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

Ebehi Iyoha*

October 2020

Contents

*Contact: e.iyoha@vanderbilt.edu. Vanderbilt University, Department of Economics, VU Station B 351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819.

1 Derivation of Bias Terms

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

$$\Delta^G y_t = \alpha_\ell \Delta^G \ell_t + \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty}\lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{1}$$

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:

$$\Delta y_t = \alpha_\ell \Delta \ell_t + \alpha_k \Delta k_t + u_t \tag{2}$$

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 (2), 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:

$$\Delta^G \widetilde{y}_t = \alpha_k \Delta^G k_t + \sum_{s=0}^{\infty} \lambda^s G_t^s \zeta_t + \Delta^G \varepsilon_t \tag{3}$$

$$\tag{4}$$

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:

$$plim\ \widehat{\alpha}_k = \frac{cov(\Delta k_t, \Delta \widetilde{y}_t)}{var(\Delta k_t)} \tag{5}$$

$$plim\ = \alpha_k \left(1 - \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s k_{t-1})}{var(\Delta k_t)} \right) + \rho \sum_{s=1}^{\infty} \lambda^s \frac{cov(\Delta k_t, G_t^s \widetilde{y}_{t-1})}{var(\Delta k_t)} \tag{6}$$

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

$$\widehat{\omega}_t = \widehat{\varphi}_t - \widehat{\alpha}_\ell \ell_t - \widehat{\alpha}_k k_t \tag{7}$$

$$plim\ \widehat{\omega}_t = \varphi_t - \alpha_\ell \ell_t - \alpha_k k_t = \omega_t \tag{8}$$

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

$$\widetilde{\omega}_t = \widehat{\varphi}_t - \widetilde{\alpha}_\ell \ell_t - \widetilde{\alpha}_k k_t = \widehat{\varphi}_t - \widehat{\alpha}_\ell \ell_t - \widehat{\alpha}_k k_t - (\alpha_\ell^{err} \ell_t + \alpha_k^{err} k_t) = \widehat{\omega}_t - \omega_t^{err} \tag{9}$$

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:[1] The true model is:

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

but we estimate:

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

2 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.

[1] 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.

2.1 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.[2]

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 $V_i(A_{ij,t}=1) - V_i(A_{ij,t}=0) > 0$.[3] 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.

$$U_{ijt}(\gamma) = \gamma_1 + \boldsymbol{z_{it}\gamma_i} + \boldsymbol{z_{jt}\gamma_j} + \boldsymbol{z_{ijt}\gamma_{ij}} + \gamma_h H_{ijt} \tag{10}$$

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 $|\boldsymbol{z_{it}} - \boldsymbol{z_{jt}}|$ 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.[4]

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

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

[2]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 allows me to explicitly highlight the dual role that productivity may play in search and interfirm spillovers. 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.

[3]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.

[4]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.

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:

$$\omega_t = h(\omega_{t-1}, \boldsymbol{x_{t-1}}, G_t\boldsymbol{x_{t-1}}) + \lambda G_t\omega_t + \Xi_t\boldsymbol{\delta} + \nu_t \tag{11}$$

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

2.2 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:

$$\begin{aligned} \mu_{it} &= \sum_{j\neq i}^{N_t} g_{ij,t}\frac{\phi(\Phi^{-1}(p))}{\Phi(\Phi^{-1}(p))} + (1-g_{ij,t})\frac{\phi(\Phi^{-1}(p))}{1-\Phi(\Phi^{-1}(p))} \\ &= \sum_{j\neq i}^{N_t} g_{ij,t}\frac{\phi(\Phi^{-1}(p))}{p} + (1-g_{ij,t})\frac{\phi(\Phi^{-1}(p))}{1-p} \end{aligned} \tag{12}$$

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 $\delta = \frac{\sigma_{\zeta\xi}}{\sigma_\xi^2}$.

2.3 Estimation Procedure

Incorporating the selection model is similar to the Olley and Pakes (1996) correction for attrition. The first stage of my benchmark procedure is unchanged with the estimation of $\widehat{\varphi}_{it}$ and $\widehat{\varepsilon}_{it}$ using the proxy variable. In the second stage, starting with the initial guesses of the labor and capital coefficients $(\alpha_k^*, \alpha_\ell^*)$, compute $\omega_{it-1}^* = \widehat{\varphi}_{it-1} - \alpha_1^* - \alpha_k^* k_{it-1} - \alpha_\ell^* \ell_{it-1}$.

Using ω_{it-1}^* and other variables that could determine the observed links between firms, estimate the selection model in equation (10) 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:

$$\omega_t^* = \sum_{s=0}^{\infty} \lambda^s G_t^s h(\omega_{t-1}^*, \boldsymbol{x_{t-1}}, G_t \boldsymbol{x_{t-1}}) + \delta \sum_{s=0}^{\infty} \lambda^s G_t^s \mu_{it}^* + u_t \tag{13}$$

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 (30) for identification of the elasticities.[5]

3 Monte Carlo Experiments

3.1 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.

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

3.1.1 Production Function

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

$$Y_{it} = \min\{L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}}, \alpha_m M_{it}\} e^{\varepsilon_{it}} \tag{14}$$

$$\implies Y_{it} = L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}+\varepsilon_{it}} = \alpha_m M_{it} e^{\varepsilon_{it}} \tag{15}$$

$$\text{In logs, } y_{it} = \alpha_\ell \ell_{it} + \alpha_k k_{it} + \omega_{it} + \varepsilon_{it} \tag{16}$$

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

3.1.2 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:

$$\omega_t = \beta_1 \iota + \rho \omega_{t-1} + \lambda G_t \omega_t + \zeta_t \tag{17}$$

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:

$$\omega_t = (\mathrm{I} - \lambda G_t)^{-1} (\beta_1 \iota + \rho \omega_{t-1} + \zeta_t) \tag{18}$$

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

3.1.3 Intermediate Input Demand

$$M_{it} = \frac{1}{\alpha_m} K_{it}^{\alpha_k} L_{it}^{\alpha_\ell} e^{\omega_{it}} \tag{19}$$

$$\text{In logs, } m_{it} = \alpha_k k_{it} + \alpha_\ell \ell_{it} + \omega_{it} - \ln(\alpha_m) \tag{20}$$

3.1.4 Labor Demand

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

$$L_{it} = \left(\alpha_\ell \frac{K_{it}^{\alpha_k}}{W_{it}} e^{\omega_{it}}\right)^{\frac{1}{1-\alpha_\ell}} \tag{21}$$

$$\text{In logs, } \ell_{it} = \frac{1}{1-\alpha_\ell} \left(\ln(\alpha_\ell) + \alpha_k k_{it} + \omega_{it} - \ln(W_{it})\right) \tag{22}$$

3.1.5 Capital and Optimal Investment

Capital is accumulated as follows:

$$K_{it} = (1-\delta)K_{it-1} + I_{t-1} \tag{23}$$

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:[6]

$$\Pi_{it} = L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}} - W_{it}L_{it} - \frac{b}{2}I_{it}^2 \tag{24}$$

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

$$V(L_{it}, K_{it}, W_{it}, \omega_{it}) = \max_{L_{it}, K_{it}} L_{it}^{\alpha_\ell} K_{it}^{\alpha_k} e^{\omega_{it}} - W_{it}L_{it} - \frac{b}{2}I_{it}^2 + \beta\, \mathbb{E}_{it}\, V(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \tag{25}$$

$$\text{such that } K_{it+1} = (1-\delta)K_{it} + I_t \tag{26}$$

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

$$bI_{it} = \beta\, \mathbb{E}_{it}\, V_K(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \tag{27}$$

The envelope condition yields:

$$V_K(L_{it}, K_{it}, W_{it}, \omega_{it}) = \alpha_k L_{it}^{\alpha_\ell} K_{it}^{\alpha_k - 1} e^{\omega_{it}} + \beta(1-\delta)\, \mathbb{E}_{it}\, V_K(L_{it+1}, K_{it+1}, W_{it+1}, \omega_{it+1}) \tag{28}$$

Substituting in (21) and (27):

$$V_K(L_{it}, K_{it}, W_{it}, \omega_{it}) = \alpha_k \alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} K_{it}^{\frac{\alpha_k+\alpha_\ell-1}{1-\alpha_\ell}} W_{it}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it}}{1-\alpha_\ell}} + b(1-\delta)I_{it} \tag{29}$$

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

$$I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}}\, \mathbb{E}_{it}\left[W_{it+1}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it+1}}{1-\alpha_\ell}}\right] + \beta(1-\delta)\, \mathbb{E}_{it}\, I_{it+1} \tag{30}$$

$$\implies I_{it} = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \sum_{\tau=0}^{\infty} \beta^\tau (1-\delta)^\tau\, \mathbb{E}_{it}\left[W_{it+1+\tau}^{\frac{-\alpha_\ell}{1-\alpha_\ell}} e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] \tag{31}$$

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

$$\begin{aligned}
\omega_{t+1+\tau} =& \rho(I-\lambda G_{t+\tau+1})^{-1}\omega_{t+\tau} + (I-\lambda G_{t+\tau+1})^{-1}\varepsilon_{t+1+\tau} \\
=& \rho^2(I-\lambda G_{t+\tau+1})^{-1}(I-\lambda G_{t+\tau})^{-1}\omega_{it+\tau-1} + \rho(I-\lambda G_{t+\tau+1})^{-1}(I-\lambda W_{t+\tau})^{-1}\varepsilon_{t+\tau} \\
&+ (I-\lambda G_{t+\tau+1})^{-1}\varepsilon_{t+\tau+1} \\
\omega_{t+1+\tau} =& \rho^{\tau+1}\prod_{r=0}^{\tau}(I-\lambda G_{t+\tau+1-r})^{-1}\omega_t + \sum_{r=0}^{\tau}\rho^r\prod_{s=0}^{r}(I-\lambda G_{t+\tau+1-s})^{-1}\varepsilon_{t+\tau+1-r} \qquad (32)
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{E}_t\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] &= \mathbb{E}_t\left[\exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell} + \frac{1}{1-\alpha_\ell}\sum_{r=0}^{\tau}\rho^r\varepsilon_{t+\tau+1-r}\right)\right] &(33)\\
&= \exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell}\right)\prod_{r=0}^{\tau}\mathbb{E}_{it}\left[\exp\left(\frac{\rho^r\varepsilon_{it+\tau+1-r}}{1-\alpha_\ell}\right)\right] &(34)\\
&= \exp\left(\frac{\rho^{\tau+1}\omega_{it}}{1-\alpha_\ell}\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2}\right) &(35)
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{E}_t\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] &= \mathbb{E}_t\left[\exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t + \frac{1}{1-\alpha_\ell}\sum_{r=0}^{\tau}\rho^r(I-\lambda\overline{G})^{-(r+1)}\varepsilon_{t+\tau+1-r}\right)\right] \\
&= \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\mathbb{E}_{it}\left[\exp\left(\frac{\rho^r}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(r+1)}\varepsilon_{t+\tau+1-r}\right)\right] \\
&= \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2}(I-\lambda\overline{G})^{-2(r+1)}\iota\right) \\
\mathbb{E}_{it}\left[e^{\frac{\omega_{it+1+\tau}}{1-\alpha_\ell}}\right] &= \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t\right)\prod_{r=0}^{\tau}\exp\left(\frac{\rho^{2r}\sigma_\zeta^2}{2(1-\alpha_\ell)^2(1-\lambda)^{2(r+1)}}\right) &(36)
\end{aligned}$$

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

$$I_t = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{\alpha_\ell^2\sigma_w^2}{2(1-\alpha_\ell)^2}\right) \times \sum_{\tau=0}^{\infty}\beta^\tau(1-\delta)^\tau \exp\left(\frac{\rho^{\tau+1}}{1-\alpha_\ell}(I-\lambda\overline{G})^{-(\tau+1)}\omega_t + \frac{\sigma_\zeta^2}{2(1-\alpha_\ell)^2(1-\lambda)^2}\sum_{r=0}^{\tau}\left(\frac{\rho}{1-\lambda}\right)^{2r}\right) \tag{37}$$

When there are no spillovers, this reduces to:

$$I_t = \frac{\beta\alpha_k}{b}\alpha_\ell^{\frac{\alpha_\ell}{1-\alpha_\ell}} \exp\left(\frac{\alpha_\ell^2\sigma_w^2}{2(1-\alpha_\ell)^2}\right)\sum_{\tau=0}^{\infty}\beta^\tau(1-\delta)^\tau \exp\left(\frac{\rho^{\tau+1}\omega_t}{1-\alpha_\ell} + \frac{\sigma_\zeta^2\sum_{r=0}^{\tau}\rho^{2r}}{2(1-\alpha_\ell)^2}\right) \tag{38}$$

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

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:

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

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

3.2 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:

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

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

For DGPs 1 to 4, I generate an exogenous directed network in each period by randomly assigning links with probability $P(A_{ijt} = 1) = \frac{indshare_{ij}}{indsize_j}$ where $indshare_{ij}$ is the compatibility of i and j's industries obtained from the industry compatibility matrix described above, while $indsize_j$ is the number of firms in j's industry. DGP 1 has no network effects ($\lambda = 0, \beta_{\bar{x}} = 0, c_{\psi_t} = 0$) and exogenous network formation, and 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 ($\lambda = 0, \beta_{\bar{x}} = 0.3, c_{\psi_t} = 0$). 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 2 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 1. 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 1: 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.

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

Next, I compare the performance my estimator against standard ACF in table 2 using four DGPs. The Monte Carlo setup is essentially the same as in experiment 1 above. 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 2.3. 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 2: Comparison of Estimates from Standard ACF and Modified ACF Procedures

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

Based on 1000 replications. 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 3.1 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.

3.4 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.[7]. The quadratic term serves as a dampener to control the size ω_t in the simulation.[8] 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 5 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.

3.5 Experiment 4: Effect of Productivity Persistence and Endogenous Effect Size on Bias and Precision

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 3.4.

When I vary λ, ACF and ACF-N perform similarly, yielding comparable estimates of the input elas-

[7]See details on optimal investment in section 3.1.5 in the appendix

[8]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.

ticities 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 3: 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 4: 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.

Table 5: Effect of Sparsity on Bias and Precision (Quadratic AR1)

Density	Estimator	Elasticities α_ℓ	Elasticities α_k	Productivity Process Coefficients β_1	Productivity Process Coefficients ρ_1	Productivity Process Coefficients ρ_2	Productivity Process Coefficients 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.

4 Extensions

4.1 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.[9]

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

[9]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 $D^{\mathcal{E}}(\ell_t, k_t, m_t) \equiv \dfrac{\partial}{\partial m_t} f(\ell_t, k_t, m_t) \times \mathcal{E}$. 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.

4.1.1 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:[10]

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

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:

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

4.2 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.

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

4.2.1 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.

4.2.2 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.

5 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 1 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: transit and ground transportation with general transportation and warehousing, and other transportation and support activities; Funds, trusts and other financial vehicles combined with securities, commodity contracts and investments; Legal services with miscellaneous professional services; Ambulatory health, hospitals, nursing and residential care with social assistance. This results in 41 industry groups.

Table 6: Sample Size by Industry and Sector

Sector	Industry	Observations
Mining	Mining, except oil and gas	445
	Oil and gas extraction	2323
	Support activities for mining	691
Utilities	Utilities	2427
Construction	Construction	481
Durables Manufacturing	Electrical equipment, appliances, and components	1063
	Fabricated metal products	1190
	Furniture and related products	303
	Machinery	2571
	Miscellaneous manufacturing	1490
	Motor vehicles, bodies and trailers, and parts	1975
	Nonmetallic mineral products	368
	Other transportation equipment	1059
	Primary metals	902
	Wood products	192
Non-Durables Manufacturing	Apparel and leather and allied products	1410
	Chemical products	3959
	Food and beverage and tobacco products	1561
	Paper products	601
	Petroleum and coal products	1068
	Plastics and rubber products	778
	Printing and related support activities	223
	Textile mills and textile product mills	448
Electronics Manufacturing	Computer and electronic products	9581
Wholesale	Wholesale trade	2215
Retail	Food and beverage stores	347
	General merchandise stores	574
	Motor vehicle and parts dealers	153
	Other retail	1303
Transport and Warehousing	Air transportation	422
	Other transportation and support activities	265
	Pipeline transportation	487
	Rail transportation	238
	Transit and ground passenger transportation	14
	Transportation and warehousing	22
	Truck transportation	322
	Warehousing and storage	13
	Water transportation	284
Information	Broadcasting and telecommunications	2213
	Data processing, internet publishing, and other information services	633
	Motion picture and sound recording industries	292
	Publishing industries, except internet (includes software)	1788
FIRE	Federal Reserve banks, credit intermediation, and related activities	554
	Funds, trusts, and other financial vehicles	25
	Insurance carriers and related activities	397
	Real estate	222
	Rental and leasing services and lessors of intangible assets	445
	Securities, commodity contracts, and investments	247
Services	Accommodation	117
	Administrative and support services	664
	Ambulatory health care services	263
	Amusements, gambling, and recreation industries	56
	Computer systems design and related services	1162
	Educational services	46
	Food services and drinking places	238
	Hospitals	76
	Legal services	4
	Miscellaneous professional, scientific, and technical services	998
	Nursing and residential care facilities	42
	Other services, except government	115
	Performing arts, spectator sports, museums, and related activities	30
	Social assistance	3
	Waste management and remediation services	159

This table reports the number firm-year observations in the sample by primary sector and industry as determined by the BEA industry classification.

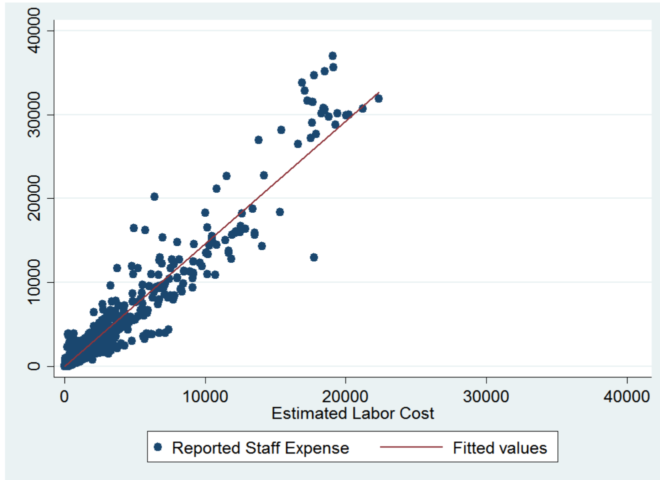

Figure 1: Estimated and Reported Labor Expenses

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

6 Additional Results and Robustness Checks

6.1 Robustness: Firm Size

6.2 Value-Added Estimates

Table 7: Productivity Spillovers by Varying Firm Size Cutoffs (Gross Output)

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

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 for GNR are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Large firms are defined by having at least as many employees as the cutoffs indicated above. The median cutoff is determined by industry and year. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 8: Productivity Spillovers by Relationship Dynamics (Value-Added)

		Dependent Variable: $\ln\text{TFP}_t$			
Period	Estimator	Continuing Customers' $\ln\text{TFP}_t$	New Customers' $\ln\text{TFP}_t$	Continuing Suppliers' $\ln\text{TFP}_t$	New Suppliers' $\ln\text{TFP}_t$
1977-1986	ACF	0.0009	0.0023	0.002	0.0005
		(0.0014)	(0.0014)	(0.0011)	(0.001)
	ACF-N	0.0008	0.0022	0.0021	0.0007
		(0.0014)	(0.0014)	(0.0011)	(0.001)
	ACF-ND	0.0008	0.001	0.003	0.0003
		(0.0016)	(0.0015)	(0.0013)	(0.0011)
1987-1996	ACF	0.0001	0.0003	-0.0006	0.001
		(0.0009)	(0.001)	(0.0007)	(0.0006)
	ACF-N	0.0002	0.0003	-0.0011	0.0005
		(0.0009)	(0.001)	(0.0007)	(0.0006)
	ACF-ND	0.0004	0.0004	-0.0012	0.0006
		(0.001)	(0.001)	(0.0007)	(0.0007)
1997-2006	ACF	0.0008	0.0004	0.0003	0.0014
		(0.0005)	(0.0005)	(0.0004)	(0.0004)
	ACF-N	0.0008	0.0004	0.0006	0.0017
		(0.0005)	(0.0005)	(0.0004)	(0.0004)
	ACF-ND	0.0005	0.0004	0.0006	0.002
		(0.0005)	(0.0004)	(0.0005)	(0.0004)
2007-2016	ACF	0.0012	0.0003	0.0011	0.0003
		(0.0006)	(0.0004)	(0.0004)	(0.0003)
	ACF-N	0.0011	0.0003	0.0013	0.0004
		(0.0006)	(0.0004)	(0.0004)	(0.0003)
	ACF-ND	0.0012	0.0003	0.0011	0.0002
		(0.0007)	(0.0004)	(0.0004)	(0.0003)
All	ACF	0.0007	0.0008	0.0006	0.0009
		(0.0003)	(0.0003)	(0.0003)	(0.0002)
	ACF-N	0.0007	0.0008	0.0008	0.001
		(0.0003)	(0.0003)	(0.0003)	(0.0002)
	ACF-ND	0.0007	0.0008	0.0007	0.0011
		(0.0004)	(0.0003)	(0.0003)	(0.0002)

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 for ACF are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 9: Productivity Spillovers by Sector (Value-Added)

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

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 for ACF are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Sectors are determined according to the BEA industry classification. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 10: Productivity Spillovers by Firm Size & Relationship Direction (Value-Added)

			Dependent Variable: $\ln\text{TFP}_t$				
Partner Size	Relationship	Firm Size	1977-1986	1987-1996	1997-2006	2007-2016	All
Large	Customers	Large	0.0002	0.0005	0.0002	0.0007	0.0008
			(0.0015)	(0.0007)	(0.0005)	(0.0004)	(0.0003)
		Small	0.0066	0.0007	-0.0021	-0.0001	0.002
			(0.0019)	(0.001)	(0.0007)	(0.0006)	(0.0004)
	Suppliers	Large	0.0002	-0.0009	0.0012	0.0007	0.0004
			(0.0013)	(0.0006)	(0.0004)	(0.0003)	(0.0002)
		Small	0.0339	0.0006	-0.0012	-0.0038	0.0028
			(0.018)	(0.005)	(0.0046)	(0.0031)	(0.0027)
Small	Customers	Large	-0.0033	-0.0046	-0.0088	-0.0043	-0.0054
			(0.0047)	(0.0048)	(0.004)	(0.0022)	(0.0023)
		Small	0.0125	-0.0059	-0.0021	0.0022	-0.0003
			(0.0071)	(0.0062)	(0.0029)	(0.0059)	(0.0028)
	Suppliers	Large	0.0013	-0.0001	0.0017	0.0005	0.0007
			(0.0011)	(0.0007)	(0.0005)	(0.0003)	(0.0002)
		Small	0.0218	-0.0004	0.0045	-0.0023	0.0039
			(0.0088)	(0.0046)	(0.0041)	(0.0046)	(0.0025)

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

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

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

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 for ACF are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Large firms are defined by having at least as many employees as the cutoffs indicated above. The median cutoff is determined by industry and year. Standard errors are in parentheses. All specifications include industry and year fixed effects.

6.3 Unweighted Estimates

Table 12: Endogenous Productivity Spillovers (Gross Output, Unweighted)

Dependent Variable: $\ln\text{TFP}_t$			
Period	Estimator	$\ln\text{TFP}_{t-1}$	Neighbors' $\ln\text{TFP}_t$
1977-1986	GNR	0.8403	0.0068
		(0.0205)	(0.0049)
	GNR-N	0.8391	0.0074
		(0.0207)	(0.0049)
	GNR-ND	0.8248	-0.0001
		(0.0228)	(0.0058)
1987-1996	GNR	0.8314	-0.0065
		(0.0244)	(0.0039)
	GNR-N	0.8312	-0.0063
		(0.0244)	(0.0039)
	GNR-ND	0.8232	-0.0093
		(0.0271)	(0.0039)
1997-2006	GNR	0.8583	0.0001
		(0.013)	(0.0043)
	GNR-N	0.8588	0.0003
		(0.0132)	(0.0043)
	GNR-ND	0.8584	0.0002
		(0.0145)	(0.0055)
2007-2016	GNR	0.8964	0.0115
		(0.0214)	(0.0048)
	GNR-N	0.8964	0.0113
		(0.0214)	(0.005)
	GNR-ND	0.889	0.0096
		(0.0229)	(0.0047)
All	GNR	0.9038	0.0076
		(0.0066)	(0.0023)
	GNR-N	0.9027	0.0086
		(0.0067)	(0.0024)
	GNR-ND	0.8998	0.0069
		(0.0073)	(0.0025)

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 for GNR are estimated using the generalized 2SLS procedure in Lee (2003); Bramoullé et al. (2009). Interaction matrices for network effects are unweighted. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 13: Productivity Spillovers by Relationship Direction (Gross Output, Unweighted)

Dependent Variable: $\ln\text{TFP}_t$			
Period	Estimator	Customers' $\ln\text{TFP}_t$	Suppliers' $\ln\text{TFP}_t$
1977-1986	GNR	0.0036	0.0167
		(0.004)	(0.005)
	GNR-N	0.0022	0.017
		(0.0033)	(0.0045)
	GNR-ND	-0.0025	0.0127
		(0.0036)	(0.004)
1987-1996	GNR	0.0063	-0.008
		(0.004)	(0.0036)
	GNR-N	0.0079	-0.0122
		(0.0039)	(0.0035)
	GNR-ND	-0.0056	-0.012
		(0.0044)	(0.0079)
1997-2006	GNR	0.0013	0.0023
		(0.0005)	(0.0005)
	GNR-N	0.001	0.0033
		(0.0005)	(0.0005)
	GNR-ND	0.0006	0.0041
		(0.0006)	(0.0006)
2007-2016	GNR	0.0006	0.0016
		(0.0004)	(0.0006)
	GNR-N	0.0004	0.0024
		(0.0004)	(0.0008)
	GNR-ND	0.0002	0.0025
		(0.0004)	(0.0008)
All	GNR	0.0025	0.0053
		(0.0008)	(0.0009)
	GNR-N	0.0024	0.0079
		(0.0008)	(0.0011)
	GNR-ND	0.0015	0.008
		(0.0008)	(0.001)

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 for GNR are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Interaction matrices for network effects are unweighted. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 14: Productivity Spillovers by Relationship Dynamics (Gross Output, Unweighted)

		Dependent Variable: $\ln\text{TFP}_t$			
Period	Estimator	Continuing Customers' $\ln\text{TFP}_t$	New Customers' $\ln\text{TFP}_t$	Continuing Suppliers' $\ln\text{TFP}_t$	New Suppliers' $\ln\text{TFP}_t$
1977-1986	GNR	0.0037	0.001	0.0149	0.0092
		(0.0034)	(0.0039)	(0.0038)	(0.0039)
	GNR-N	0.0013	0.0004	0.0138	0.0095
		(0.0024)	(0.0026)	(0.0029)	(0.0028)
	GNR-ND	0.0001	-0.0028	0.0157	0.0088
		(0.0031)	(0.0031)	(0.0032)	(0.0035)
1987-1996	GNR	0.0077	0.0033	-0.003	-0.0121
		(0.0035)	(0.0039)	(0.0038)	(0.0037)
	GNR-N	-0.0036	0.0165	-0.0041	-0.0295
		(0.0048)	(0.0068)	(0.0081)	(0.0077)
	GNR-ND	-0.0116	0.0101	0.0006	-0.0214
		(0.0046)	(0.0065)	(0.009)	(0.0088)
1997-2006	GNR	-0.0007	0.0005	0.0003	0.0013
		(0.0004)	(0.0004)	(0.0003)	(0.0003)
	GNR-N	-0.0009	0.0003	0.0015	0.0026
		(0.0004)	(0.0004)	(0.0004)	(0.0004)
	GNR-ND	-0.0011	0.0002	0.002	0.0031
		(0.0005)	(0.0005)	(0.0005)	(0.0004)
2007-2016	GNR	0.0005	0.0005	0.0015	0.0013
		(0.0004)	(0.0004)	(0.0005)	(0.0004)
	GNR-N	0.0004	0.0004	0.0025	0.0022
		(0.0004)	(0.0004)	(0.0007)	(0.0005)
	GNR-ND	0.0001	0.0003	0.0025	0.0022
		(0.0004)	(0.0004)	(0.0008)	(0.0006)
All	GNR	0.0008	0.0017	0.0029	0.0038
		(0.0006)	(0.0006)	(0.0006)	(0.0006)
	GNR-N	0.0009	0.0017	0.0066	0.0076
		(0.0007)	(0.0007)	(0.0009)	(0.0008)
	GNR-ND	0.0005	0.0012	0.007	0.0081
		(0.0007)	(0.0008)	(0.001)	(0.0009)

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 for GNR are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Interaction matrices for network effects are unweighted. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 15: Productivity Spillovers by Sector (Gross Output, Unweighted)

	Dependent Variable: $\ln\text{TFP}_t$											
	Firm's Sector											
Partners' Sector	Mining	Utilities	Constr	Durables	Non-Durables	Electronics	Wholesale	Retail	Trans & WH	Info	FIRE	Services
00Mining	-0.0082	-0.0093	-0.0062	0.0017	0.0071	-0.0155	0.0165	0.0088	0.0098	0.0343	0.0042	0.0027
	(0.0062)	(0.0037)	(0.0317)	(0.0048)	(0.0026)	(0.0312)	(0.01)	(0.0077)	(0.0088)	(0.0107)	(0.0082)	(0.0077)
01Utilities	-0.0031	-0.0006	0.006	0.0029	0.0097	0.0127	0.0041	-0.0027	0.0018	-0.0023	-0.0035	0.01
	(0.0057)	(0.0027)	(0.0102)	(0.0026)	(0.002)	(0.0079)	(0.008)	(0.0081)	(0.0063)	(0.0121)	(0.0173)	(0.0058)
02Construction	-0.0189	0.0012	0.0144	0.0052	-0.0057	-0.0076	0.0109	-0.0072	0.0164	0.0041	-0.0002	0.0276
	(0.0303)	(0.004)	(0.0112)	(0.0057)	(0.0035)	(0.0128)	(0.0073)	(0.0087)	(0.0222)	(0.0034)	(0.0059)	(0.0213)
03Durables Mfg	0.0074	0.0002	0.0255	0.0024	0.0012	-0.0036	0.0076	-0.0016	0.0028	0.0037	0.0049	0.003
	(0.0063)	(0.0029)	(0.0223)	(0.0021)	(0.0016)	(0.0023)	(0.0026)	(0.0016)	(0.0034)	(0.0033)	(0.0061)	(0.0033)
04Non-Durables Mfg	-0.0059	-0.0014	-0.0028	0.0	-0.0009	-0.0065	0.0033	-0.006	-0.0063	-0.0025	-0.0169	-0.0098
	(0.005)	(0.0027)	(0.0129)	(0.0017)	(0.0013)	(0.0046)	(0.0026)	(0.0015)	(0.0037)	(0.0043)	(0.0082)	(0.0037)
05Electronics Mfg	-0.0413	-0.015	-0.0262	-0.0001	-0.004	0.0231	0.0126	0.0021	0.0076	0.0005	0.0203	0.0002
	(0.0478)	(0.0053)	(0.034)	(0.0044)	(0.0036)	(0.0028)	(0.0029)	(0.0031)	(0.0055)	(0.0029)	(0.0089)	(0.0041)
06Wholesale	-0.005	0.0122	0.0115	0.0021	0.0031	0.016	0.0096	0.0016	0.0032	0.0021	-0.0125	0.0026
	(0.0114)	(0.0073)	(0.0157)	(0.002)	(0.0012)	(0.002)	(0.0037)	(0.0013)	(0.0101)	(0.0023)	(0.0114)	(0.0038)
07Retail	0.0109	-0.0035	-0.0072	0.0071	0.0029	0.0187	0.0101	0.0012	-0.0012	0.0086	0.0175	0.0016
	(0.0114)	(0.01)	(0.0154)	(0.0027)	(0.0013)	(0.0035)	(0.0018)	(0.0025)	(0.0038)	(0.004)	(0.0045)	(0.0043)
08Transport and Warehousing	0.0075	0.0111	0.0536	0.0035	0.0031	0.0063	0.0074	0.0021	-0.003	0.0007	-0.0084	0.0
	(0.009)	(0.0032)	(0.0183)	(0.004)	(0.0019)	(0.0057)	(0.0105)	(0.0038)	(0.004)	(0.0079)	(0.007)	(0.01)
09Information	0.0235	-0.0023	0.0083	-0.0034	-0.001	0.0145	0.0062	-0.001	-0.0174	0.004	0.0064	-0.0071
	(0.0114)	(0.0066)	(0.0165)	(0.0038)	(0.0027)	(0.0024)	(0.0047)	(0.0025)	(0.0046)	(0.0027)	(0.0046)	(0.0038)
10Finance, Insur & Real Estate	0.0003	0.0048	-0.013	-0.0	-0.0044	0.0129	-0.0024	0.0017	0.0043	0.004	-0.0019	0.0047
	(0.0108)	(0.0137)	(0.0124)	(0.0043)	(0.0024)	(0.0033)	(0.0144)	(0.0019)	(0.0042)	(0.003)	(0.0042)	(0.0034)
11Services	0.0032	-0.0086	0.0082	0.0084	0.004	0.0021	0.008	0.004	0.0023	0.005	0.0088	0.0024
	(0.0113)	(0.0035)	(0.0245)	(0.0022)	(0.0022)	(0.0027)	(0.0069)	(0.0025)	(0.0048)	(0.003)	(0.0047)	(0.0036)

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 for GNR are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Sectors are determined according to the BEA industry classification. Interaction matrices for network effects are unweighted. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 16: Productivity Spillovers by Firm Size & Relationship Direction (Gross Output, Unweighted)

			Dependent Variable: $\ln\text{TFP}_t$				
Partner Size	Relationship	Firm Size	1977-1986	1987-1996	1997-2006	2007-2016	All
Large	Customers	Large	0.001	0.0028	-0.0001	0.0003	0.0019
			(0.0034)	(0.0073)	(0.0005)	(0.0004)	(0.0008)
		Small	-0.0039	-0.0019	0.0006	-0.0004	0.0019
			(0.0042)	(0.0056)	(0.0006)	(0.0006)	(0.001)
	Suppliers	Large	0.0151	-0.0252	0.0025	0.0026	0.0081
			(0.0036)	(0.0108)	(0.0004)	(0.0007)	(0.001)
		Small	0.0569	-0.0686	0.0047	-0.005	0.0139
			(0.0304)	(0.0216)	(0.0025)	(0.0025)	(0.0076)
Small	Customers	Large	-0.0134	-0.0778	-0.0063	-0.0059	-0.009
			(0.0117)	(0.0468)	(0.0028)	(0.0064)	(0.0075)
		Small	0.0002	-0.051	0.0	-0.0008	-0.0044
			(0.0143)	(0.0392)	(0.0022)	(0.0027)	(0.0052)
	Suppliers	Large	0.0138	-0.0125	0.0023	0.0022	0.008
			(0.0039)	(0.0058)	(0.0005)	(0.0006)	(0.0009)
		Small	0.0581	-0.0154	0.0033	0.0	0.0074
			(0.0272)	(0.0137)	(0.002)	(0.003)	(0.0052)

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 for ACF 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.Interaction matrices for network effects are unweighted. Standard errors are in parentheses. All specifications include industry and year fixed effects.

Table 17: Productivity Spillovers by Varying Firm Size Cutoffs (Gross Output, Unweighted)

			Dependent Variable: $\ln\text{TFP}_t$			
			Firm's Sector			
Partner Size	Relationship	Firm Size	500	1000	5000	Median
Large	Customers	Large	0.0019	0.0016	0.0014	0.001
			(0.0008)	(0.0008)	(0.0008)	(0.0008)
		Small	0.0019	0.0033	0.0017	0.0014
			(0.001)	(0.0011)	(0.001)	(0.001)
	Suppliers	Large	0.0081	0.0078	0.006	0.0087
			(0.001)	(0.001)	(0.0008)	(0.001)
		Small	0.0139	0.0093	0.0084	0.0091
			(0.0076)	(0.0035)	(0.0027)	(0.0016)
Small	Customers	Large	-0.009	-0.001	-0.0039	0.0018
			(0.0075)	(0.006)	(0.0038)	(0.0015)
		Small	-0.0044	-0.0051	0.0007	0.0026
			(0.0052)	(0.0036)	(0.0012)	(0.0012)
	Suppliers	Large	0.008	0.0086	0.0079	0.0073
			(0.0009)	(0.001)	(0.0012)	(0.0011)
		Small	0.0074	0.0071	0.009	0.0087
			(0.0052)	(0.0032)	(0.0013)	(0.0013)

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 for GNR are estimated using the generalized 2SLS procedure for heterogenous peer effects in Dieye and Fortin (2017); Patacchini et al. (2017). Large firms are defined by having at least as many employees as the cutoffs indicated above. The median cutoff is determined by industry and year. Interaction matrices for network effects are unweighted. Standard errors are in parentheses. All specifications include industry and year fixed effects.

7 Bootstrap for Network Data

7.1 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.

7.2 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 bootstrap 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}.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451. (Cited on pages 7, 13, 15, 18, 19, 20, 29, 30, 31, and 32.)

Arduini, T., E. Patacchini, E. Rainone, et al. (2015). Parametric and semiparametric IV estimation of network models with selectivity. Technical report, Einaudi Institute for Economics and Finance (EIEF). (Cited on pages 4, 5, and 6.)

Bernard, A. B. and A. Moxnes (2018). Networks and trade. *Annual Review of Economics* (0). (Cited on page 16.)

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

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*, 41–55. (Cited on page 34.)

Collard-Wexler, A. and J. De Loecker (2016). Production function estimation with measurement error in inputs. Technical report, National Bureau of Economic Research. (Cited on pages 7 and 9.)

Dieye, R. and B. Fortin (2017). Gender peer effects heterogeneity in obesity. (Cited on pages 24, 28, 29, 30, 31, 32, 34, 35, 36, 37, and 38.)

Erdős, P. and A. Rényi (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5*(1), 17–60. (Cited on pages 8 and 16.)

Gandhi, A., S. Navarro, and D. A. Rivers (2020). On the identification of gross output production functions. *Journal of Political Economy 128*(8), 2973–3016. (Cited on pages 21, 28, 34, 35, 36, 37, and 38.)

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264. (Cited on pages 5 and 6.)

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica 85*(4), 1033–1063. (Cited on page 5.)

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. (Cited on page 5.)

İmrohoroğlu, A. and Ş. Tüzel (2014). Firm-level productivity, risk, and return. *Management Science 60*(8), 2073–2090. (Cited on page 25.)

Jin, F. and L. F. Lee (2012). On the bootstrap for spatial econometric models. *Working Paper*. (Cited on page 39.)

Lee, L. F. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews 22*(4), 307–335. (Cited on page 34.)

Liu, X. and L. F. Lee (2010). GMM estimation of social interaction models with centrality. *Journal of Econometrics 159*(1), 99–115. (Cited on page 24.)

Liu, X., E. Patacchini, and Y. Zenou (2014). Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior & Organization 103*, 39–59. (Cited on page 24.)

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment. *Econometrica 64*(6), 1263–1297. (Cited on pages 4 and 7.)

Patacchini, E., E. Rainone, and Y. Zenou (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior & Organization 134*, 190–227. (Cited on pages 24, 28, 29, 30, 31, 32, 34, 35, 36, 37, and 38.)

Qu, X., L. F. Lee, and J. Yu (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics 197*(2), 173–201. (Cited on pages 4 and 5.)

Serpa, J. C. and H. Krishnan (2018). The impact of supply chains on firm-level productivity. *Management Science 64*(2), 511–532. (Cited on page 5.)

Snijders, T. A. and S. P. Borgatti (1999). Non-parametric standard errors and tests for network statistics. *Connections 22*(2), 161–170. (Cited on page 39.)

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *The Journal of Industrial Economics 55*(3), 529–569. (Cited on pages 7 and 9.)