omega_{it-1}^* = \widehat{\varphi}_{it-1} - \alpha_1^* - \alpha_k^* k_{it-1} - \alpha_\ell^* \ell_{it-1}$.

Using ω_{it-1}^* and other variables that could determine the observed links between firms, estimate the selection model in equation (34) 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:

$$\omega_t^* = \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}^*, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \delta \sum_{s=0}^{\infty} \lambda^s G_t^s \mu_{it}^* + u_t \tag{37}$$

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 (32) for identification of the elasticities.[20]

[20] In principle, the selection model would be re-estimated for each value of ω_{it-1}^* as the values $(\alpha_1^*, \alpha_k^*, \alpha_\ell^*)$ are updated in each iteration.However, this significantly increases the computational cost of the procedure. As long as 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.

6 Monte Carlo Experiments

I conduct three 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 examine how each type of network effect individually affects the bias and efficiency of capital and labor elasticity estimates obtained using the ACF procedure. Next, I demonstrate how my modified procedure performs when endogenous, contextual and correlated effects are cumulatively present and consider the sensitivity of the estimates from my benchmark procedure to misspecification. Finally, I compare the performance of ACF and my benchmark procedure as the size of the endogenous effect, the persistence of productivity and the density of the network vary.

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

$$y_{it} = \alpha_1 + \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it}$$

where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. I set $\alpha_1 = 0, \alpha_\ell = 0.6, \alpha_k = 0.4$ and $\sigma_\varepsilon^2 = 1$.[21] 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. 2 sells to 50 percent each to 3 and 4, while industry 3 sells all its output to 4. The fourth industry 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.

6.1 Experiment 1: Bias in Standard ACF estimates due to Network Effects

I simulate five data generating processes (DGPs) to demonstrate the bias in standard ACF estimates of the input elasticities from each type network effect—endogenous, contextual and correlated—and network endogeneity separately.

The productivity process is:

$$\omega_t = \beta_1 \iota + \rho \omega_{t-1} + \beta_x x_t + \lambda G_t \omega_t + \beta_{\bar{x}} G_t x_t + c_{\psi_t} + \zeta_t \tag{38}$$

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

[21] See section B for further details on the Monte Carlo setup.

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 ACF estimates should be consistent. DGP 2 features only the endogenous effect ($\lambda = 0.3, \beta_{\bar{x}} = 0, c_{\psi_t} = 0$) while DGP 3 features only the contextual effect ($\lambda = 0, \beta_{\bar{x}} = 0.3, c_{\psi_t} = 0$). DGP 2 features only the endogenous effect ($\lambda = 0.3, \beta_{\bar{x}} = 0, c_{\psi_t} = 0$) while DGP 3 features only the contextual effect ($\lambda = 0, \beta_{\bar{x}} = 0.3, c_{\psi_t} = 0$). 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 ($\lambda = 0, \beta_{\bar{x}} = 0, c_{\psi_t}\mathcal{N}(1, 1)$). 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 $\delta = \frac{\sigma_{\zeta\xi}}{\sigma_\xi^2} = 0.003$, while there are no other network effects ($\lambda = 0, \beta_{\bar{x}} = 0, c_{\psi_t} = 0$).

I estimate the production function using standard ACF with a second-degree polynomial approximation in the first and second stages. The results are shown in table 4. The largest bias comes from the presence of an endogenous effect. It leads to a capital coefficient estimate that is almost 25% higher than the true value. In comparison, a contextual effect of the same size has a negligible impact on the capital coefficient. As expected, correlated effects reduce precision but do not have a sizable impact on bias. In the absence of any other network effects, endogenous network formation has no impact on bias or efficiency of the estimated input elasticities. Therefore, of all the network effects, ignoring endogenous spillovers introduces the greatest bias in the production function elasticities.

Table 4: Bias due to Network Effects with Standard ACF Procedure

DGP		α_ℓ	α_k
	True values	0.6	0.4
No network effects	Mean	0.599	0.4
	Std. Dev.	(0.025)	(0.061)
Endogenous Effect ($\lambda = 0.3$)	Mean	0.596	0.495
	Std. Dev.	(0.032)	(0.066)
Contextual Effect ($\beta_{\bar{x}} = 0.3$)	Mean	0.598	0.414
	Std. Dev.	(0.04)	(0.063)
Correlated Effect ($c_{\psi_t} \sim \mathcal{N}(1, 1)$)	Mean	0.58	0.38
	Std. Dev.	(0.171)	(0.271)
Endogenous Network $\left(\frac{\sigma_{\zeta\xi}}{\sigma_\xi^2} = 0.003\right)$	Mean	0.6	0.391
	Std. Dev.	(0.014)	(0.082)

Based on 1000 replications.This table reports production function elasticites obtained using the procedure in Ackerberg et al. (2015). Each row includes a separate network effect in the law of motion on productivity.

6.2 Experiment 2: Comparison of Estimates from Standard and Modified ACF Procedures

Next, I compare the performance my estimator against standard ACF in table 5 using four DGPs. The productivity process is:

$$\omega_t = \beta_1 \iota + \rho \omega_{t-1} + \beta_x x_t + \lambda G_t \omega_t + \beta_{\bar{x}} G_t x_t + c_{\psi_t} + \zeta_t \tag{39}$$

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 $\sigma_{\widetilde{x}}^2 = 5$.

In this experiment, however, I introduce network effects cumulatively rather than individually. DGP 1 favors the ACF procedure with no network effects and exogenous network formation, while DGP 2 introduces both endogenous and contextual network effects. DGP 3 is similar to the second DGP but with the addition of correlated effects, while DGP 4 has all the previous network effects with endogenous network formation.

I consider 4 estimators. The first is a standard ACF that assumes no network effects. 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.

Under DGP 1, all estimators perform well when estimating both the production 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 introduce spurious 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. In these respects, ACF performs better because when correlated effects are unaccounted for, all network terms containing G_t introduce bias because they are correlated 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 with ACF-ND, estimates of the productivity process and input elasticities are not different from when selection is accounted for with ACF-NDS.

Table 5: Comparison of Estimates from Standard ACF and Modified ACF Procedures

DGP	Estimator		Elasticities		Productivity Process Coefficients					
			α_ℓ	α_k	β_1	ρ	β_x	$\beta_{\bar{x}}$	λ	$\frac{\sigma_{\zeta\xi}}{\sigma_\xi^2}$
DGP 1		True values	0.6	0.4	0.5	0.6	0.4	0.0	0.0	0.0
	ACF	Mean	0.599	0.4	0.905	0.6	0.401	0.	-0.001	
		Std. Dev.	(0.025)	(0.061)	(0.182)	(0.015)	(0.026)	(0.009)	(0.01)	
	ACF-N	Mean	0.602	0.392	0.916	0.601	0.398	0.	-0.001	
		Std. Dev.	(0.018)	(0.061)	(0.185)	(0.016)	(0.019)	(0.009)	(0.01)	
	ACF-ND	Mean	0.603	0.389	0.92	0.601	0.397	-0.	-0.	
		Std. Dev.	(0.024)	(0.064)	(0.19)	(0.016)	(0.024)	(0.01)	(0.011)	
	ACF-NDS	Mean	0.603	0.39	0.919	0.601	0.397	-0.	0.	-0.
		Std. Dev.	(0.024)	(0.064)	(0.189)	(0.016)	(0.025)	(0.01)	(0.012)	(0.002)
DGP 2		True values	0.6	0.4	0.5	0.6	0.4	0.1	0.3	0.0
	ACF	Mean	0.595	0.516	0.507	0.556	0.402	0.092	0.332	
		Std. Dev.	(0.035)	(0.07)	(0.127)	(0.017)	(0.035)	(0.016)	(0.042)	
	ACF-N	Mean	0.601	0.401	0.82	0.596	0.399	0.121	0.242	
		Std. Dev.	(0.018)	(0.046)	(0.113)	(0.016)	(0.018)	(0.013)	(0.026)	
	ACF-ND	Mean	0.602	0.398	0.921	0.595	0.397	0.118	0.249	
		Std. Dev.	(0.028)	(0.055)	(0.141)	(0.016)	(0.028)	(0.014)	(0.026)	
	ACF-NDS	Mean	0.602	0.396	0.921	0.596	0.397	0.115	0.257	-0.004
		Std. Dev.	(0.027)	(0.055)	(0.142)	(0.016)	(0.028)	(0.014)	(0.026)	(0.002)
DGP 3		True values	0.6	0.4	0.5	0.6	0.4	0.1	0.3	0.0
	ACF	Mean	0.616	0.496	1.21	0.479	0.362	0.121	0.357	
		Std. Dev.	(0.169)	(0.417)	(1.207)	(0.171)	(0.161)	(0.102)	(0.496)	
	ACF-N	Mean	0.741	0.162	1.587	0.514	0.257	0.082	0.222	
		Std. Dev.	(0.154)	(0.215)	(0.851)	(0.269)	(0.154)	(0.072)	(0.62)	
	ACF-ND	Mean	0.614	0.368	1.974	0.605	0.385	0.109	0.266	
		Std. Dev.	(0.032)	(0.052)	(0.157)	(0.017)	(0.032)	(0.012)	(0.018)	
	ACF-NDS	Mean	0.614	0.368	1.966	0.605	0.385	0.108	0.269	-0.002
		Std. Dev.	(0.032)	(0.052)	(0.157)	(0.018)	(0.032)	(0.012)	(0.018)	(0.002)
DGP 4		True values	0.6	0.4	0.5	0.6	0.4	0.1	0.3	0.003
	ACF	Mean	0.607	0.35	1.707	0.603	0.374	0.128	0.255	
		Std. Dev.	(0.138)	(0.239)	(1.156)	(0.147)	(0.166)	(0.109)	(0.122)	
	ACF-N	Mean	0.705	0.183	1.508	0.637	0.291	0.067	0.236	
		Std. Dev.	(0.137)	(0.213)	(0.756)	(0.184)	(0.142)	(0.056)	(0.2)	
	ACF-ND	Mean	0.619	0.368	1.794	0.61	0.383	0.091	0.281	
		Std. Dev.	(0.073)	(0.116)	(0.276)	(0.056)	(0.07)	(0.023)	(0.037)	
	ACF-NDS	Mean	0.621	0.362	1.78	0.612	0.38	0.09	0.28	0.001
		Std. Dev.	(0.078)	(0.129)	(0.297)	(0.064)	(0.076)	(0.026)	(0.037)	(0.002)

Based on 1000 replications. Estimators are based on Ackerberg et al. (2015) with ACF denoting the standard procedure, while N, D, and S indicate modifications to account for network effects, network differencing, and network selection respectively. Data generating processes are outlined above (see appendix B for details). DGP1 has no network effects, DGP2 has correlated and endogenous effects, DGP3 includes correlated, endogenous and network fixed effects, while DGP4 features all 3 network effects and an endogenous network formation process.

6.3 Experiment 3: Effect of Network Density on Bias and Precision

Since the first experiment shows that most important source of bias is the endogenous effect, 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:

$$\omega_t = \beta_1 + \rho_1\omega_{t-1} + \rho_2\omega_{t-1}^2 + \lambda G_t\omega_t + \zeta_t \tag{40}$$

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 simulated values of productivity grow quite large for some firms, and the resulting investment series soon tends to infinity for highly productivity firms.[22]. The quadratic term serves as a dampener to control the size ω_t in the simulation.[23] Additionally, it allows for the comparison of ACF and my modified procedure when the productivity is process not linear.

To vary network density, I draw random exogenous networks using Erdős and Rényi (1960) graphs, also known as binomial graphs. Firms are edges are formed $A_{ijt} \overset{i.i.d.}{\sim} \text{Bern}(p)$ and the density of the graph is equal to the probability of an 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 to follow a Pareto (power-law) degree distribution (Bernard and Moxnes, 2018).

Table 6 shows the results of varying network density. 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 fewer triads in the network.

In section C in the appendix, I also examine how my approach performs as the persistence of productivity and the size of the endogenous network effect vary. Unsurprisingly, my procedure yields more consistent and precise estimates as productivity gets more persistent, increasing the relevance of $G_t\omega_{t-1}$ as an instrument, and as the endogenous spillover gets larger, raising the relevance of the $G_t^2\omega_{t-1}$ instrument. Importantly, even when $\lambda = 0$, my procedure still yields consistent estimates as long as ρ is sufficiently large.

[22] See details on optimal investment in section B.5 in the appendix

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

Table 6: Effect of Sparsity on Bias and Precision (Quadratic AR1)

Density	Estimator	Elasticities		Productivity Process Coefficients			
		α_ℓ	α_k	β_1	ρ_1	ρ_2	λ
		0.6	0.4	0.5	0.8	-0.01	0.3
0.0001	ACF	0.603	0.358	-0.125	0.809	-0.01	0.087
		(0.024)	(0.239)	(2.369)	(0.216)	(0.003)	(0.106)
	ACF-N	0.617	0.413	-0.23	0.76	-0.01	0.226
		(0.057)	(0.165)	(2.845)	(0.196)	(0.022)	(0.109)
0.0003	ACF	0.604	0.359	0.122	0.81	-0.01	0.113
		(0.024)	(0.216)	(1.975)	(0.19)	(0.003)	(0.12)
	ACF-N	0.632	0.381	0.379	0.764	-0.011	0.241
		(0.093)	(0.113)	(1.456)	(0.195)	(0.038)	(0.238)
0.0005	ACF	0.605	0.377	0.209	0.798	-0.01	0.132
		(0.024)	(0.195)	(1.691)	(0.169)	(0.003)	(0.126)
	ACF-N	0.641	0.371	0.412	0.753	-0.009	0.25
		(0.106)	(0.113)	(1.226)	(0.217)	(0.034)	(0.097)
0.0007	ACF	0.606	0.387	0.271	0.791	-0.01	0.159
		(0.027)	(0.182)	(1.509)	(0.158)	(0.003)	(0.137)
	ACF-N	0.646	0.362	0.51	0.745	-0.007	0.243
		(0.116)	(0.116)	(0.506)	(0.237)	(0.046)	(0.1)
0.0009	ACF	0.606	0.411	0.266	0.771	-0.01	0.18
		(0.03)	(0.168)	(1.359)	(0.147)	(0.004)	(0.145)
	ACF-N	0.635	0.371	0.532	0.767	-0.011	0.252
		(0.101)	(0.098)	(0.308)	(0.2)	(0.037)	(0.085)
0.001	ACF	0.606	0.423	0.243	0.761	-0.01	0.196
		(0.032)	(0.161)	(1.295)	(0.147)	(0.007)	(0.152)
	ACF-N	0.63	0.377	0.539	0.778	-0.009	0.225
		(0.09)	(0.088)	(0.305)	(0.203)	(0.031)	(1.04)
0.003	ACF	0.602	0.617	-0.343	0.523	-0.017	0.261
		(0.053)	(0.114)	(1.261)	(0.137)	(0.011)	(0.184)
	ACF-N	0.611	0.389	0.585	0.815	-0.01	0.283
		(0.038)	(0.049)	(0.256)	(0.06)	(0.009)	(0.035)
0.005	ACF	0.605	0.637	-0.109	0.462	-0.019	0.312
		(0.067)	(0.138)	(0.987)	(0.161)	(0.017)	(0.201)
	ACF-N	0.608	0.388	0.509	0.818	-0.01	0.291
		(0.03)	(0.057)	(0.298)	(0.05)	(0.007)	(0.042)
0.007	ACF	0.606	0.639	-0.405	0.47	-0.026	0.545
		(0.073)	(0.149)	(13.879)	(0.694)	(0.207)	(6.696)
	ACF-N	0.607	0.385	0.437	0.818	-0.01	0.294
		(0.027)	(0.068)	(0.339)	(0.053)	(0.002)	(0.027)
0.009	ACF	0.606	0.638	0.132	0.448	-0.018	0.339
		(0.073)	(0.154)	(2.354)	(0.217)	(0.05)	(1.188)
	ACF-N	0.606	0.386	0.405	0.815	-0.01	0.302
		(0.03)	(0.077)	(0.393)	(0.06)	(0.005)	(0.175)
0.01	ACF	0.606	0.639	0.011	0.452	-0.019	0.388
		(0.072)	(0.153)	(1.8)	(0.2)	(0.046)	(0.925)
	ACF-N	0.606	0.386	0.404	0.815	-0.01	0.301
		(0.031)	(0.078)	(0.404)	(0.061)	(0.005)	(0.14)
0.03	ACF	0.605	0.644	-0.23	0.478	-0.011	0.317
		(0.069)	(0.149)	(5.307)	(1.216)	(0.238)	(2.909)
	ACF-N	0.605	0.387	0.417	0.813	-0.01	0.298
		(0.032)	(0.083)	(0.398)	(0.064)	(0.004)	(0.061)
0.05	ACF	0.606	0.643	0.054	0.414	-0.024	0.446
		(0.072)	(0.152)	(3.357)	(0.782)	(0.151)	(1.833)
	ACF-N	0.605	0.388	0.417	0.813	-0.01	0.299
		(0.032)	(0.084)	(0.399)	(0.065)	(0.004)	(0.062)
0.07	ACF	0.606	0.643	0.005	0.423	-0.022	0.419
		(0.073)	(0.154)	(1.994)	(0.457)	(0.085)	(1.053)
	ACF-N	0.604	0.388	0.42	0.813	-0.01	0.297
		(0.03)	(0.084)	(0.4)	(0.064)	(0.003)	(0.049)
0.09	ACF	0.606	0.643	0.001	0.426	-0.021	0.413
		(0.071)	(0.154)	(1.774)	(0.406)	(0.074)	(0.928)
	ACF-N	0.604	0.388	0.42	0.813	-0.01	0.297
		(0.03)	(0.084)	(0.401)	(0.063)	(0.003)	(0.049)
0.1	ACF	0.606	0.642	-0.003	0.425	-0.021	0.417
		(0.073)	(0.157)	(1.812)	(0.417)	(0.077)	(0.952)
	ACF-N	0.604	0.388	0.42	0.814	-0.01	0.296
		(0.028)	(0.083)	(0.403)	(0.061)	(0.003)	(0.032)
0.3	ACF	0.605	0.644	-0.048	0.437	-0.019	0.388
		(0.07)	(0.15)	(1.033)	(0.184)	(0.025)	(0.454)
	ACF-N	0.603	0.389	0.422	0.813	-0.01	0.296
		(0.027)	(0.084)	(0.409)	(0.062)	(0.003)	(0.032)
0.5	ACF	0.606	0.644	-0.054	0.435	-0.019	0.392
		(0.072)	(0.153)	(1.048)	(0.189)	(0.027)	(0.468)
	ACF-N	0.604	0.389	0.421	0.813	-0.01	0.296
		(0.027)	(0.083)	(0.412)	(0.062)	(0.003)	(0.032)
0.7	ACF	0.607	0.643	-0.025	0.427	-0.021	0.416
		(0.074)	(0.156)	(1.63)	(0.365)	(0.066)	(0.827)
	ACF-N	0.604	0.388	0.42	0.814	-0.01	0.296
		(0.027)	(0.084)	(0.413)	(0.062)	(0.003)	(0.032)
0.9	ACF	0.607	0.642	-0.054	0.435	-0.019	0.392
		(0.074)	(0.156)	(1.063)	(0.189)	(0.027)	(0.478)
	ACF-N	0.604	0.388	0.422	0.813	-0.01	0.297
		(0.029)	(0.085)	(0.414)	(0.063)	(0.004)	(0.052)

Based on 1000 replications. Estimators are based on Ackerberg et al. (2015) with ACF denoting the standard procedure and ACF-N indicating the modified procedure to account for network effects. Networks are exogenous erdos-renyi (binomial) graphs with densities as shown. The data-generating process for productivity is quadratic AR1 with endogenous network effects. True values of the parameters are at the top of the table. Standard deviations are in parentheses.

7 Extensions

7.1 Gross 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 the production function:

$$\begin{aligned} Y_t &= F(L_t, K_t, M_t)e^{\omega_t+\varepsilon_t} \\ \iff y_t &= f(\ell_t, k_t, m_t) + \omega_t + \varepsilon_t \end{aligned} \tag{41}$$

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:

$$\max_{M_t} P_t E[F(L_t, K_t, M_t)e^{\omega_t+\varepsilon_t}] - P_t^M M_t \tag{42}$$

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:

$$P_t \frac{\partial}{\partial M_t} F(L_t, K_t, M_t)e^{\omega_t}\mathcal{E} = P_t^M \tag{43}$$

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

$$\begin{aligned} &\frac{\partial}{\partial M_t} F(L_t, K_t, M_t)e^{\omega_t}\mathcal{E} = \frac{P_t^M}{P_t} \\ &\frac{M_t}{Y_t}\frac{\partial}{\partial M_t} F(L_t, K_t, M_t)e^{\omega_t}\mathcal{E} = \frac{P_t^M M_t}{P_t Y_t} \\ &\ln\left(\frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t)\right) - \varepsilon_t + \ln(\mathcal{E}) = s_t \end{aligned} \tag{44}$$

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

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.

$$E[\varepsilon_t \mid \mathcal{I}_t] = 0 \implies E[s_t \mid \mathcal{I}_t] = \ln\left(\frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t)\right) + \ln(\mathcal{E}) \tag{45}$$

Let $D^{\mathcal{E}}(\ell_t, k_t, m_t) \equiv \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) \times \mathcal{E}$. Then given the moment of ε_t in (45) 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:

$$\begin{aligned} \varepsilon_t = \ln D^{\mathcal{E}}(\ell_t, k_t, m_t) - s_t \implies e^{\varepsilon_t} &= D^{\mathcal{E}}(\ell_t, k_t, m_t) e^{-s_t} \\ \mathcal{E} = E[e^{\varepsilon_t}] &= E[D^{\mathcal{E}}(\ell_t, k_t, m_t) e^{-s_t}] \end{aligned} \tag{46}$$

Using the estimates of $D^{\mathcal{E}}$ from the share regression, we can replace the moment in (46) with its empirical equivalent and compute the constant $\mathcal{E}$. This enables us obtain an estimate of the materials elasticity:

$$D(\ell_t, k_t, m_t) = \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) = \frac{D^{\mathcal{E}}(\ell_t, k_t, m_t)}{\mathcal{E}} \tag{47}$$

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:

$$\int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t = f(\ell_t, k_t, m_t) + \mathcal{C}(\ell_t, k_t) \tag{48}$$

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 (41):

$$\begin{aligned} \int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t &= y_t - \omega_t - \varepsilon_t + \mathcal{C}(\ell_t, k_t) \\ \mathcal{Y}_t \equiv y_t - \int \frac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) dm_t - \varepsilon_t &= -\mathcal{C}(\ell_t, k_t) + \omega_t \end{aligned} \tag{49}$$

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:

$$\omega_t = h(\omega_{t-1}) + \eta_t, \qquad \text{where } E[\eta_t \mid \mathcal{I}_{t-1}] = 0 \tag{50}$$

$$\begin{aligned} \omega_{t-1} &= \mathcal{Y}_{t-1} + \mathcal{C}(\ell_{t-1}, k_{t-1}) \\ \implies \mathcal{Y}_t &= -\mathcal{C}(\ell_t, k_t) + h(\mathcal{Y}_{t-1} + \mathcal{C}(\ell_{t-1}, k_{t-1})) + \eta_t \end{aligned} \tag{51}$$

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

$$\begin{aligned} E[\eta_t \ell_t^{\tau_\ell} k_t^{\tau_k}] &= 0 \\ E[\eta_t \mathcal{Y}_{t-1}^{\tau_\mathcal{Y}}] &= 0 \end{aligned} \tag{52}$$

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.

7.1.1 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 (50) 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 (51) becomes:

$$\mathcal{Y}_t = -\mathcal{C}(\ell_t, k_t) + \sum_{s=0}^{\infty} \lambda^s G_t^s h(\mathcal{Y}_{t-1} + \mathcal{C}(\ell_{t-1}, k_{t-1})) + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t \tag{53}$$

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

$$E[\zeta_t G_t^s \mathcal{Y}_{t-1}^{\tau_\mathcal{Y}}] = 0 \qquad \text{where } s \geq 1 \tag{54}$$

7.2 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.

[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.

7.2.1 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:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, A_t \boldsymbol{x_{t-1}}) + \lambda A_t \omega_t + \zeta_t \tag{55}$$

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:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda A_t \omega_t + \zeta_t \tag{56}$$

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

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \lambda_A A_t \omega_t + \lambda_G G_t \omega_t + \zeta_t \tag{57}$$

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.

7.2.2 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:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, \{G_{b,t} \boldsymbol{x_{t-1}}\}_{b=1}^{B}) + \sum_{b=1}^{B} \lambda_b G_{b,t} \omega_t + \zeta_t \tag{58}$$

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.

8 Results

In this section, I apply my empirical framework to examining the magnitude and direction of endogenous vertical productivity spillovers 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. I also consider a value-added Cobb Douglas 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 obtain 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, 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. After estimating TFP, I obtain network effect coefficients by applying 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 apply global differencing described in section 4.2 to both standard and modified procedures.

8.1 Production Function Elasticites

Tables 7 and 8 report the estimated production function elasticities. GNR/ACF refer to the standard procedures, GNR-N/ACF-N denote my modified approach that accounts for endogenous productivity spillovers, and GNR-ND/ACF-ND indicate specifications that account for both endogenous network effects and component-year fixed effects. Because I assume that the network does not affect intermediate input demand in the gross output specification, the elasticity of output to materials does not vary across specifications.

Estimated capital and labor elasticities are also quite similar with and without accounting for network effects. The relative importance of each input varies over time; in the gross output specification, the estimated capital elasticity rises from 0.17 between 1977-1986 to 0.22 in the 2007-2016 period, while labor elasticity falls from 0.36 to 0.22 in the same interval. Surprisingly, capital and labor elasticities from the value-added specification move in the opposite direction, with the former decreasing from 0.42 to 0.32 and the latter increasing from 0.6 to 0.66.

Table 7: Gross Production Function Elasticities

Period	Estimator	Capital	Labor	Materials
1977-1986	GNR	0.161	0.355	0.506
	GNR-N	0.158	0.357	0.506
	GNR-ND	0.169	0.351	0.506
1987-1996	GNR	0.192	0.243	0.576
	GNR-N	0.167	0.260	0.575
	GNR-ND	0.191	0.244	0.575
1997-2006	GNR	0.195	0.234	0.569
	GNR-N	0.195	0.233	0.569
	GNR-ND	0.193	0.240	0.569
2007-2016	GNR	0.229	0.211	0.550
	GNR-N	0.215	0.223	0.550
	GNR-ND	0.217	0.217	0.550
All	GNR	0.182	0.295	0.545
	GNR-N	0.183	0.295	0.545
	GNR-ND	0.185	0.297	0.545

This table reports the average input elasticities from a gross output production function estimated on US firms in Compustat. Estimators are based on Gandhi et al. (2020): GNR denotes the standard procedure with a linear first stage, a second-degree polynomial in the second stage, and a linear productivityprocess. GNR-N and GNR-ND are modifications to accommodate network effects and network differencing respectively. All specifications include industry and year fixed effects in the productivity process.

Table 8: Value-Added Production Function Elasticities

Period	Estimator	Capital	Labor
1977-1986	ACF	0.401	0.618
	ACF-N	0.406	0.612
	ACF-ND	0.421	0.602
1987-1996	ACF	0.441	0.596
	ACF-N	0.437	0.606
	ACF-ND	0.446	0.592
1997-2006	ACF	0.369	0.651
	ACF-N	0.369	0.652
	ACF-ND	0.366	0.658
2007-2016	ACF	0.330	0.650
	ACF-N	0.330	0.650
	ACF-ND	0.323	0.661
All	ACF	0.380	0.635
	ACF-N	0.380	0.637
	ACF-ND	0.376	0.642

This table reports input elasticities of a Cobb-Douglas value-added production function (in logs) estimated on US firms in Compustat. Estimators are based on Ackerberg et al. (2015): ACF denotes the standard procedure with a second-degree polynomial in the first stage and a linear productivity evolution process. ACF-N and ACF-ND are modifications to accommodate network effects and network differencing respectively. All specifications include industry and year fixed effects in the productivity process.

8.2 Endogenous Productivity Spillovers

I now turn to estimates of productivity spillovers. First, I define the network as undirected: a firm j belongs in firm i's neighborhood if it either buys from or sells to the firm. Table 9 and figure 7 show the endogenous network effects from gross output productivity. Without accounting for component-year fixed effects, the results suggest that firms' productivity increases by 0.06 percent when its average buyer or seller gets 10 percent more productive. Accounting for correlated effects lowers the estimate from GNR to 0.036 percent and my modified procedure's to 0.048 percent.

These estimates are driven by large positive spillovers between 1997 and 2006. As seen in figure 5, this was the onset of a period of rapid growth in sales for firms in the sample, as well as a shift in the central sectors from manufacturing to retail and wholesale. In contrast, spillovers during the other sub-periods are negative and mostly insignificant, suggesting free-riding on more efficient suppliers or buyers. An alternative explanation is that, since productivity was measured using deflated revenues rather than physical quantities, negative spillovers may also capture downward pressure on firms' output prices by relatively larger and more profitable customers.

To understand the economic importance of these spillover estimates, I conduct two exercises. First, I compute a back-of-the-envelope estimate of the impact of the most connected firm in each year on aggregate TFP through spillovers. Let j denote the most central firm in year t. I sum up j's contribution to the network average for each of its connections, and multiply that by the spillover estimate. That is:

$$\text{Contribution}_{jt} = \widehat{\lambda} \sum_{i} G_{ijt} \tag{59}$$

Figure 11 shows that for $\widehat{\lambda} = 0.0048$, a 10 percent increase in the TFP of the most central firm would correspond with a 2 to 4 percent increase in aggregrate TFP through spillovers alone.

Secondly, I simulate a growth path for the average firm in 1978 under four assumptions: it is connected to a single firm in the 10^{th}, 50^{th} or 90^{th} percentile of the 1978 productivity distribution, or it enjoys no spillovers. For simplicity, I assume there are no shocks to productivity and the average firm is also its partner's only connection, and this relationship remains the same for all periods. Then I compute:

$$\widetilde{\omega}_t = \left(I - \widehat{\lambda} G\right)^{-1} \left(\widehat{\beta}_1 + \widehat{\rho}\widetilde{\omega}_{t-1}\right), \qquad G = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{60}$$

where $t = \{1979 \ldots 2016\}$, $\widehat{\beta}_1 = 0.408$, $\widehat{\rho} = 0.899$ and $\widehat{\lambda} = 0.0048$.[26] Figure 12 shows that even in this simple setting, compared to a scenario in which there are no spillovers, the average firm in 1978 would be 20% more productive by 2016 due to spillovers alone. Because spillovers go in both directions and are mutually reinforcing, even a connection to a less productive partner

[26] $\widehat{\rho}$ is AR1 parameter estimate from GNR-ND table 9 while $\widehat{\beta}_1$ is the constant from GNR-N, since the constant is differenced out in GNR-ND.

is an improvement over experiencing no spillovers at all. This also explains the convergence in productivity growth by 2016 regardless of the initial TFP of the firm's partner. An average firm i linked to a less productive firm j in the 10^{th} percentile would not initially not benefit much from that relationship, but firm j's efficiency would be boosted by its connection to i. Over time, as j becomes more productive due to this relationship, i benefits more from the increased efficiency. By contrast, if j is in the 90^{th} percentile of the productivity distribution, firm i initially enjoys a substantial efficiency boost from the relationship. However, because j does not gain much from its connection to the less productive i, its own gains from spillovers slows, which in turn tamps down the benefits to i.

Estimates of endogenous spillovers using a value-added production function differ considerably from the results above. Without accounting for correlated network effects, there appears to be a significant positive effect of having more productive neighbors. However, after differencing to eliminate component-year fixed effects, spillover estimates become negative and insignificant. Like the gross output results, however, positive spillovers largely appear to have occurred from 1997 to 2006, while the other sub-periods featured negative spillovers.

Across all specifications, estimates from standard and modified approaches do not differ significantly, although there sometimes are small differences in the point estimates. This is consistent with the discussion in section A and results from the Monte Carlo experiment in table 6: standard approaches yielded estimates of productivity spillovers that were closest to the true effect when the network density was between 0.1 and 0.3 percent. As shown in figure 4, the density of the observed network in my sample ranges from 0.12 to 0.28 percent and lies within the region with the least bias in estimated spillovers.

8.2.1 Relationship Direction and Dynamics

Next, I examine how spillovers depend on the nature of the relationship. In figure 9 and table 11, it is clear that productive suppliers have almost twice the impact on their customers as productive buyers have on their suppliers: a 10 percent more productive supplier raises efficiency by 0.1 percent while customers raise productivity by 0.06 percent.

To investigate how much this is driven by maintaining old relationships as opposed to forming new ones, I decompose the interaction matrix into buyers/sellers that the firm traded with in both the current and the previous year (G_{t-1}), and changes from forming new links or severing old ones ($G_t - G_{t-1}$). The results in table 13 and figure 13 suggest that there is little appreciable difference in spillovers from new relationships or old ones.

Point estimates from value-added specifications do not show a significant difference between spillovers from buyers to sellers or vice versa. Most spillovers are negative except for 1997-2006 and only estimates from 1977-1986 are statistically significant, suggesting a 0.03 percent productivity penalty from transacting with more productive buyers and sellers, which does not vary by relationship persistence.

8.2.2 Heterogeneity by Sector

In figure 15 and table 15, I investigate how the spillovers transmit within and across sectors, as defined by Primary SIC. I also classify sectors 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 downstream, while the impact of sector v on firms in u is considered an upstream spillover. I focus on estimates that are significant at the 5% level.

Manufacturing firms enjoy productivity spillovers from all sectors except mining and construction. Their impact on other industries is predominantly as suppliers to firms in construction, and transportation and utilities. Wholesalers are a significant source of productivity spillovers to their upstream suppliers, and benefit from having more productive retailers. The production synergies between the Wholesale Trade and Mining sectors are primarily to crude oil exploration companies like Houston Exploration Co. and petroleum product transporters and resalers such as Adams Resources Energy Inc. Selling to a 10 percent more productive marketer would boost the efficiency of an exploration company by 0.2 percent.

While firms in Services are a source of both downstream and upstream spillovers, they do not benefit significantly from spillovers from other sectors. The construction sector is the only one with significant intra-sector spillovers, although spillovers between manufacturing firms are significant at the 6% level.

8.2.3 The Role of Firm Size

Finally, I consider the role of firm size in the transmission of efficiency gains through the production network. I classify firms as large if they have 500 or more employees, the definition used by the US BEA. The results are reported in table 17 and figure 16 highlights estimates that are significant at the 5% level. Large productive suppliers are an important source of productivity gains for both large and small customers, and the impact of having a more productive large supplier does not vary substantially by the customer's size. Likewise, large customers have a similar effect on their large and small suppliers. Small customers, however, do not have a substantial effect on their suppliers, but they enjoy additional benefits from their relationships with small suppliers.

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 19, 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 the impact of small suppliers on large customers becomes larger and statistically significant.

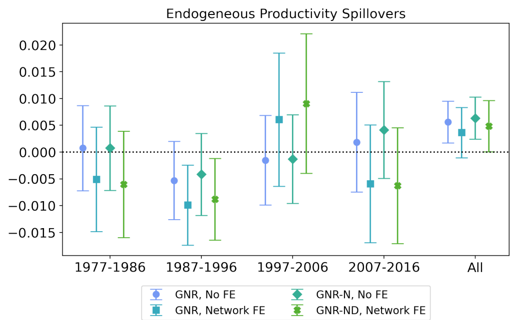


Figure 7: Spillover Estimates (Gross-Output)
This figure shows point estimates and 95% confidence intervals of endogenous productivity spillovers from a gross output production function. See table 9 for standard errors.

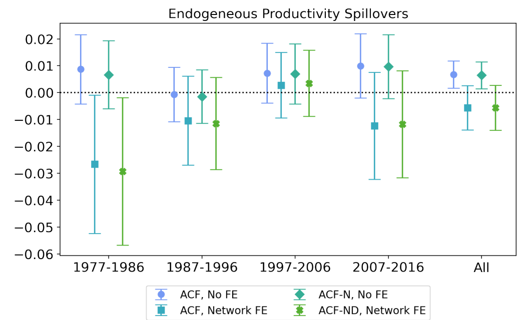


Figure 8: Spillover Estimates (Value-Added)
This figure shows point estimates and 95% confidence intervals of endogenous productivity spillovers from a value-added production function. See table 10 for standard errors.

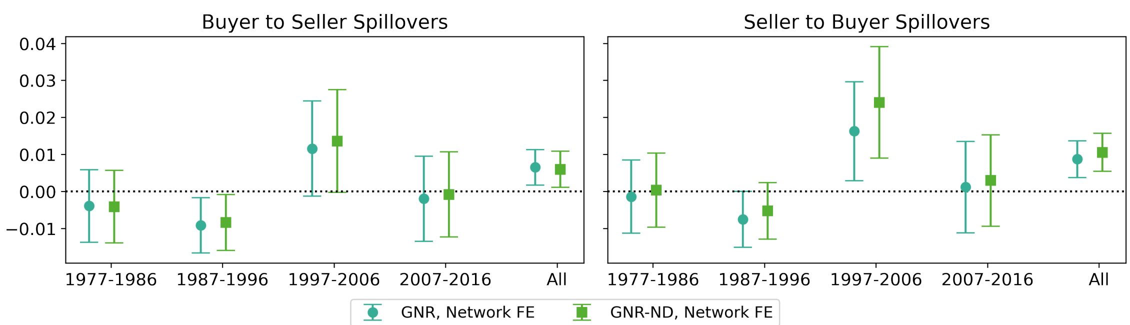


Figure 9: Spillover Estimates by Relationship Direction (Gross-Output)
This figure shows point estimates and 95% confidence intervals of endogenous productivity spillovers by direction of the relationship from a gross output production function. See table 11 for standard errors.

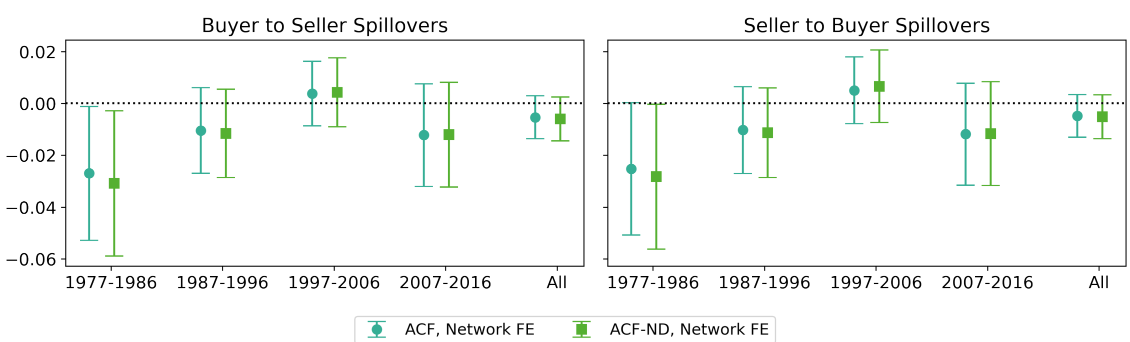


Figure 10: Spillover Estimates by Relationship Direction (Value-Added)
This figure shows point estimates and 95% confidence intervals of endogenous productivity spillovers by direction of the relationship from a value-added production function. See table 12 for standard errors.

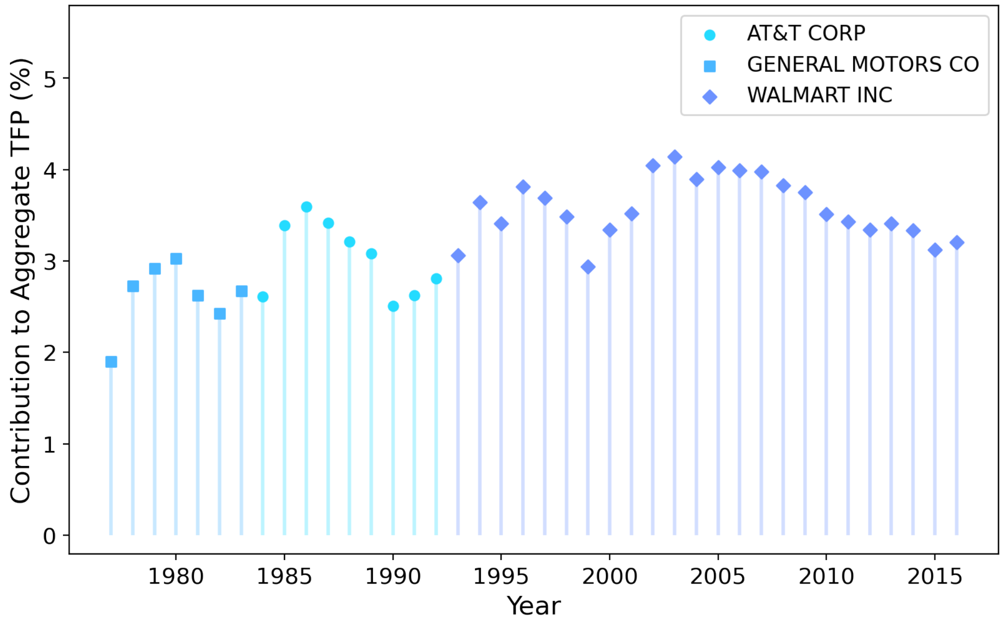


Figure 11: Contribution of Most Central Firms to Aggregate TFP
This figure shows the impact of a 10% increase in the TFP of the most central firm in each year to aggregate productivity through spillovers. The contribution of the most central firm j in year t is calculated as $\widehat{\lambda}\sum_i G_{ijt}$ where $\widehat{\lambda} = 0.048$.

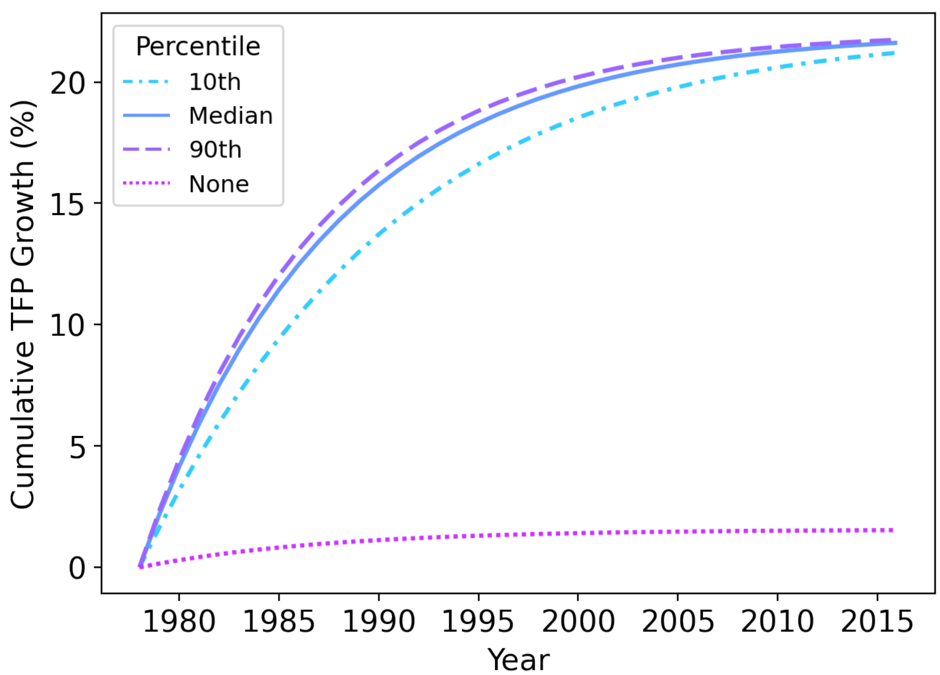


Figure 12: Cumulative Impact of Endogenous Productivity Spillovers over Time
This figure depicts a simulated path of log(TFP) for a firm with productivity equal to the 1978 average, assuming it is connected to a single firm that was at the 10th, 50th or 90th percentile of the productivity distribution in 1978, and endogeneous productivity spillovers = 0.0048. The bottom dotted line assumes that the firm experiences no spillovers.

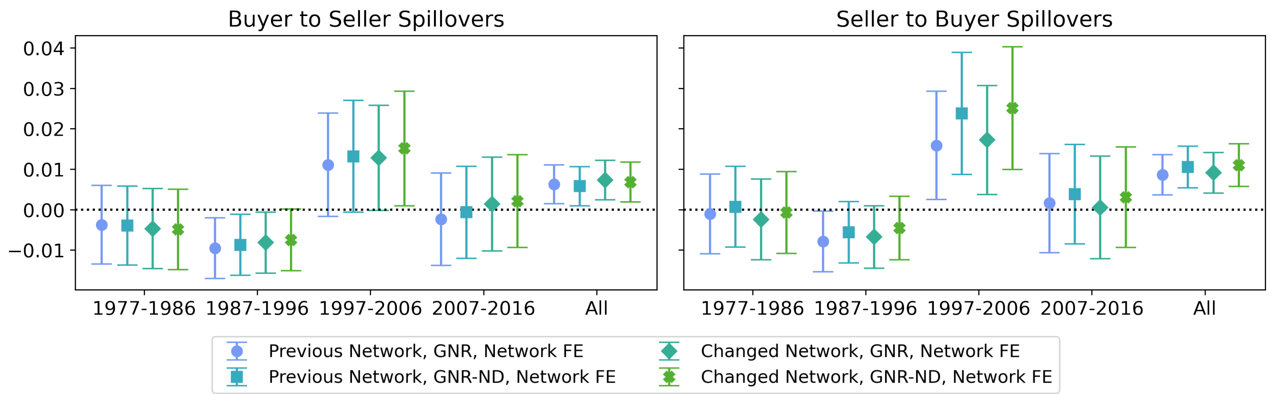


Figure 13: Spillover Estimates by Relationship Direction and Dynamics (Gross-Output)
This figure shows point estimates and 95% confidence intervals of endogenous productivity spillovers by direction and persistence of the relationship, from a gross output production function. See table 13 for standard errors.

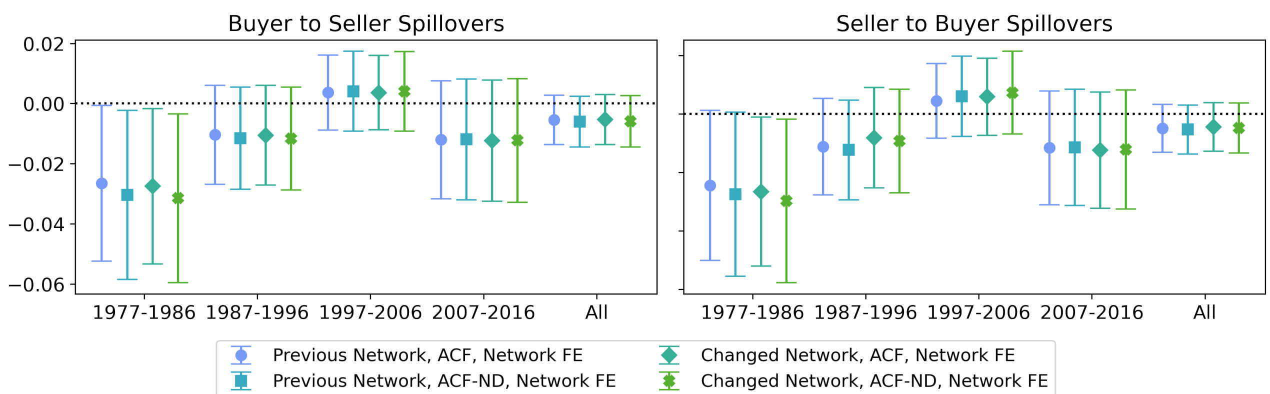
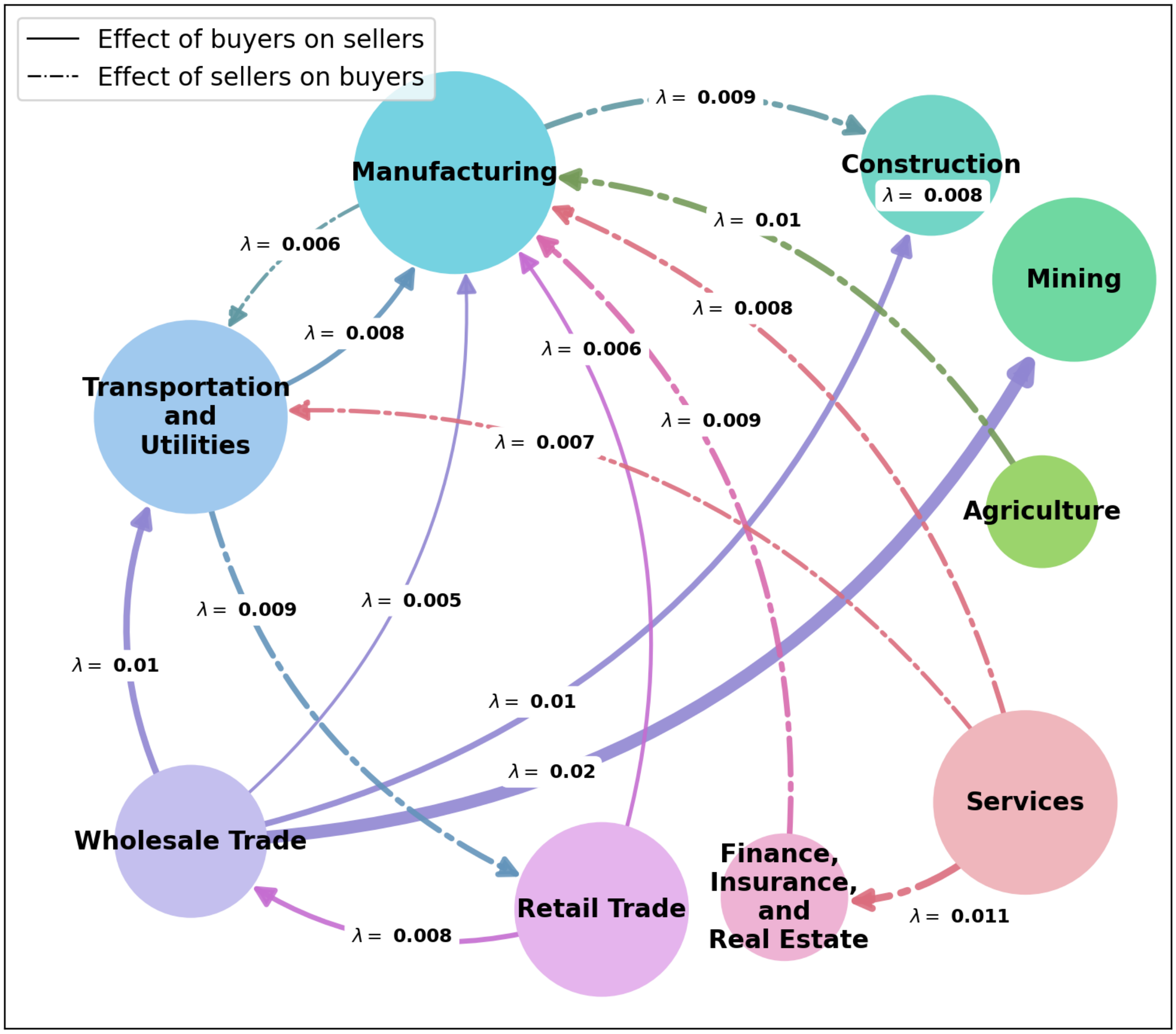


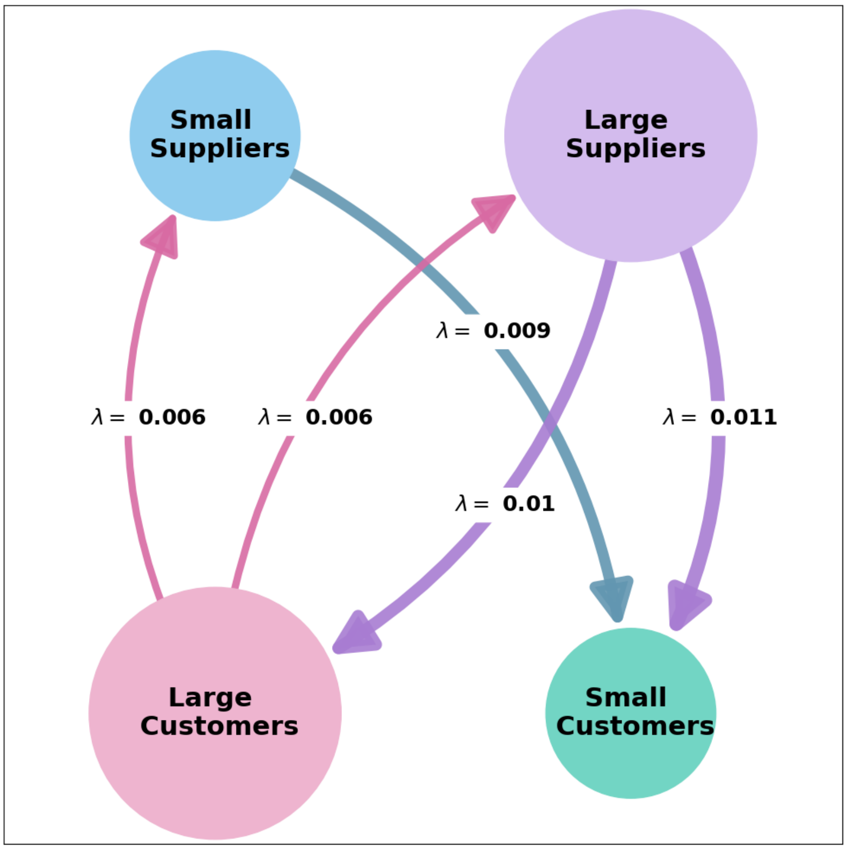
Figure 14: Spillover Estimates by Relationship Direction and Dynamics (Value-Added)
This figure shows point estimates and 95% confidence intervals of endogenous productivity spillovers by direction and persistence of the relationship, from a gross output production function. See table 14 for standard errors.

Figure 15: Endogenous Productivity Spillovers by Industry



This figure shows endogenous productivity spillovers (λ) that vary by the industry (Primary SIC) of the firm and its network. Estimates are significant at the 5% level. Industry nodes are weighted by the total number of connections originating from or going to firms in the industry, across all time periods. See table 15 for the full set of coefficients.

Figure 16: Endogenous Productivity Spillovers by Firm Size



This figure shows endogenous productivity spillovers that vary by firm size. Estimates are significant at the 5% level. See table 11 for the full set of coefficients.

Table 9: Endogenous Productivity Spillovers (Gross Output)

		Dependent Variable: lnTFP$_t$			
Period	TFP Estimator	lnTFP$_{t-1}$		Average Neighbors' lnTFP$_t$	
1977-1986	GNR	0.8293	0.8333	0.0007	-0.0051
		(0.0174)	(0.0158)	(0.0041)	(0.005)
	GNR-N/ND	0.8295	0.8315	0.0007	-0.006
		(0.0175)	(0.0158)	(0.004)	(0.0051)
1987-1996	GNR	0.7768	0.7639	-0.0053	-0.0099
		(0.0155)	(0.0144)	(0.0037)	(0.0038)
	GNR-N/ND	0.7696	0.7637	-0.0042	-0.0088
		(0.0137)	(0.0144)	(0.0039)	(0.0039)
1997-2006	GNR	0.8467	0.8372	-0.0015	0.0061
		(0.0125)	(0.012)	(0.0043)	(0.0063)
	GNR-N/ND	0.8464	0.8355	-0.0013	0.0091
		(0.0125)	(0.0121)	(0.0042)	(0.0067)
2007-2016	GNR	0.9013	0.8962	0.0018	-0.0059
		(0.0105)	(0.0101)	(0.0048)	(0.0056)
	GNR-N/ND	0.8922	0.8913	0.0041	-0.0063
		(0.0108)	(0.0103)	(0.0046)	(0.0055)
All	GNR	0.9064	0.8999	0.0056	0.0036
		(0.0043)	(0.0043)	(0.002)	(0.0024)
	GNR-N/ND	0.906	0.8993	0.0063	0.0048
		(0.0044)	(0.0043)	(0.002)	(0.0024)
Network Differencing		No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the standard Gandhi et al. (2020) procedure (GNR), or with modifications to accommodate network effects (GNR-N) and network differencing (GNR-ND). Network effects are estimated using the generalized 2SLS procedure in Lee (2003); Bramoullé et al. (2009). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 10: Endogenous Productivity Spillovers (Value-Added)

		Dependent Variable: $\ln\text{TFP}_t$			
Period	TFP Estimator	$\ln\text{TFP}_{t-1}$		Average Neighbors' $\ln\text{TFP}_t$	
1977-1986	ACF	0.8161	0.8243	0.0087	-0.0266
		(0.0225)	(0.0208)	(0.0066)	(0.0131)
	ACF-N/ND	0.8158	0.8217	0.0067	-0.0293
		(0.0226)	(0.0209)	(0.0065)	(0.014)
1987-1996	ACF	0.8566	0.8506	-0.0007	-0.0104
		(0.0143)	(0.0131)	(0.0052)	(0.0085)
	ACF-N/ND	0.8572	0.8495	-0.0015	-0.0115
		(0.0142)	(0.0132)	(0.0051)	(0.0087)
1997-2006	ACF	0.8802	0.8824	0.0073	0.0028
		(0.0123)	(0.0107)	(0.0057)	(0.0062)
	ACF-N/ND	0.8803	0.8835	0.007	0.0035
		(0.0123)	(0.0107)	(0.0057)	(0.0063)
2007-2016	ACF	0.8673	0.867	0.01	-0.0124
		(0.0396)	(0.0386)	(0.0061)	(0.0101)
	ACF-N/ND	0.8674	0.8686	0.0097	-0.0118
		(0.0396)	(0.0383)	(0.0061)	(0.0102)
All	ACF	0.8736	0.873	0.0068	-0.0056
		(0.0109)	(0.0102)	(0.0026)	(0.0042)
	ACF-N/ND	0.8737	0.8736	0.0065	-0.0056
		(0.0109)	(0.0101)	(0.0026)	(0.0042)
Network Differencing		No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a value-added production function (in logs) estimated with the standard Ackerberg et al. (2015) procedure (ACF), or with modifications to accommodate network effects (ACF-N) and network differencing (ACF-ND). Network effects are estimated using the generalized 2SLS procedure in Lee (2003); Bramoullé et al. (2009). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 11: Productivity Spillovers by Relationship Direction (Gross Output)

		Dependent Variable: $\ln\text{TFP}_t$					
Period	TFP Estimator	$\ln\text{TFP}_{t-1}$		Customers' $\ln\text{TFP}_t$		Suppliers' $\ln\text{TFP}_t$	
1977-1986	GNR	0.8261	0.8295	0.0016	-0.004	0.0041	-0.0014
		(0.0175)	(0.0159)	(0.0041)	(0.005)	(0.0042)	(0.005)
	GNR-N/ND	0.8251	0.8268	0.0021	-0.0041	0.006	0.0003
		(0.018)	(0.0163)	(0.0041)	(0.005)	(0.0043)	(0.0051)
1987-1996	GNR	0.775	0.7628	-0.0042	-0.0092	-0.0016	-0.0076
		(0.0156)	(0.0145)	(0.0037)	(0.0038)	(0.0038)	(0.0038)
	GNR-N/ND	0.7655	0.7623	-0.0017	-0.0084	0.0006	-0.0053
		(0.0138)	(0.0146)	(0.0039)	(0.0039)	(0.004)	(0.0039)
1997-2006	GNR	0.8435	0.8324	0.0027	0.0116	0.0059	0.0163
		(0.0126)	(0.0122)	(0.0045)	(0.0065)	(0.0048)	(0.0068)
	GNR-N/ND	0.8416	0.8328	0.0033	0.0136	0.0081	0.024
		(0.0124)	(0.0121)	(0.0046)	(0.0071)	(0.0049)	(0.0077)
2007-2016	GNR	0.8968	0.8937	0.0084	-0.002	0.0134	0.0011
		(0.0112)	(0.0105)	(0.0054)	(0.0059)	(0.0062)	(0.0063)
	GNR-N/ND	0.8889	0.8897	0.0136	-0.0008	0.0187	0.0029
		(0.0116)	(0.0108)	(0.0056)	(0.0059)	(0.0063)	(0.0063)
All	GNR	0.9037	0.8975	0.0088	0.0065	0.0112	0.0087
		(0.0045)	(0.0044)	(0.0021)	(0.0025)	(0.0023)	(0.0025)
	GNR-N/ND	0.9041	0.8981	0.0085	0.006	0.0127	0.0105
		(0.0045)	(0.0044)	(0.0021)	(0.0025)	(0.0023)	(0.0026)
Network Differencing		No	Yes	No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the standard Gandhi et al. (2020) procedure (GNR), or with modifications to accommodate network effects (GNR-N) and network differencing (GNR-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 12: Productivity Spillovers by Relationship Direction (Value-Added)

		Dependent Variable: $\ln\text{TFP}_t$					
Period	TFP Estimator	$\ln\text{TFP}_{t-1}$		Customers' $\ln\text{TFP}_t$		Suppliers' $\ln\text{TFP}_t$	
1977-1986	ACF	0.8157	0.8238	0.0082	-0.027	0.0096	-0.0253
		(0.0225)	(0.0208)	(0.0066)	(0.0132)	(0.0066)	(0.0131)
	ACF-N/ND	0.8155	0.8212	0.0055	-0.0309	0.0081	-0.0283
		(0.0226)	(0.021)	(0.0064)	(0.0143)	(0.0066)	(0.0143)
1987-1996	ACF	0.8567	0.8507	-0.0008	-0.0105	-0.0005	-0.0103
		(0.0143)	(0.0131)	(0.0052)	(0.0084)	(0.0052)	(0.0085)
	ACF-N/ND	0.8572	0.8496	-0.0014	-0.0116	-0.0015	-0.0114
		(0.0142)	(0.0132)	(0.0051)	(0.0087)	(0.0051)	(0.0088)
1997-2006	ACF	0.8802	0.8826	0.0077	0.0038	0.0083	0.0051
		(0.0123)	(0.0107)	(0.0057)	(0.0064)	(0.0057)	(0.0066)
	ACF-N/ND	0.8811	0.8856	0.0076	0.0043	0.0086	0.0066
		(0.0123)	(0.0106)	(0.0057)	(0.0068)	(0.0058)	(0.0071)
2007-2016	ACF	0.8672	0.8669	0.0102	-0.0122	0.0108	-0.0119
		(0.0396)	(0.0387)	(0.0061)	(0.0101)	(0.0063)	(0.01)
	ACF-N/ND	0.8681	0.8688	0.0097	-0.0121	0.0109	-0.0117
		(0.0397)	(0.0384)	(0.0061)	(0.0103)	(0.0064)	(0.0102)
All	ACF	0.8735	0.8729	0.007	-0.0054	0.0075	-0.0048
		(0.0109)	(0.0102)	(0.0026)	(0.0042)	(0.0026)	(0.0042)
	ACF-N/ND	0.8741	0.8739	0.0063	-0.006	0.0071	-0.0052
		(0.0109)	(0.0101)	(0.0025)	(0.0043)	(0.0026)	(0.0043)
Network Differencing		No	Yes	No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a value-added production function (in logs) estimated with the standard Ackerberg et al. (2015) procedure (ACF), or with modifications to accommodate network effects (ACF-N) and network differencing (ACF-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 13: Productivity Spillovers by Relationship Dynamics (Gross Output)

		Dependent Variable: $\ln\text{TFP}_t$									
Period	TFP estimator	$\ln\text{TFP}_{t-1}$		Previous Customers' $\ln\text{TFP}_t$		Changed Customers' $\ln\text{TFP}_t$		Previous Suppliers' $\ln\text{TFP}_t$		Changed Suppliers' $\ln\text{TFP}_t$	
1977-1986	GNR	0.8262	0.8295	0.0018	-0.0037	0.0013	-0.0047	0.0041	-0.0011	0.0043	-0.0025
		(0.0176)	(0.0159)	(0.0041)	(0.005)	(0.0042)	(0.0051)	(0.0042)	(0.005)	(0.0043)	(0.0051)
	GNR-N/ND	0.8252	0.8268	0.0022	-0.004	0.0018	-0.0049	0.0058	0.0007	0.006	-0.0007
		(0.0181)	(0.0163)	(0.0041)	(0.005)	(0.0041)	(0.0051)	(0.0043)	(0.0051)	(0.0044)	(0.0052)
1987-1996	GNR	0.7746	0.7623	-0.0045	-0.0095	-0.0032	-0.0082	-0.0016	-0.0079	-0.0016	-0.0068
		(0.0156)	(0.0145)	(0.0037)	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.0039)	(0.0039)
	GNR-N/ND	0.774	0.7621	-0.0043	-0.0087	-0.0031	-0.0075	0.0002	-0.0056	0.0	-0.0046
		(0.0157)	(0.0146)	(0.0037)	(0.0039)	(0.0038)	(0.0039)	(0.0038)	(0.0039)	(0.0039)	(0.004)
1997-2006	GNR	0.8439	0.8327	0.0021	0.0111	0.004	0.0128	0.0052	0.0159	0.0074	0.0172
		(0.0126)	(0.0122)	(0.0045)	(0.0065)	(0.0045)	(0.0066)	(0.0048)	(0.0068)	(0.0048)	(0.0069)
	GNR-N/ND	0.8441	0.8331	0.0028	0.0132	0.0046	0.0151	0.0085	0.0238	0.0105	0.0251
		(0.0126)	(0.0122)	(0.0046)	(0.0071)	(0.0046)	(0.0072)	(0.005)	(0.0077)	(0.005)	(0.0077)
2007-2016	GNR	0.8968	0.8935	0.008	-0.0024	0.0111	0.0014	0.0134	0.0016	0.0138	0.0005
		(0.0112)	(0.0105)	(0.0054)	(0.0058)	(0.0055)	(0.0059)	(0.0062)	(0.0063)	(0.0064)	(0.0065)
	GNR-N/ND	0.8894	0.89	0.0135	-0.0007	0.0158	0.0021	0.0189	0.0038	0.0191	0.003
		(0.0116)	(0.0108)	(0.0056)	(0.0058)	(0.0056)	(0.0058)	(0.0063)	(0.0063)	(0.0064)	(0.0063)
All	GNR	0.9038	0.8977	0.0086	0.0063	0.0097	0.0073	0.011	0.0086	0.0119	0.0091
		(0.0045)	(0.0044)	(0.0021)	(0.0025)	(0.0021)	(0.0025)	(0.0023)	(0.0025)	(0.0023)	(0.0026)
	GNR-N/ND	0.9043	0.8981	0.0083	0.0058	0.0094	0.0068	0.0125	0.0105	0.0134	0.011
		(0.0045)	(0.0044)	(0.0021)	(0.0025)	(0.0022)	(0.0025)	(0.0023)	(0.0026)	(0.0024)	(0.0027)
Network Differencing		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the standard Gandhi et al. (2020) procedure (GNR), or with modifications to accommodate network effects (GNR-N) and network differencing (GNR-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 14: Productivity Spillovers by Relationship Dynamics (Value-Added)

		Dependent Variable: $\ln TFP_t$									
Period	TFP estimator	$\ln TFP_{t-1}$		Previous Customers' $\ln TFP_t$		Changed Customers' $\ln TFP_t$		Previous Suppliers' $\ln TFP_t$		Changed Suppliers' $\ln TFP_t$	
1977-1986	ACF	0.8157	0.8238	0.0082	-0.0265	0.0081	-0.0275	0.0096	-0.0245	0.0094	-0.0266
		(0.0225)	(0.0208)	(0.0066)	(0.0132)	(0.0067)	(0.0132)	(0.0066)	(0.0131)	(0.0068)	(0.013)
	ACF-N/ND	0.8155	0.8212	0.0055	-0.0304	0.0054	-0.0315	0.0081	-0.0275	0.0079	-0.0297
		(0.0226)	(0.0211)	(0.0064)	(0.0143)	(0.0065)	(0.0143)	(0.0065)	(0.0143)	(0.0067)	(0.0142)
1987-1996	ACF	0.8566	0.8505	-0.001	-0.0104	-0.0008	-0.0106	-0.001	-0.0112	0.0004	-0.0081
		(0.0142)	(0.013)	(0.0052)	(0.0084)	(0.0052)	(0.0084)	(0.0052)	(0.0084)	(0.0053)	(0.0088)
	ACF-N/ND	0.8569	0.8495	-0.0015	-0.0116	-0.0013	-0.0117	-0.0021	-0.0123	-0.0007	-0.0092
		(0.0141)	(0.0131)	(0.0052)	(0.0087)	(0.0051)	(0.0087)	(0.0051)	(0.0087)	(0.0052)	(0.009)
1997-2006	ACF	0.8801	0.8826	0.0077	0.0036	0.0077	0.0036	0.0081	0.0045	0.0086	0.0059
		(0.0123)	(0.0107)	(0.0057)	(0.0064)	(0.0057)	(0.0063)	(0.0057)	(0.0065)	(0.0057)	(0.0067)
	ACF-N/ND	0.881	0.8856	0.0076	0.0041	0.0075	0.004	0.0084	0.006	0.0088	0.0073
		(0.0123)	(0.0106)	(0.0057)	(0.0068)	(0.0057)	(0.0067)	(0.0058)	(0.007)	(0.0058)	(0.0072)
2007-2016	ACF	0.8674	0.8669	0.0104	-0.0121	0.0099	-0.0124	0.011	-0.0116	0.0108	-0.0124
		(0.0395)	(0.0385)	(0.0063)	(0.01)	(0.006)	(0.0103)	(0.0063)	(0.0099)	(0.0064)	(0.0101)
	ACF-N/ND	0.8684	0.8688	0.0099	-0.0119	0.0094	-0.0123	0.0111	-0.0114	0.0108	-0.0121
		(0.0396)	(0.0383)	(0.0062)	(0.0102)	(0.006)	(0.0105)	(0.0064)	(0.0102)	(0.0064)	(0.0104)
All	ACF	0.8734	0.8729	0.0068	-0.0054	0.0071	-0.0053	0.0073	-0.005	0.0077	-0.0044
		(0.0109)	(0.0102)	(0.0026)	(0.0042)	(0.0026)	(0.0042)	(0.0027)	(0.0042)	(0.0027)	(0.0043)
	ACF-N/ND	0.8739	0.8739	0.0062	-0.006	0.0064	-0.006	0.0069	-0.0053	0.0073	-0.0048
		(0.0109)	(0.0101)	(0.0026)	(0.0043)	(0.0025)	(0.0044)	(0.0026)	(0.0043)	(0.0027)	(0.0044)
Network Differencing		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a value-added production function (in logs) estimated with the standard Ackerberg et al. (2015) procedure (ACF), or with modifications to accommodate network effects (ACF-N) and network differencing (ACF-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 15: Productivity Spillovers by Industry (Gross Output)

Dependent Variable: $\ln\text{TFP}_t$										
		Firm's Industry								
Neighbors' Industry	TFP Estimator	Agriculture	Mining	Construction	Mfg	Transport & Utilities	Wholesale Trade	Retail Trade	Finance, Insurance, & Real Estate	Services
Agriculture	GNR	0.006	-0.012	-0.002	0.0087	0.0001	-0.0008	0.0078	0.0087	-0.0016
		(0.0123)	(0.0112)	(0.0096)	(0.0046)	(0.0041)	(0.0108)	(0.007)	(0.0144)	(0.0057)
	GNR-ND	0.007	-0.0126	-0.0035	0.01	0.0001	-0.001	0.0084	0.0111	-0.0031
		(0.0141)	(0.0126)	(0.0116)	(0.0051)	(0.0045)	(0.0124)	(0.0078)	(0.0166)	(0.006)
Mining	GNR	-0.0023	0.003	0.0046	0.003	0.0044	0.0082	0.0064	0.0003	0.0036
		(0.0126)	(0.0037)	(0.0118)	(0.0032)	(0.0033)	(0.0116)	(0.0133)	(0.0058)	(0.0091)
	GNR-ND	-0.002	0.0035	0.0059	0.0036	0.0051	0.0095	0.0078	0.0005	0.0048
		(0.0142)	(0.0039)	(0.0135)	(0.0034)	(0.0035)	(0.0131)	(0.0149)	(0.0064)	(0.0105)
Construction	GNR	-0.0111	0.0295	0.0073	0.0041	0.0034	0.0055	-0.0	-0.0015	-0.0001
		(0.0079)	(0.0166)	(0.0034)	(0.0031)	(0.0033)	(0.0035)	(0.0032)	(0.0086)	(0.0036)
	GNR-ND	-0.0137	0.0346	0.0081	0.0046	0.0039	0.007	-0.0005	-0.0029	-0.0003
		(0.0092)	(0.019)	(0.0037)	(0.0033)	(0.0035)	(0.0038)	(0.0034)	(0.0097)	(0.0039)
Manufacturing	GNR	0.0011	0.0046	0.0085	0.0043	0.0051	0.0024	-0.0011	0.0034	0.0029
		(0.0057)	(0.0035)	(0.0036)	(0.0026)	(0.0027)	(0.0028)	(0.0027)	(0.0034)	(0.0027)
	GNR-ND	0.0014	0.0052	0.0094	0.005	0.0058	0.0026	-0.0016	0.0041	0.0033
		(0.0064)	(0.0037)	(0.0039)	(0.0026)	(0.0027)	(0.0029)	(0.0027)	(0.0036)	(0.0028)
Transport & Utilities	GNR	0.0026	0.0038	0.0044	0.0069	0.0042	0.0036	0.0085	0.0025	0.0032
		(0.0049)	(0.0037)	(0.0036)	(0.0026)	(0.0028)	(0.0057)	(0.0034)	(0.004)	(0.0028)
	GNR-ND	0.0027	0.0043	0.0048	0.008	0.0048	0.0044	0.0091	0.0028	0.0038
		(0.0054)	(0.0039)	(0.0039)	(0.0026)	(0.0029)	(0.0064)	(0.0036)	(0.0043)	(0.0028)
Wholesale Trade	GNR	0.0011	0.0178	0.0087	0.0044	0.0091	0.0042	-0.0005	-0.0026	-0.0036
		(0.01)	(0.0086)	(0.004)	(0.0024)	(0.0046)	(0.0035)	(0.0026)	(0.0089)	(0.004)
	GNR-ND	0.0012	0.0205	0.0097	0.005	0.0104	0.0042	-0.0007	-0.0034	-0.0041
		(0.0113)	(0.0097)	(0.0043)	(0.0024)	(0.0051)	(0.0037)	(0.0027)	(0.0098)	(0.0044)
Retail Trade	GNR	0.0009	-0.0041	0.004	0.0053	0.005	0.0069	0.0021	0.0063	0.0043
		(0.0048)	(0.0105)	(0.0032)	(0.0026)	(0.0033)	(0.0026)	(0.0029)	(0.0039)	(0.003)
	GNR-ND	0.001	-0.004	0.0045	0.0062	0.0057	0.0081	0.0024	0.0074	0.005
		(0.0053)	(0.012)	(0.0033)	(0.0026)	(0.0035)	(0.0027)	(0.003)	(0.0042)	(0.0031)
Finance, Insurance & Real Estate	GNR	0.0145	0.0065	0.0062	0.008	0.0057	0.0015	0.0024	-0.0002	0.0002
		(0.0119)	(0.0075)	(0.0063)	(0.0039)	(0.0045)	(0.0065)	(0.0031)	(0.0041)	(0.0034)
	GNR-ND	0.017	0.008	0.0067	0.009	0.0067	0.0008	0.0026	-0.0005	0.0003
		(0.0135)	(0.0083)	(0.0071)	(0.0042)	(0.0049)	(0.0071)	(0.0033)	(0.0045)	(0.0036)
Services	GNR	-0.0018	0.0127	0.001	0.0069	0.0061	0.01	0.0041	0.0104	0.004
		(0.0059)	(0.0089)	(0.0044)	(0.0029)	(0.0032)	(0.0065)	(0.0034)	(0.0033)	(0.003)
	GNR-ND	-0.0034	0.014	0.0004	0.0081	0.0069	0.0119	0.0043	0.0115	0.0046
		(0.0056)	(0.0102)	(0.0048)	(0.0029)	(0.0034)	(0.0076)	(0.0036)	(0.0035)	(0.0031)

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the standard Gandhi et al. (2020) procedure (GNR), or with modifications to accommodate network effects and network differencing (GNR-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017), accounting for network fixed effects. Industries are defined as firms' Primary SIC code. Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 16: Productivity Spillovers by Industry (Value-Added)

Dependent Variable: $\ln\mathrm{TFP}_t$										
		Firm's Industry								
Neighbors' Industry	TFP Estimator	Agriculture	Mining	Construction	Mfg	Transport & Utilities	Wholesale Trade	Retail Trade	Finance, Insurance, & Real Estate	Services
Agriculture	ACF	-0.0234	-0.0176	0.0106	-0.0079	-0.0119	-0.0212	-0.0261	0.0067	-0.0123
		(0.0105)	(0.0079)	(0.0103)	(0.0061)	(0.0058)	(0.0086)	(0.0068)	(0.0104)	(0.0114)
	ACF-ND	-0.0228	-0.0166	0.0116	-0.0071	-0.0112	-0.0206	-0.0258	0.0077	-0.0111
		(0.0105)	(0.0078)	(0.0107)	(0.006)	(0.0057)	(0.0085)	(0.0067)	(0.0104)	(0.0113)
Mining	ACF	-0.0047	-0.0077	-0.0233	-0.0086	-0.005	-0.0069	-0.0169	-0.0097	-0.0058
		(0.0081)	(0.0053)	(0.0096)	(0.0052)	(0.0052)	(0.0062)	(0.0072)	(0.0062)	(0.0093)
	ACF-ND	-0.0039	-0.007	-0.0225	-0.0078	-0.0043	-0.0062	-0.0162	-0.0088	-0.0051
		(0.008)	(0.0052)	(0.0096)	(0.0051)	(0.0052)	(0.0061)	(0.0071)	(0.0061)	(0.0092)
Construction	ACF	-0.0007	-0.0109	-0.0109	-0.0099	-0.0057	0.0024	-0.0177	-0.0193	-0.0161
		(0.0094)	(0.0108)	(0.0055)	(0.0054)	(0.0056)	(0.0103)	(0.0058)	(0.007)	(0.0068)
	ACF-ND	0.0	-0.0102	-0.0102	-0.0091	-0.0051	0.0029	-0.0172	-0.0188	-0.0153
		(0.0096)	(0.0108)	(0.0054)	(0.0053)	(0.0055)	(0.0101)	(0.0057)	(0.0069)	(0.0067)
Manufacturing	ACF	-0.0101	-0.0066	-0.0085	-0.0107	-0.008	-0.0115	-0.0151	-0.0084	-0.0113
		(0.0073)	(0.0055)	(0.0057)	(0.0054)	(0.0053)	(0.0054)	(0.0057)	(0.0055)	(0.0055)
	ACF-ND	-0.0092	-0.0058	-0.008	-0.0099	-0.0074	-0.0109	-0.0146	-0.0077	-0.0106
		(0.0072)	(0.0054)	(0.0056)	(0.0054)	(0.0052)	(0.0053)	(0.0056)	(0.0054)	(0.0055)
Transport & Utilities	ACF	-0.0133	-0.0057	-0.0128	-0.0086	-0.0072	-0.0076	-0.0173	-0.0082	-0.0084
		(0.0059)	(0.0054)	(0.0057)	(0.0051)	(0.0053)	(0.0061)	(0.0061)	(0.0055)	(0.0053)
	ACF-ND	-0.0125	-0.0049	-0.0121	-0.0079	-0.0065	-0.0069	-0.0166	-0.0074	-0.0078
		(0.0058)	(0.0053)	(0.0056)	(0.005)	(0.0052)	(0.006)	(0.006)	(0.0055)	(0.0052)
Wholesale Trade	ACF	-0.0213	-0.0037	-0.009	-0.0078	-0.0065	-0.0084	-0.0141	-0.0132	-0.0139
		(0.0083)	(0.0076)	(0.0055)	(0.005)	(0.0063)	(0.0074)	(0.0055)	(0.0221)	(0.0065)
	ACF-ND	-0.0206	-0.0028	-0.0083	-0.0071	-0.0059	-0.0078	-0.0136	-0.0122	-0.013
		(0.0083)	(0.0076)	(0.0055)	(0.0049)	(0.0062)	(0.0073)	(0.0054)	(0.0221)	(0.0064)
Retail Trade	ACF	-0.0164	-0.0106	-0.0137	-0.0086	-0.01	-0.0051	-0.0122	-0.0083	-0.0096
		(0.0063)	(0.0064)	(0.0057)	(0.0055)	(0.006)	(0.0051)	(0.0055)	(0.006)	(0.0055)
	ACF-ND	-0.0159	-0.0098	-0.0129	-0.0079	-0.0093	-0.0044	-0.0116	-0.0077	-0.0089
		(0.0062)	(0.0064)	(0.0056)	(0.0054)	(0.0059)	(0.005)	(0.0055)	(0.006)	(0.0054)
Finance, Insurance & Real Estate	ACF	0.0024	0.003	-0.0096	-0.0063	-0.0059	-0.009	-0.0097	-0.0145	-0.009
		(0.008)	(0.0069)	(0.0054)	(0.0053)	(0.0053)	(0.0064)	(0.0052)	(0.0061)	(0.0052)
	ACF-ND	0.0029	0.0039	-0.0088	-0.0057	-0.0052	-0.0084	-0.0093	-0.0138	-0.0083
		(0.0079)	(0.0068)	(0.0054)	(0.0052)	(0.0052)	(0.0063)	(0.0051)	(0.006)	(0.0052)
Services	ACF	-0.0241	-0.0144	-0.0149	-0.0097	-0.0093	-0.0043	-0.0151	-0.0064	-0.0092
		(0.0102)	(0.0069)	(0.0061)	(0.0055)	(0.0055)	(0.0118)	(0.0056)	(0.0055)	(0.0056)
	ACF-ND	-0.0232	-0.0138	-0.0143	-0.009	-0.0087	-0.0036	-0.0146	-0.0059	-0.0086
		(0.0101)	(0.0068)	(0.006)	(0.0054)	(0.0054)	(0.0117)	(0.0055)	(0.0054)	(0.0055)

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the standard Ackerberg et al. (2015) procedure (ACF), or with modifications to accommodate network effects and network differencing (ACF-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017), accounting for network fixed effects. Industries are defined as firms' Primary SIC code. Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 17: Productivity Spillovers by Firm Size & Relationship Direction (Gross Output)

				Dependent Variable: $\ln\text{TFP}_t$									
Neighbor Size	Relationship	Firm Size	TFP Estimator	1977-1986		1987-1996		1997-2006		2007-2016		All	
Large	$\ln\text{TFP}_{t-1}$		GNR	0.8262	0.8294	0.7745	0.7623	0.844	0.8327	0.897	0.8939	0.9038	0.8977
				(0.0175)	(0.0158)	(0.0156)	(0.0145)	(0.0127)	(0.0122)	(0.0112)	(0.0106)	(0.0045)	(0.0044)
			GNR-N/ND	0.8258	0.8276	0.7733	0.762	0.8433	0.8324	0.8888	0.8898	0.9041	0.8982
				(0.0182)	(0.0163)	(0.0157)	(0.0146)	(0.0127)	(0.0123)	(0.0118)	(0.011)	(0.0045)	(0.0044)
	Customers	Large	GNR	0.0013	-0.0041	-0.0052	-0.0102	0.0021	0.0106	0.008	-0.0024	0.0085	0.0062
				(0.0041)	(0.005)	(0.0037)	(0.0038)	(0.0045)	(0.0065)	(0.0053)	(0.0058)	(0.0021)	(0.0025)
			GNR-N/ND	0.0017	-0.0041	-0.0046	-0.0087	0.003	0.0132	0.0138	-0.0003	0.0083	0.006
				(0.0041)	(0.005)	(0.0037)	(0.0039)	(0.0046)	(0.0071)	(0.0056)	(0.0059)	(0.0021)	(0.0025)
		Small	GNR	0.002	-0.0039	-0.0032	-0.0083	0.004	0.0127	0.0097	-0.0007	0.0094	0.0072
				(0.0041)	(0.0051)	(0.0038)	(0.0038)	(0.0045)	(0.0066)	(0.0055)	(0.0059)	(0.0021)	(0.0025)
			GNR-N/ND	0.0008	-0.0054	-0.0032	-0.008	0.0044	0.0144	0.0132	-0.0007	0.0086	0.006
				(0.0041)	(0.0051)	(0.0038)	(0.0039)	(0.0046)	(0.0072)	(0.0057)	(0.006)	(0.0021)	(0.0025)
	Suppliers	Large	GNR	0.0043	-0.0018	-0.002	-0.0081	0.0054	0.0153	0.014	0.0008	0.011	0.0083
				(0.0042)	(0.0051)	(0.0039)	(0.0039)	(0.0048)	(0.0068)	(0.0062)	(0.0063)	(0.0023)	(0.0025)
			GNR-N/ND	0.0052	-0.0001	-0.0002	-0.0055	0.0087	0.0236	0.0189	0.0032	0.0125	0.0102
				(0.0043)	(0.0051)	(0.0039)	(0.0039)	(0.005)	(0.0077)	(0.0063)	(0.0063)	(0.0023)	(0.0026)
		Small	GNR	0.0034	-0.0015	-0.0014	-0.0073	0.0066	0.0171	0.0127	0.0022	0.0114	0.0095
				(0.0043)	(0.0051)	(0.0038)	(0.0039)	(0.0049)	(0.0069)	(0.0063)	(0.0065)	(0.0023)	(0.0026)
			GNR-N/ND	0.0046	-0.0001	0.0	-0.005	0.0093	0.0246	0.0188	0.0041	0.0128	0.0114
				(0.0043)	(0.0052)	(0.0038)	(0.004)	(0.0051)	(0.0078)	(0.0065)	(0.0065)	(0.0024)	(0.0027)
Small	Customers	Large	GNR	0.0082	0.005	-0.0103	-0.0196	-0.009	-0.0113	0.008	-0.0206	0.0043	0.0002
				(0.0097)	(0.0135)	(0.0065)	(0.0077)	(0.0095)	(0.0118)	(0.0139)	(0.0184)	(0.004)	(0.0051)
			GNR-N/ND	0.0073	0.0044	-0.0092	-0.0155	-0.009	-0.0144	0.0136	-0.0129	0.004	-0.0011
				(0.008)	(0.0129)	(0.0063)	(0.0075)	(0.0098)	(0.014)	(0.009)	(0.0122)	(0.0042)	(0.0057)
		Small	GNR	-0.0033	-0.0075	-0.0067	-0.0124	0.0059	0.02	0.0096	0.0031	0.0077	0.0074
				(0.0077)	(0.0083)	(0.0047)	(0.0055)	(0.0062)	(0.0086)	(0.0095)	(0.0101)	(0.0027)	(0.0033)
			GNR-N/ND	-0.0026	-0.0086	-0.0065	-0.0116	0.0068	0.0242	0.0141	0.0033	0.007	0.0064
				(0.0066)	(0.0081)	(0.0047)	(0.0053)	(0.0064)	(0.0101)	(0.0072)	(0.0079)	(0.0029)	(0.0036)
	Suppliers	Large	GNR	0.0239	0.0186	-0.0003	-0.0109	0.0041	0.0065	0.0041	-0.0181	0.0109	0.0059
				(0.0143)	(0.0149)	(0.0089)	(0.0077)	(0.0066)	(0.0093)	(0.01)	(0.0147)	(0.0039)	(0.0045)
			GNR-N/ND	0.0192	0.0169	-0.0004	-0.009	0.0046	0.007	0.0121	-0.0099	0.0105	0.0047
				(0.0115)	(0.0142)	(0.0088)	(0.0072)	(0.0068)	(0.0109)	(0.0074)	(0.0098)	(0.0041)	(0.005)
		Small	GNR	0.0052	-0.0005	-0.002	-0.0118	0.0117	0.0263	0.02	0.0084	0.0127	0.0099
				(0.0062)	(0.0076)	(0.0046)	(0.0053)	(0.0062)	(0.0083)	(0.0089)	(0.0102)	(0.0028)	(0.0033)
			GNR-N/ND	0.0037	-0.002	-0.002	-0.0105	0.0125	0.0314	0.02	0.0057	0.0123	0.0091
				(0.0057)	(0.0076)	(0.0045)	(0.0052)	(0.0065)	(0.0097)	(0.0072)	(0.0082)	(0.0029)	(0.0036)
Network Differencing				No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the standard Gandhi et al. (2020) procedure (GNR), or with modifications to accommodate network effects (GNR-N) and network differencing (GNR-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Large firms are businesses with 500 or more employees. Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 18: Productivity Spillovers by Firm Size & Relationship Direction (Value-Added)

				Dependent Variable: $\ln\text{TFP}_t$									
Neighbor Size	Relationship	Firm Size	TFP Estimator	1977-1986		1987-1996		1997-2006		2007-2016		All	
	$\ln\text{TFP}_{t-1}$		ACF	0.8156	0.8234	0.8564	0.8507	0.8798	0.8822	0.8673	0.8673	0.8733	0.8729
				0.0224	0.0206	0.0144	0.0131	0.0123	0.0108	(0.0401)	(0.0389)	(0.0109)	(0.0102)
			ACF-N/ND	0.8155	0.821	0.857	0.851	0.8814	0.8864	0.8687	0.8703	0.8737	0.8739
				0.0222	0.0207	0.0143	0.0133	0.0124	0.0108	(0.0403)	(0.0387)	(0.0109)	(0.0102)
Large	Customers	Large	ACF	0.006	-0.0291	0.0003	-0.0096	0.0086	0.0043	0.009	-0.0146	0.0066	-0.006
				0.0065	0.013	0.0052	0.0085	0.0054	0.0066	(0.0058)	(0.01)	(0.0026)	(0.0042)
			ACF-N/ND	0.0047	-0.0317	-0.0008	-0.0101	0.0094	0.0069	0.009	-0.0145	0.0062	-0.0064
				0.0064	0.0138	0.0052	0.009	0.0055	0.0075	(0.0059)	(0.0105)	(0.0026)	(0.0044)
		Small	ACF	0.009	-0.0268	-0.0008	-0.0106	0.0083	0.0042	0.0089	-0.0148	0.007	-0.0057
				0.0066	0.0133	0.0052	0.0086	0.0054	0.0065	(0.0057)	(0.0102)	(0.0026)	(0.0042)
			ACF-N/ND	0.0097	-0.0289	-0.0017	-0.0127	0.008	0.0049	0.0079	-0.0152	0.0066	-0.0066
				0.0065	0.0141	0.0052	0.0091	0.0055	0.0074	(0.0056)	(0.0108)	(0.0025)	(0.0044)
	Suppliers	Large	ACF	0.0083	-0.0281	-0.0007	-0.0107	0.0089	0.0046	0.0097	-0.0141	0.007	-0.0058
				0.0065	0.013	0.0052	0.0084	0.0054	0.0067	(0.0059)	(0.0101)	(0.0026)	(0.0042)
			ACF-N/ND	0.0059	-0.0311	-0.002	-0.0103	0.0108	0.0089	0.0106	-0.0137	0.0066	-0.0059
				0.0063	0.0138	0.0051	0.009	0.0056	0.0077	(0.0063)	(0.0105)	(0.0026)	(0.0044)
		Small	ACF	0.0081	-0.0252	0.001	-0.0091	0.009	0.0062	0.0098	-0.0138	0.0075	-0.0047
				0.0065	0.0129	0.0053	0.0088	0.0055	0.0068	(0.0059)	(0.01)	(0.0026)	(0.0042)
			ACF-N/ND	0.0057	-0.0279	-0.0003	-0.009	0.0107	0.0102	0.0106	-0.0135	0.0071	-0.0049
				0.0062	0.0137	0.0052	0.0095	0.0057	0.008	(0.0063)	(0.0104)	(0.0026)	(0.0044)
Small	Customers	Large	ACF	-0.0085	-0.0265	-0.0028	-0.0253	-0.0018	-0.0159	0.0065	-0.0222	0.0003	-0.0166
				0.0108	0.022	0.0092	0.0123	0.0066	0.0103	(0.0064)	(0.0114)	(0.0041)	(0.0059)
			ACF-N/ND	-0.0102	-0.0278	-0.004	-0.0249	-0.0003	-0.0107	0.0067	-0.0221	-0.0	-0.0169
				0.011	0.023	0.0091	0.0129	0.0066	0.0104	(0.0066)	(0.0119)	(0.0041)	(0.006)
		Small	ACF	0.0038	-0.0183	0.0003	-0.0176	0.0048	0.017	0.0139	-0.0104	0.0076	-0.004
				0.0104	0.0169	0.0066	0.01	0.0097	0.0106	(0.0085)	(0.0123)	(0.0037)	(0.0052)
			ACF-N/ND	0.0044	-0.0192	-0.0004	-0.0196	0.0049	0.0178	0.0128	-0.0106	0.0074	-0.0049
				0.0105	0.0175	0.0066	0.0106	0.0097	0.0109	(0.0084)	(0.0127)	(0.0037)	(0.0053)
	Suppliers	Large	ACF	0.0452	0.0142	0.0001	-0.0176	0.0066	-0.0079	0.0063	-0.0227	0.0084	-0.0098
				0.0217	0.0249	0.0077	0.01	0.0071	0.0094	(0.0068)	(0.0113)	(0.0038)	(0.0053)
			ACF-N/ND	0.0458	0.0136	-0.0009	-0.0186	0.0068	-0.0053	0.0059	-0.0228	0.0081	-0.0105
				0.0221	0.0256	0.0076	0.0104	0.0071	0.0098	(0.0068)	(0.0119)	(0.0038)	(0.0053)
		Small	ACF	0.0183	-0.0096	0.0037	-0.0191	0.0109	0.024	0.0075	-0.0194	0.0095	-0.005
				0.0118	0.0175	0.0066	0.0101	0.0091	0.0142	(0.0082)	(0.0115)	(0.0036)	(0.0052)
			ACF-N/ND	0.0183	-0.0103	0.0027	-0.021	0.0109	0.0244	0.0071	-0.0195	0.0092	-0.0058
				0.012	0.019	0.0065	0.0107	0.0092	0.0143	(0.0083)	(0.012)	(0.0035)	(0.0054)
Network Differencing				No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a value-added production function (in logs) estimated with the standard Ackerberg et al. (2015) procedure (ACF), or with modifications to accommodate network effects (ACF-N) and network differencing (ACF-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses.All specifications include industry and year fixed effects.

Table 19: Productivity Spillovers by Varying Firm Size Cutoffs (Gross Output)

Dependent Variable: $\ln\text{TFP}_t$						
		Large Firm Employees Cutoff:	500	1000	5000	Median
Neighbor Size	Relationship	Firm Size				
	$\ln\text{TFP}_{t-1}$		0.8982 (0.0044)	0.898 (0.0044)	0.8979 (0.0044)	0.8979 (0.0044)
Large	Customers	Large	0.006 (0.0025)	0.0059 (0.0025)	0.0059 (0.0025)	0.0054 (0.0025)
		Small	0.006 (0.0025)	0.0067 (0.0025)	0.0065 (0.0025)	0.0054 (0.0025)
	Suppliers	Large	0.0102 (0.0026)	0.0102 (0.0026)	0.0104 (0.0027)	0.0112 (0.0027)
		Small	0.0114 (0.0027)	0.0113 (0.0027)	0.0105 (0.0027)	0.0103 (0.0027)
Small	Customers	Large	-0.0011 (0.0057)	0.0043 (0.0044)	0.0051 (0.0034)	0.0059 (0.0027)
		Small	0.0064 (0.0036)	0.0079 (0.0033)	0.009 (0.0027)	0.0078 (0.0026)
	Suppliers	Large	0.0047 (0.005)	0.0075 (0.0037)	0.0104 (0.0033)	0.0099 (0.0029)
		Small	0.0091 (0.0036)	0.0114 (0.0032)	0.0129 (0.0028)	0.0108 (0.0027)

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a gross output production function (in logs) estimated with the Gandhi et al. (2020) procedure modified to account for endogenous and correlated network effects (GNR-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects inDieye and Fortin (2017); Patacchini et al. (2017). Large firms are defined by having at least as many employees as the cutoffs above. The median cutoff is determined by industry and year. Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

Table 20: Productivity Spillovers by Varying Firm Size Cutoffs (Value-Added)

		Dependent Variable: $\ln\text{TFP}_t$				
		Large Firm Employees Cutoff:	500	1000	5000	Median
Neighbor Size	Relationship	Firm Size				
Large	$\ln\text{TFP}_{t-1}$		0.8739 (0.0102)	0.8733 (0.0102)	0.8699 (0.0102)	0.8713 (0.0102)
Large	Customers	Large	-0.0064 (0.0044)	-0.0058 (0.0043)	-0.0066 (0.0044)	-0.0074 (0.0045)
Large	Customers	Small	-0.0066 (0.0044)	-0.0053 (0.0043)	-0.0046 (0.0043)	-0.0063 (0.0045)
Large	Suppliers	Large	-0.0059 (0.0044)	-0.0055 (0.0043)	-0.0071 (0.0044)	-0.0072 (0.0045)
Large	Suppliers	Small	-0.0049 (0.0044)	-0.0048 (0.0043)	-0.0063 (0.0043)	-0.007 (0.0045)
Small	Customers	Large	-0.0169 (0.006)	-0.0103 (0.0048)	-0.0058 (0.0042)	-0.0067 (0.0042)
Small	Customers	Small	-0.0049 (0.0053)	-0.005 (0.0046)	-0.0035 (0.0041)	-0.0039 (0.0042)
Small	Suppliers	Large	-0.0105 (0.0053)	-0.0052 (0.0046)	-0.0021 (0.0042)	-0.0019 (0.0043)
Small	Suppliers	Small	-0.0058 (0.0054)	-0.0034 (0.0046)	-0.0009 (0.0041)	-0.0036 (0.0043)

This table reports coefficients of a linear productivity evolution process with endogenous network effects estimated on US firms in Compustat. Each TFP measure is from a value-added production function (in logs) estimated with the standard Ackerberg et al. (2015) procedure modified to account for endogenous and correlated network effects (ACF-ND). Network effects are estimated using the generalized 2SLS procedure for heterogenous peer effects inDieye and Fortin (2017); Patacchini et al. (2017). Large firms are defined by having at least as many employees as the cutoffs above. The median cutoff is determined by industry and year. Heteroskedasticity-robust standard errors from the second step of the G2SLS estimator are in parentheses. All specifications include industry and year fixed effects.

9 Concluding Remarks

This paper examines how efficiency gains are transmitted through vertical relationships. The existence of spillovers implies a form of firm interdependence that matters for consistent estimation of production functions. Using Monte Carlo experiments, I show that endogenous spillovers—the effect of the average productivity of a firm's neighbors on its own productivity—are an important source of bias when not accounted for in TFP estimation. Furthermore, the direction of this bias cannot always be clearly predicted *a priori*, and varies by the density of the network and the persistence of productivity. However, for moderately sparse networks, standard approaches may deliver reasonably unbiased estimates of production function elasticities and spillovers.

Under additional assumptions on firms' information sets and the structure of the network, I propose a methodology that can flexibly accommodate various network effects and endogenous network formation, and can be applied to both gross output and value-added production functions. I show experimentally that it performs better than standard approaches as long as the network is not too dense and productivity is sufficiently persistent.

Using data from *Compustat* on supplier-customer relationships in the US, I investigate the extent of productivity spillovers in from 1977 to 2016. I find that firms benefit from having more productive buyers and sellers, with both large and small suppliers having a larger effect than customers. Furthermore, the cumulative impact of spillovers over the 4 decades in the sample could mean a 20 percent difference in efficiency when compared to a no-spillover scenario.

The sectoral composition of the production network plays a large role in the size and transmission of productivity gains. Positive spillovers are largely driven by the period from 1997 to 2006, which was the onset of stretch of rapid expansion for firms in the sample and a shift in the centrality of manufacturing firms in the production network, to retailers and wholesalers. I find substantial heterogeneity in the size and spillovers between and within sectors, with manufacturing firms benefiting from efficiency gains from most sectors and wholesalers boosting other industries.

Estimates suggest that if the most connected firm in a given year was 10 percent more productive, spillovers would lead to an increase in aggregate TFP of 2 to 4 percent. This also works in the opposite direction: a significant decline in productivity of central firms could mean substantial second-order impacts to US aggregate efficiency due to the interdependence of firms' activities through supply chains. This suggests that industrial and trade policies that could potentially affect the productivity of well-connected firms needs to account for potential indirect effects both upstream and downstream.

Consistent with my Monte Carlo experiments, estimates from standard approaches empirically yielded estimates of network effects that were similar to those obtained from my procedure, because the observed network density fell within the region where bias in spillover estimates was minimized. While this is reassuring for studies conducted on networks with similar levels of sparsity, caution should be taken when networks are much sparser or denser.

Results differed between gross output and value-added specifications. As discussed in Gandhi et al. (2017), value-added and gross output productivity measures may vary significantly and lead to substantively different policy implications about the dispersion of firm productivity. This study reveals that the choice of production function also matters for the estimation of productivity spillovers.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Alfaro-Urena, A., I. Manelici, and J. P. Vasquez (2019). The effects of joining multinational supply chains: New evidence from firm-to-firm linkages. *Working Paper*.

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

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).

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.

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

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

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

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*, 41–55.

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

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

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

Dieye, R. and B. Fortin (2017). Gender peer effects heterogeneity in obesity.

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.

Gandhi, A., S. Navarro, and D. Rivers (2017). How heterogeneous is productivity? a comparison of gross output and value added.

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.

Garcia-Marin, A. and N. Voigtländer (2019). Exporting and plant-level efficiency gains: It's in the measure. *Journal of Political Economy*.

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.

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

Hsieh, C.-S. and L. F. Lee (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics 31*(2), 301–319.

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

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.

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

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

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.

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.

Lee, L. F. and J. Yu (2016). Identification of spatial durbin panel models. *Journal of Applied Econometrics 31*(1), 133–162.

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

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

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

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

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

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

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

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.

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

Snijders, T. A. and S. P. Borgatti (1999). Non-parametric standard errors and tests for network statistics. *Connections 22*(2), 161–170.

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

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *The Journal of Industrial Economics 55*(3), 529–569.

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

Appendices

A Biases

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

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

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:

$$\Delta^G y_t = \alpha_\ell \Delta^G \ell_t + \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{61}$$

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:

$$\Delta y_t = \alpha_\ell \Delta \ell_t + \alpha_k \Delta k_t + u_t \tag{62}$$

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 (62), 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 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 case where the labor elasticity is estimated in the first stage, the second stage is equivalent to estimating:

$$\Delta^G \widetilde{y}_t = \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{63}$$

$$\tag{64}$$

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:

$$plim\ \widehat{\alpha}_k = \frac{cov(\Delta k_t, \Delta \widetilde{y}_t)}{var(\Delta k_t)} \tag{65}$$

$$plim\ = \alpha_k \left(1 - \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s k_{t-1})}{var(\Delta k_t)}\right) + \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s \widetilde{y}_{t-1})}{var(\Delta k_t)} \tag{66}$$

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, and

$$\widehat{\omega}_t = \widehat{\varphi}_t - \widehat{\alpha}_\ell \ell_t - \widehat{\alpha}_k k_t \tag{67}$$

$$plim\ \widehat{\omega}_t = \varphi_t - \alpha_\ell \ell_t - \alpha_k k_t = \omega_t \tag{68}$$

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

$$\widetilde{\omega}_t = \widehat{\varphi}_t - \widetilde{\alpha}_\ell \ell_t - \widetilde{\alpha}_k k_t = \widehat{\varphi}_t - \widehat{\alpha}_\ell \ell_t - \widehat{\alpha}_k k_t - (\alpha_\ell^{err} \ell_t + \alpha_k^{err} k_t) = \widehat{\omega}_t - \omega_t^{err} \tag{69}$$

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:[27] The true model is:

$$\omega_t = \rho \omega_{t-1} + \lambda G_t \omega_t + \zeta_t$$

but we estimate:

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

B 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.

[27] 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.

B.1 Production Function

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

$$Y_{it} = \min\{L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\alpha_1+\omega_{it}}, \alpha_m M_{it}\} e^{\varepsilon_{it}} \tag{70}$$

$$\implies Y_{it} = L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\alpha_1+\omega_{it}+\varepsilon_{it}} = \alpha_m M_{it} e^{\varepsilon_{it}} \tag{71}$$

$$\text{In logs, } y_{it} = \alpha_1 + \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it} \tag{72}$$

where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. I set $\alpha_1 = 0, \alpha_\ell = 0.6, \alpha_k = 0.4$ and $\sigma_\varepsilon^2 = 1$

B.2 Productivity Process and Network

Productivity evolves according to an AR1 process that allows for contemporaneous endogenous productivity spillovers. For ease of notation, I write the equation in vectorized form:

$$\omega_t = \beta_1 \iota + \rho \omega_{t-1} + \lambda G_t \omega_t + \zeta_t \tag{73}$$

where $\zeta_{it} \sim \mathcal{N}(0, \sigma_\zeta^2)$. I set $\sigma_\zeta^2 = 5$. I generate productivity using the reduced form of the above equation:

$$\omega_t = (\mathrm{I} - \lambda G_t)^{-1} (\beta_1 \iota + \rho \omega_{t-1} + \zeta_t) \tag{74}$$

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 $A_{ijt} \overset{i.i.d.}{\sim} \text{Bern}(p)$.

B.3 Intermediate Input Demand

$$M_{it} = \frac{1}{\alpha_m} K_{it}^{\alpha_k} L_{it}^{\alpha_\ell} e^{\alpha_1+\omega_{it}} \tag{75}$$

$$\text{In logs, } m_{it} = \alpha_1 + \alpha_k k_{it} + \alpha_\ell \ell_{it} + \omega_{it} - \ln(\alpha_m) \tag{76}$$

B.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:

$$L_{it} = \left(\alpha_\ell \frac{K_{it}^{\alpha_k}}{W_{it}} e^{\alpha_1+\omega_{it}}\right)^{\frac{1}{1-\alpha_\ell}} \tag{77}$$

$$\text{In logs, } \ell_{it} = \frac{1}{1-\alpha_\ell} \left(\ln(\alpha_\ell) + \alpha_1 + \alpha_k k_{it} + \omega_{it} - \ln(W_{it})\right) \tag{78}$$

B.5 Capital and Optimal Investment

Capital is accumulated as follows:

$$K_{it} = (1-\delta)K_{it-1} + I_{t-1} \tag{79}$$

I set the depreciation rate at $\delta = 0.2$.

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:[28]

$$\Pi_{it} = L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\alpha_1+\omega_{it}} - W_{it}L_{it} - \frac{b}{2}I_{it}^2 \tag{80}$$

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

$$V(L_{it}, K_{it}, W_{it}, \omega_{it}) = \max_{L_{it},K_{it}} L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\alpha_1+\omega_{it}} - W_{it}L_{it} - \frac{b}{2}I_{it}^2 + \beta\, \mathbb{E}_{it}\, V(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \tag{81}$$

$$\text{such that } K_{it+1} = (1-\delta)K_{it} + I_t \tag{82}$$

β is the discount factor and is fixed at 0.95. Optimal investment solves the Euler equation $\dfrac{\partial V}{\partial I} = 0$:

$$bI_{it} = \beta\, \mathbb{E}_{it}\, V_K(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \tag{83}$$

The envelope condition yields:

$$V_K(L_{it}, K_{it}, W_{it}, \omega_{it}) = \alpha_k L_{it}^{\alpha_\ell} K_{it}^{\alpha_k - 1} e^{\alpha_1+\omega_{it}} + \beta(1-\delta)\, \mathbb{E}_{it}\, V_K(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \tag{84}$$

Substituting in (77) and (83):

$$V_K(L_{it}, K_{it}, W_{it}, \omega_{it}) = \alpha_k \alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} K_{it}^{\frac{\alpha_k+\alpha_\ell-1}{1-\alpha_\ell}} W_{it}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\alpha_1+\omega_{it}}{1-\alpha_\ell}} + b(1-\delta)I_{it} \tag{85}$$

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

$$I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\alpha_1}{1-\alpha_\ell}}\, \mathbb{E}_{it}\left[W_{it+1}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it+1}}{1-\alpha_\ell}}\right] + \beta(1-\delta)\, \mathbb{E}_{it}\, I_{it+1} \tag{86}$$

$$\implies I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\alpha_1}{1-\alpha_\ell}} \sum_{\tau=0}^{\infty} \beta^\tau (1-\delta)^\tau\, \mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] \tag{87}$$

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

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 $\mathbb{E}_{it}\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right]$ depends on the productivity process:

$$\begin{aligned}
\omega_{t+1+\tau} =& \rho(I - \lambda G_{t+\tau+1})^{-1}\omega_{t+\tau} + (I - \lambda G_{t+\tau+1})^{-1}\varepsilon_{t+1+\tau} \\
=& \rho^2(I - \lambda G_{t+\tau+1})^{-1}(I - \lambda G_{t+\tau})^{-1}\omega_{it+\tau-1} + \rho(I - \lambda G_{t+\tau+1})^{-1}(I - \lambda W_{t+\tau})^{-1}\varepsilon_{t+\tau} \\
&+ (I - \lambda G_{t+\tau+1})^{-1}\varepsilon_{t+\tau+1} \\
\omega_{t+1+\tau} =& \rho^{\tau+1}\prod_{r=0}^{\tau}(I - \lambda G_{t+\tau+1-r})^{-1}\omega_t + \sum_{r=0}^{\tau}\rho^r\prod_{s=0}^{r}(I - \lambda G_{t+\tau+1-s})^{-1}\varepsilon_{t+\tau+1-r}
\end{aligned} \tag{88}$$

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

$$\mathbb{E}_{it}\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] = \mathbb{E}_{it}\left[\exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell} + \frac{1}{1-\alpha_\ell}\sum_{r=0}^{\tau}\rho^r\varepsilon_{t+\tau+1-r}\right)\right] \tag{89}$$

$$= \exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell}\right)\prod_{r=0}^{\tau}\mathbb{E}_{it}\left[\exp\left(\frac{\rho^r\varepsilon_{it+\tau+1-r}}{1-\alpha_\ell}\right)\right] \tag{90}$$

$$= \exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell}\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2}\right) \tag{91}$$

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 $G_{t+1} \in \mathcal{I}_t$:

$$\begin{aligned}
\mathbb{E}_{it}\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] &= \mathbb{E}_{it}\left[\exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I - \lambda\overline{G})^{-(\tau+1)}\omega_t + \frac{1}{1-\alpha_\ell}\sum_{r=0}^{\tau}\rho^r(I - \lambda\overline{G})^{-(r+1)}\varepsilon_{t+\tau+1-r}\right)\right] \\
&= \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I - \lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\mathbb{E}_{it}\left[\exp\left(\frac{\rho^r}{1-\alpha_\ell}(I - \lambda\overline{G})^{-(r+1)}\varepsilon_{t+\tau+1-r}\right)\right] \\
&= \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I - \lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2}(I - \lambda\overline{G})^{-2(r+1)}\iota\right) \\
\mathbb{E}_{it}\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] &= \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I - \lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2(1-\lambda)^{2(r+1)}}\right)
\end{aligned} \tag{92}$$

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

$$I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{\alpha_1}{1-\alpha_\ell} + \frac{\alpha_\ell^2\sigma_w^2}{2(1-\alpha_\ell)^2}\right) \tag{93}$$
$$\times \sum_{\tau=0}^{\infty} \beta^\tau(1-\delta)^\tau \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t + \frac{\sigma_\zeta^2}{2(1-\alpha_\ell)^2(1-\lambda)^2}\sum_{r=0}^{\tau}\left(\frac{\rho}{1-\lambda}\right)^{2r}\right)$$

When there are no spillovers, this reduces to:

$$I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{\alpha_1}{1-\alpha_\ell} + \frac{\alpha_\ell^2\sigma_w^2}{2(1-\alpha_\ell)^2}\right)\sum_{\tau=0}^{\infty}\beta^\tau(1-\delta)^\tau \exp\left(\frac{\rho^{\tau+1}\omega_t}{1-\alpha_\ell} + \frac{\sigma_\zeta^2\sum_{r=0}^{\tau}\rho^{2r}}{2(1-\alpha_\ell)^2}\right) \tag{94}$$

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 $\mathbb{E}_{it}\left[exp\left(\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}\right)\right] = \frac{1}{100}\sum_{s=0}^{100} exp\left(\frac{\omega_{it+1+\tau,s}}{1-\alpha_\ell}\right)$.

C Additional Monte Carlo Experiments

In this section, I consider how bias and precision change with the size of the endogenous network effect, and the persistence of productivity over time. The Monte Carlo setup is the same as in section 6.3.

When I vary λ, ACF and ACF-N perform similarly, yielding comparable estimates of the input elasticities and spillover effects. Low values of λ are difficult to detect, while at very high values, there is a sharp decline in efficiency, with the decline greater under ACF. Under the linear process with negative spillovers, ACF-N appears to perform better than ACF as λ rises in magnitude.

Finally, variations in ρ_1 have striking effects on the estimation of λ because it determines the strength of $G_t\omega_{t-1}$ as an instrument for $G_t\omega_t$. Intuitively, if productivity is not persistent, then neighbors' lagged TFP is a weak instrument for the contemporaneous effect of neighbors's productivity, because the intertemporal correlation is not strong. ACF-N is not immune to this issue, and loses efficiency in its estimates of λ unless ρ_1 is sufficiently high. However the input elasticities are relatively well estimated by ACF-N, while ACF leads to biased estimates for high values of ρ_1: overestimating the capital coefficient when $\rho_1 = 0.8$ and underestimating it when $\rho_1 = 0.9$.

Table 21: Effect of λ on Bias and Precision (Quadratic AR1)

λ	Estimator		Elasticities		Productivity Process Coefficients			
			α_ℓ	α_k	β_1	ρ_1	ρ_2	λ
		True	0.6	0.4	0.5	0.8	-0.01	
0.01	ACF	Mean	0.603	0.353	-0.198	0.809	-0.01	0.
		Std. Dev.	(0.024)	(0.251)	(2.706)	(0.23)	(0.003)	(0.061)
	ACF-N	Mean	0.608	0.355	-0.038	0.811	-0.009	-0.02
		Std. Dev.	(0.037)	(0.234)	(2.394)	(0.212)	(0.027)	(0.337)
0.03	ACF	Mean	0.603	0.353	-0.182	0.809	-0.01	0.019
		Std. Dev.	(0.024)	(0.247)	(2.457)	(0.224)	(0.003)	(0.061)
	ACF-N	Mean	0.608	0.354	-0.03	0.812	-0.01	-0.009
		Std. Dev.	(0.035)	(0.23)	(2.218)	(0.206)	(0.003)	(0.686)
0.05	ACF	Mean	0.603	0.354	-0.16	0.808	-0.01	0.038
		Std. Dev.	(0.024)	(0.242)	(2.212)	(0.218)	(0.003)	(0.062)
	ACF-N	Mean	0.608	0.353	-0.027	0.813	-0.01	0.038
		Std. Dev.	(0.034)	(0.225)	(2.002)	(0.201)	(0.003)	(0.27)
0.07	ACF	Mean	0.604	0.354	-0.132	0.808	-0.01	0.057
		Std. Dev.	(0.024)	(0.237)	(1.978)	(0.212)	(0.003)	(0.062)
	ACF-N	Mean	0.607	0.354	0.013	0.814	-0.01	0.055
		Std. Dev.	(0.033)	(0.219)	(1.746)	(0.193)	(0.003)	(0.143)
0.09	ACF	Mean	0.604	0.357	-0.099	0.805	-0.01	0.076
		Std. Dev.	(0.024)	(0.23)	(1.759)	(0.204)	(0.003)	(0.063)
	ACF-N	Mean	0.607	0.355	0.044	0.814	-0.01	0.071
		Std. Dev.	(0.033)	(0.212)	(1.546)	(0.184)	(0.003)	(0.136)
0.1	ACF	Mean	0.604	0.36	-0.08	0.803	-0.01	0.086
		Std. Dev.	(0.024)	(0.226)	(1.65)	(0.2)	(0.003)	(0.063)
	ACF-N	Mean	0.607	0.356	0.059	0.814	-0.01	0.08
		Std. Dev.	(0.033)	(0.208)	(1.443)	(0.179)	(0.003)	(0.151)
0.3	ACF	Mean	0.606	0.643	0.054	0.414	-0.024	0.446
		Std. Dev.	(0.072)	(0.152)	(3.357)	(0.782)	(0.151)	(1.833)
	ACF-N	Mean	0.605	0.388	0.417	0.813	-0.01	0.299
		Std. Dev.	(0.032)	(0.084)	(0.399)	(0.065)	(0.004)	(0.062)
0.5	ACF	Mean	0.645	0.362	-5.659	0.723	-0.012	0.651
		Std. Dev.	(0.151)	(0.156)	(270.251)	(0.645)	(0.063)	(8.7)
	ACF-N	Mean	0.648	0.351	-3.503	0.77	-0.012	0.704
		Std. Dev.	(0.1)	(0.101)	(204.646)	(0.311)	(0.034)	(6.386)
0.7	ACF	Mean	0.682	0.318	-3.839	0.668	-0.012	1.148
		Std. Dev.	(0.19)	(0.19)	(217.5)	(4.683)	(0.616)	(11.871)
	ACF-N	Mean	0.685	0.314	4.003	0.592	-0.001	0.671
		Std. Dev.	(0.144)	(0.144)	(91.859)	(0.669)	(0.099)	(2.519)

Based on 1000 replications. Estimators are based on Ackerberg et al. (2015) with ACF denoting the standard procedure and ACF-N indicating the modified procedure to account for network effects. Networks are exogenous erdos-renyi (binomial) graphs with 0.05 density. The data-generating process for productivity is quadratic AR1 with endogenous network effects.

Table 22: Effect of ρ on Bias and Precision (Quadratic AR1)

ρ	Estimator		Elasticities		Productivity Process Coefficients			
			α_ℓ	α_k	β_1	ρ_1	ρ_2	λ
		True	0.6	0.4	0.5		-0.01	0.3
0.1	ACF	Mean	0.644	0.413	-271.942	0.485	-0.047	52.923
		Std. Dev.	(0.138)	(0.384)	(8697.007)	(12.757)	(1.038)	(1683.731)
	ACF-N	Mean	0.65	0.39	2.029	0.113	-0.009	0.151
		Std. Dev.	(0.1)	(0.365)	(90.035)	(0.514)	(0.041)	(10.864)
0.2	ACF	Mean	0.634	0.362	3.014	0.255	-0.003	-0.094
		Std. Dev.	(0.131)	(0.234)	(30.083)	(1.139)	(0.109)	(5.893)
	ACF-N	Mean	0.647	0.343	-0.796	0.205	-0.01	0.14
		Std. Dev.	(0.099)	(0.219)	(37.532)	(0.127)	(0.031)	(7.268)
0.3	ACF	Mean	0.628	0.369	1.427	0.286	-0.001	0.069
		Std. Dev.	(0.109)	(0.161)	(6.855)	(0.587)	(0.166)	(3.382)
	ACF-N	Mean	0.645	0.345	-0.132	0.312	-0.009	1.074
		Std. Dev.	(0.095)	(0.151)	(31.101)	(0.103)	(0.03)	(24.603)
0.4	ACF	Mean	0.617	0.381	1.412	0.314	0.028	-0.083
		Std. Dev.	(0.086)	(0.122)	(15.72)	(2.994)	(1.188)	(11.549)
	ACF-N	Mean	0.638	0.354	0.251	0.409	-0.01	0.45
		Std. Dev.	(0.084)	(0.118)	(21.734)	(0.086)	(0.028)	(8.122)
0.5	ACF	Mean	0.609	0.389	0.831	0.518	-0.013	0.268
		Std. Dev.	(0.061)	(0.094)	(1.139)	(0.193)	(0.104)	(0.935)
	ACF-N	Mean	0.628	0.364	1.104	0.512	-0.011	0.105
		Std. Dev.	(0.072)	(0.103)	(13.659)	(0.066)	(0.022)	(7.008)
0.6	ACF	Mean	0.606	0.398	0.758	0.609	-0.009	0.253
		Std. Dev.	(0.044)	(0.082)	(0.549)	(0.053)	(0.017)	(0.146)
	ACF-N	Mean	0.619	0.373	1.087	0.613	-0.01	0.186
		Std. Dev.	(0.057)	(0.09)	(6.734)	(0.064)	(0.016)	(1.727)
0.7	ACF	Mean	0.601	0.465	0.575	0.665	-0.012	0.29
		Std. Dev.	(0.043)	(0.086)	(0.38)	(0.081)	(0.014)	(0.1)
	ACF-N	Mean	0.608	0.385	0.642	0.714	-0.01	0.279
		Std. Dev.	(0.036)	(0.081)	(2.109)	(0.053)	(0.005)	(0.47)
0.8	ACF	Mean	0.606	0.643	0.054	0.414	-0.024	0.446
		Std. Dev.	(0.072)	(0.152)	(3.357)	(0.782)	(0.151)	(1.833)
	ACF-N	Mean	0.605	0.388	0.417	0.813	-0.01	0.299
		Std. Dev.	(0.032)	(0.084)	(0.399)	(0.065)	(0.004)	(0.062)
0.9	ACF	Mean	0.708	0.113	-5.964	0.65	-0.028	0.346
		Std. Dev.	(0.093)	(0.273)	(33.081)	(4.953)	(0.367)	(4.293)
	ACF-N	Mean	0.603	0.386	0.024	0.922	-0.01	0.296
		Std. Dev.	(0.028)	(0.09)	(1.165)	(0.122)	(0.002)	(0.033)

Based on 1000 replications. Estimators are based on Ackerberg et al. (2015) with ACF denoting the standard procedure and ACF-N indicating the modified procedure to account for network effects. Networks are exogenous erdos-renyi (binomial) graphs with 0.05 density. The data-generating process for productivity is quadratic AR1 with endogenous network effects.

D 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. Figure 17 shows that 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.
- Exports: International Sales as reported in the geographic segments information on annual reports. These figures are often reported by location of the final customer, but do not always differentiate between exports from the US and sales by multinational firms within foreign countries. However, to the extent that this contains some measure of exporting, a dummy for exporting based on positive values of this variable should have minimal measurement error.
- 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: Farms combined with forestry and fishing; 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 54 industry groups.

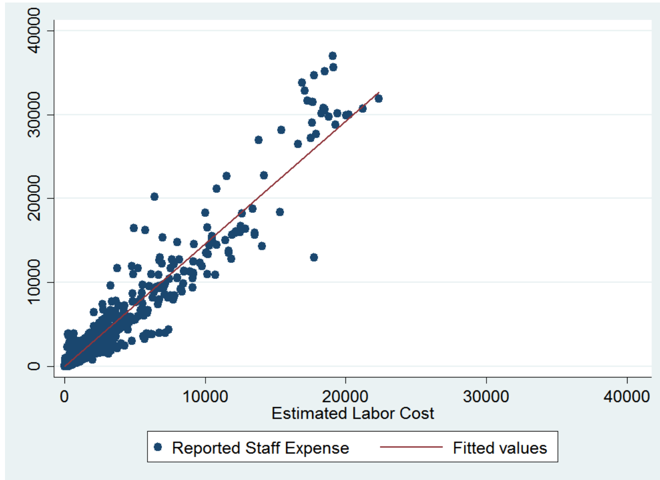


Figure 17: Estimated and Reported Labor Expenses

This figure shows the correlation between labor expenses reported in *Compustat* and labor costs estimated using industry wage expenditure.

E Boostrapping Network Data

E.1 Residual-based resampling

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 pseudo-samples, 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.

E.2 Vertex Resampling

An alternative procedure is the vertex bootstrap introduced by Snijders and Borgatti (1999). Although this method is potentially more robust to model misspecification, the resulting adjacency matrices are not guaranteed to satisfy the linear independence conditions for consistency of the G2SLS peer effects estimator. The procedure is as follows: Let M be the set of unique firms across all years in the data, with cardinality m and let R be the number of bootstrap repetitions.

For each bootstrap repetition r, randomly select m firms from M with replacement to form a bootstrap sample M_r. Each firm k in M_r corresponds to a firm $i(k) \in M$; I include observations from all years in which $i(k)$ appears in the original dataset. This is the standard block bootstrapping procedure for panel data, which maintains the dependence structure across time within a firm.

Next, for each year, construct the adjacency matrix A_{rt} from the original A_t. Every pair of firms (k, h) in M_r corresponds to $(i(k), i(h))$ in M. Therefore, if $i(k) \neq i(h)$, then we can set

$$A_{kh,rt} = A_{i(k)i(h),t}$$

However, A_t does not provide information on edges between duplicated nodes ($i(k) = i(h)$), because in the original network, a firm could not buy from itself. But in the boostrap sample, k and h are considered different firms. Therefore, I fill in these edges by uniformly sampling from all the elements of A_t. Finally, the interaction matrix G_{rt} is constructed by row-normalizing A_{rt}.