Firm-to-firm Trade in Sticky Production Networks

Kevin Lim*

January 2018

Abstract

This paper develops a structural model of trade between heterogeneous firms in which the network of firm-level input-output linkages is determined both dynamically and endogenously. Firms vary in the size of their customer and supplier bases, occupy heterogeneous positions in different supply chains, and adjust their sets of trade partners over time. Despite the rich heterogeneity and dynamics, the model remains computationally tractable. Using both cross-sectional and panel data on trading relationships between US firms, I estimate the model's key parameters via a simulated method of moments technique and assess its fit to the data. Simulations of the model are then used to study how the structure and dynamics of the production network matter for the propagation of firm-level supply and demand shocks and their translation into aggregate effects.

*kvn.lim@utoronto.ca. I am extremely grateful to Gene Grossman, Steve Redding, and Esteban Rossi-Hansberg for their support and guidance in writing this paper. I also thank Oleg Itskhoki, Ezra Oberfield, Davin Chor, as well as numerous seminar participants and visiting faculty at Princeton University for helpful discussions and comments. Later drafts of this paper benefited enormously from feedback and comments during the job market at numerous seminars, as well as from discussions with the trade communities at Dartmouth College and the University of Toronto. This research benefited from financial support from the International Economics Section at Princeton University.

1 Introduction

Many of the goods and services that are traded between firms lack centralized markets or intermediaries facilitating their exchange, and instead are traded through direct connections between buyers and sellers.[1] Firm-level supply and demand shocks propagate via these connections, through the network of firm-to-firm relationships, and translate into aggregate effects. The nature of this propagation depends in principle on several empirically stark features of the production network that are often abstracted from in existing theories of production. First, firms vary in the extent to which they are connected to other firms (relationship heterogeneity). Second, firms occupy different positions in different supply chains (supply chain heterogeneity). Third, the set of active trading relationships changes over time (relationship dynamics). In this paper, I study the extent to which accounting for these characteristics of the production network matters for our understanding of firm-level supply and demand shock propagation.

To do so, I first develop a structural model of trade between heterogeneous firms in which the network of firm-level input-output linkages is endogenously determined. In the model, active relationships face a time-varying cost, and frictions impede the ability of forward-looking firms to change their sets of trade partners. These assumptions deliver a model of a production network where firms vary in the size of their customer and supplier bases, occupy heterogeneous positions in different supply chains, and adjust their sets of active relationships dynamically. I develop tractable computational algorithms to solve for the model's steady-state as well as its transition dynamics, and use both cross-sectional and panel data on firm-level trading relationships in the US to estimate the model's key parameters via a simulated method of moments technique. Finally, simulations of the model are used to study how relationship heterogeneity, supply chain heterogeneity, and relationship dynamics matter for the aggregate welfare effects of shocks to firm-level productivity and demand.

The key findings of the analysis are as follows. First, accounting for the heterogeneous distribution of relationships leads to lower predicted welfare effects of shocks to small firms, and larger predicted effects of shocks to large firms. This results intuitively from the fact that large firms are central to the production network not only because they are large in size, but also because they are more connected to other firms in the economy than smaller firms. Second, if one takes the production network as fixed, the higher-order propagation of firm-level shocks multiple stages upstream or downstream of supply chains appears to be quantitatively unimportant. In the simulations studied, over 90% of the short-run welfare

[1] Using the Rauch (1999) classification of products, for example, only about 15% of trade by US firms in 2014 was in goods that have organized exchanges, while another 15% was in goods that have reference prices (implying the existence of specialized traders engaging in price arbitrage).

effects of firm-level shocks are accounted for by propagation one stage upstream or downstream of where the shock hits. Third, it is the dynamic propagation of firm-level shocks that is quantitatively important instead, as the predicted welfare effects can differ dramatically once the endogenous adjustment of the network is taken into account.

In modeling the dynamics of firm-level trading relationships, this paper is most closely related to the models of Oberfield (2015) and Chaney (2014, 2015). In both of these models, as in this paper, the network of firm-level input-output linkages is an endogenous and dynamic outcome of a stochastic process by which potential buyer-supplier pairs receive trading opportunities over time. In Oberfield (2015), however, the number of suppliers per firm is exogenously fixed, while in Chaney (2014, 2015), every firm has the same number of suppliers even though the number of suppliers per firm grows over time. Relationship heterogeneity is therefore shut down in these models.

In modeling the matching between buying and selling firms, this paper is also closely related to the models of Bernard, Moxnes, and Saito (2015) and Bernard, Moxnes, and Ulltveit-Moe (2015). In both of these models, variation in the extensive margin of firm-to-firm relationships is similarly generated by assuming that relationships are costly. However, these papers address the static formation of relationships between one group of buyers and one group of sellers, in essence capturing one tier of relationships between firms instead of the entire network. Supply chain heterogeneity and relationship dynamics are therefore absent in these models.

In addition, this paper builds on the existing literature studying how microeconomic shocks translate into aggregate fluctuations. Acemoglu et al (2012) argue that the network structure of linkages between sectors matters for how idiosyncratic sector-level shocks translate into aggregate movements, while Magerman et al (2016) make an analogous argument by studying the production network between firms. However, neither of these papers seeks to explain what determines the network structure of the economy in the first place, nor how the network structure evolves in response to changes in the economic environment. The theory developed in this paper endogeneizes the formation of the production network, and therefore allows us to address these questions.

In this last regard, the theory developed here is related to the broader theoretical literature on social and economic network formation, within which there are two qualitatively different approaches to modeling the formation of ties between atomistic agents.[2] The first approach posits an exogenous stochastic algorithm for the formation of links, and then proceeds to study the resulting network properties.[3] As these models of network formation

[2]See Jackson (2005, 2011) for more in-depth surveys of the network formation literature.

[3]Well-known examples from the graph theory literature are the Erdös-Rényi (1959) random network, the

are non-structural, however, they cannot be used to study how networks of trade between firms respond to changes in economic incentives. The second approach to modeling network formation assumes that the creation and destruction of links are the result of strategic interactions between agents.[4] These game-theoretic approaches therefore explicitly take into account optimizing behavior by the agents constituting the network, but the complexity of solving these models beyond simple illustrative examples precludes quantitative analysis.

The modeling of network formation in this paper can thus be viewed as a combination of the two approaches discussed above, or in the terminology of Currarini, Jackson, and Pin (2010), a combination of "chance and choice": firms receive the opportunity to adjust relationships according to an exogenous stochastic process, but the activation or termination of a trading relationship conditional on having the opportunity to do so is an endogenous outcome. This hybrid approach is similar in spirit to the dynamic network formation models in Bala and Goyal (2000), Watts (2001), and Jackson and Watts (2002), but within the context of a structural model of trade between heterogeneous producers that can be used for quantitative analysis.[5]

The outline of this paper is as follows. I begin in section 2 by developing a static version of the theoretical model, in which the set of buyer-supplier relationships is taken as given. I characterize how firm size, firm-to-firm trade volumes, and aggregate outcomes such as household welfare depend on the existing production network, and show how to solve for the market equilibrium of the model given any network of relationships. In section 3, I then endogeneize the formation of linkages between firms in the economy by introducing a dynamic matching process between potential buyers and sellers, and discuss how to solve for both the model's steady-state as well as its transition dynamics. In section 4, I discuss the data used for structural estimation of the model's parameters, the simulated method of moments estimation approach, and the fit of the model to data. Section 5 then discusses the simulation exercises, and section 6 concludes.

Watts-Strogatz (1998) small world model, and the Barabási-Albert (1999) preferential attachment model. In the economics literature, Atalay et al (2011) combine the random and preferential attachment algorithms to model the buyer-supplier network in the US economy.

[4]Aumann and Myerson (1988) and Myerson (1991) model network formation as extensive-form and simultaneous move games respectively. Jackson and Wolinsky (1996) adopt a cooperative game theoretic approach, while Kranton and Minehart (2001) study buyer-seller networks in which ascending-bid auctions are used to determine the formation of links.

[5]Bala and Goyal (2000), Watts (2001), and Jackson and Watts (2002) also assume for tractability that agents are myopic in their decisions about which links to form, whereas firms in this paper are forward-looking and optimally select relationships taking into account their future costs and benefits.

2 Static Model

To study the dynamic formation of firm-to-firm linkages, it is useful to first understand how firms behave conditional on these relationships. I therefore begin by describing a static version of the model in which the network of trading relationships between firms is fixed.

2.1 Model environment

The economy consists of a representative household and an exogenously-given unit continuum of firms that each produce a unique good. Firms are heterogeneous over states $\chi = (\phi, \delta)$, where ϕ and δ are what are referred to as the *fundamental productivity* of a firm's production process and the *fundamental demand* for a firm's product respectively. The exogenous cumulative distribution function over firm states is denoted by G_χ, with density g_χ and support S_χ a bounded subset of $\mathbb{R}^2_+$.[6] For brevity, I also refer to firms with state χ as χ-firms.

2.1.1 Households

The representative household supplies L units of labor inelastically and has constant-elasticity-of-substitution (CES) preferences over all goods in the economy, given by:

$$U = \left[\int_{S_\chi} \left[\delta x_H (\chi) \right]^{\frac{\sigma-1}{\sigma}} dG_\chi (\chi) \right]^{\frac{\sigma}{\sigma-1}} \tag{2.1}$$

Here, σ denotes the elasticity of substitution across varieties, and $x_H (\chi)$ is the household's consumption of χ-firm varieties. Given the price $p_H (\chi)$ charged by χ-firms to the household, household demand is given by:

$$x_H (\chi) = \Delta_H \delta^{\sigma-1} \left[p_H (\chi) \right]^{-\sigma} \tag{2.2}$$

Note that conditional on prices, households demand a greater amount of goods for which fundamental demand δ is higher. The household's demand shifter can then be written as:

$$\Delta_H \equiv U P_H^\sigma \tag{2.3}$$

[6] Note that given the unit mass of firms, integrals of all firm-level variables over the distribution G_χ are equal to both the average as well as the total value of that variable across firms.

and the consumer price index is equal to:

$$P_H = \left[\int_{S_\chi} \left[\frac{p_H(\chi)}{\delta}\right]^{1-\sigma} dG_\chi(\chi)\right]^{\frac{1}{1-\sigma}} \tag{2.4}$$

2.1.2 Firm production technology

Each firm produces its output using labor and the output of other firms. However, firm-to-firm trade is characterized by relationship frictions, such that every $\chi-$firm is only able to purchase inputs from a given χ'-firm with probability $m(\chi, \chi')$. Given that there exists a continuum of firms of every state, $m(\chi, \chi')$ is also equal to the fraction of χ'-firms that supply a given χ-firm, as well as the fraction of χ-firms that purchase from a given χ'-firm. I refer to m as the *matching function* of the economy, which completely specifies the extensive margin of firm-to-firm trading relationships in the economy.

Given the matching function, the output of a χ-firm is then given by the following constant returns to scale CES production function:

$$X(\chi) = \left[[\phi l(\chi)]^{\frac{\sigma-1}{\sigma}} + \int_{S_\chi} m\left(\chi, \chi'\right)\left[\alpha x\left(\chi, \chi'\right)\right]^{\frac{\sigma-1}{\sigma}} dG_\chi\left(\chi'\right)\right]^{\frac{\sigma}{\sigma-1}} \tag{2.5}$$

where $l(\chi)$ is the quantity of labor demanded and $x(\chi, \chi')$ is the quantity of each χ'-good used as inputs. Note that the fundamental productivity ϕ of the firm can be interpreted as a measure of its labor productivity, while the parameter α captures how efficiently the output of one firm can be transformed into the output of another firm. To rule out explosive production, it is assumed that $\alpha < 1$.[7] As is standard in the literature, I also assume that the elasticity of substitution across inputs for intermediate demand is the same as that for final demand.

Taking the wage as numeraire and given prices $\{p(\chi, \chi')\}_{\chi' \in S_\chi}$ charged by other firms, the marginal cost of each χ-firm is therefore given by:

$$\eta(\chi) = \left[\phi^{\sigma-1} + \alpha^{\sigma-1}\int_{S_\chi} m\left(\chi, \chi'\right)\left[p\left(\chi, \chi'\right)\right]^{1-\sigma} dG_\chi(\chi)\right]^{\frac{1}{1-\sigma}} \tag{2.6}$$

[7]When $\alpha \geq 1$, it becomes feasible for a pair of firms that are connected to each other both as buyer and seller to use only each other's output as inputs for production, thereby generating infinite output and profits.

while the quantities of labor and intermediate inputs demanded are given respectively by:

$$l(\chi) = X(\chi)\,\eta(\chi)^{\sigma}\,\phi^{\sigma-1} \tag{2.7}$$

$$x\left(\chi,\chi^{'}\right) = X(\chi)\,\eta(\chi)^{\sigma}\,\alpha^{\sigma-1}p\left(\chi,\chi^{'}\right)^{-\sigma} \tag{2.8}$$

Note that conditional on prices, firms with greater fundamental productivity ϕ have lower marginal costs.

2.1.3 Market structure and firm pricing

The market structure for all firm sales is assumed to be monopolistic competition. This assumption affords the model a great degree of tractability, as it implies that regardless of the complexity of the matching function, the markups that firms charge over their marginal costs are identical in equilibrium. This follows from the fact that every buyer (including the household) faces a continuum of sellers, and that the demand functions (2.2) and (2.8) exhibit a constant price elasticity. Consequently, the profit-maximizing price charged by each firm is equal to the standard CES markup over marginal cost:

$$p_H(\chi) = \mu\eta(\chi) \tag{2.9}$$

$$p\left(\chi,\chi^{'}\right) = \mu\eta\left(\chi^{'}\right) \tag{2.10}$$

where $\mu \equiv \frac{\sigma}{\sigma-1}$.

2.1.4 Market clearing

Market clearing for labor requires:

$$\int_{S_\chi} l(\chi)\,dG_\chi(\chi) = L - L_f \tag{2.11}$$

where $L_f < L$ is the aggregate quantity of labor hired to maintain firm-to-firm relationships in the economy. In this section, we take L_f as given, whereas in section 3 when the dynamic formation of the production network is considered, L_f becomes an endogenous variable. Finally, market clearing for the output of a χ-firm requires:

$$X(\chi) = x_H(\chi) + \int_{S_\chi} m\left(\chi^{'},\chi\right) x\left(\chi^{'},\chi\right) dG_\chi(\chi') \tag{2.12}$$

2.2 Static market equilibrium

2.2.1 Firm network characteristics

As described above, the parameters ϕ and δ capture exogenous productivity and demand characteristics that are fundamental to the firm, in the sense that they are independent of the firm's connection to other firms. Firm-level outcomes in equilibrium, however, such as the overall size and profit of a firm, depend not only on a firm's fundamental characteristics but also on the characteristics of other firms that it is connected to in the production network. For an arbitrary matching function, a given firm-level outcome may therefore in principle be a function of very complicated moments of the production network, which would render the model intractable.

To circumvent this problem, I rely on the structure of the CES production function specified in (2.5) to derive sufficient statistics at the firm level, from which all variables of interest can be easily computed. In contrast with firm fundamental characteristics ϕ and δ, it is therefore useful to characterize the static market equilibrium of the model in terms of what I call a χ-firm's *network productivity and demand*, defined respectively by:

$$\Phi(\chi) \equiv \eta(\chi)^{1-\sigma} \tag{2.13}$$

$$\Delta(\chi) \equiv \frac{1}{\Delta_H} X(\chi)\, \eta(\chi)^{\sigma} \tag{2.14}$$

Note that $\Phi(\chi)$ is an inverse measure of a χ-firm's marginal cost, while $\Delta(\chi)$ is the demand shifter in a χ-firm's intermediate demand function (2.8) relative to the household's demand shifter Δ_H.

Combining the demand equations (2.2) and (2.8), the firm marginal cost equation (2.6), the goods market clearing condition (2.12), and the pricing conditions (2.9) and (2.10), we obtain the following system of equations that determines firms' network characteristics:

$$\Phi(\chi) = \phi^{\sigma-1} + \mu^{1-\sigma}\alpha^{\sigma-1} \int_{S_\chi} m\left(\chi, \chi'\right) \Phi\left(\chi'\right) dG_\chi\left(\chi'\right) \tag{2.15}$$

$$\Delta(\chi) = \mu^{-\sigma}\delta^{\sigma-1} + \mu^{-\sigma}\alpha^{\sigma-1} \int_{S_\chi} m\left(\chi', \chi\right) \Delta\left(\chi'\right) dG_\chi\left(\chi'\right) \tag{2.16}$$

Note that (2.15) and (2.16) constitute a pair of decoupled linear functional equations in Φ and Δ respectively, and show how a firm's network characteristics depend on both its fundamental characteristics as well as on the network characteristics of its suppliers and customers. Conditional on ϕ and δ, firms that are connected to firms with larger network productivities and demands also have higher network productivities and demands themselves.

Furthermore, since $\alpha < 1$, $\mu > 1$, and $m\left(\chi, \chi^{'}\right) \leq 1$ for all $\left(\chi, \chi^{'}\right) \in S_\chi^2$, it is easily verified via Blackwell's sufficient conditions that (2.15) and (2.16) constitute decoupled contraction mappings in Φ and Δ. The contraction mapping theorem therefore immediately implies the existence and uniqueness of a solution to the firm network characteristic functions, and also guarantees that iteration on Φ and Δ converges to this solution. This offers a tractable method of solving for the model's static equilibrium regardless of the complexity of the matching function.

Proposition 1. *There exist unique network productivity and demand functions $\Phi : S_\chi \to \mathbb{R}_+$ and $\Delta : S_\chi \to \mathbb{R}_+$ for any matching function $m : S_\chi \times S_\chi \to [0, 1]$.*

Note that we can also rewrite equations (2.15) and (2.16) to express the network productivity and demand of a χ-firm respectively as:

$$\Phi(\chi) = \int_{S_\chi} \left[\sum_{d=0}^{\infty} \left(\frac{\alpha}{\mu}\right)^{d(\sigma-1)} m^{(d)}\left(\chi, \chi^{'}\right) \right] \left(\phi^{'}\right)^{\sigma-1} dG_\chi\left(\chi^{'}\right) \tag{2.17}$$

$$\Delta(\chi) = \mu^{-\sigma} \int_{S_\chi} \left[\sum_{d=0}^{\infty} \frac{1}{\mu^d} \left(\frac{\alpha}{\mu}\right)^{d(\sigma-1)} m^{(d)}\left(\chi^{'}, \chi\right) \right] \left(\delta^{'}\right)^{\sigma-1} dG_\chi\left(\chi^{'}\right) \tag{2.18}$$

where $m^{(d)}$ is the d^{th}-degree matching function, defined recursively by:

$$m^{(0)}\left(\chi, \chi^{'}\right) = \begin{cases} \frac{1}{g_\chi(\chi)}, & \text{if } \chi = \chi^{'} \\ 0, & \text{if } \chi \neq \chi^{'} \end{cases} \tag{2.19}$$

$$m^{(1)}\left(\chi, \chi^{'}\right) = m\left(\chi, \chi^{'}\right) \tag{2.20}$$

$$m^{(d)}\left(\chi, \chi^{'}\right) = \int_{S_\chi} m^{(d-1)}\left(\chi, \chi^{''}\right) m\left(\chi^{''}, \chi^{'}\right) dG_\chi\left(\chi^{''}\right) \tag{2.21}$$

Intuitively, one can think of $m^{(d)}\left(\chi, \chi^{'}\right)$ for $d \geq 1$ as the probability that a χ-firm buys indirectly from a $\chi^{'}$-firm through a supply chain that is of length d. With this interpretation, equations (2.17) and (2.18) show how the network characteristics of a firm depend on its connections to all other firms via supply chains of all lengths. Note that the rate at which the value of an indirect relationship decays with the length of the supply chain is decreasing in input suitability α and increasing in the markup μ.

2.2.2 Firm size and inter-firm trade

Once firm network characteristics are known, the total revenue, variable profit, and variable employment of a χ-firm are completely determined up to the scale factor Δ_H. These are given respectively by:

$$R(\chi) = \mu \Delta_H \Delta(\chi) \Phi(\chi) \tag{2.22}$$

$$\pi(\chi) = (\mu - 1) \Delta_H \Delta(\chi) \Phi(\chi) \tag{2.23}$$

$$l(\chi) = \Delta_H \Delta(\chi) \phi^{\sigma-1} \tag{2.24}$$

Intuitively, if a firm is twice as productive and produces a product for which there is twice as much demand from the perspective of the entire networked economy, its revenue and profit (gross of fixed operating costs) is quadrupled. Total output of a χ-firm is also completely determined by firm fundamental and network characteristics up to a scale factor:

$$X(\chi) = \Delta_H \Delta(\chi) \Phi(\chi)^{\frac{\sigma}{\sigma-1}} \tag{2.25}$$

as are the value and quantity of output traded from χ'- to χ-firms:

$$r\left(\chi, \chi'\right) = \left(\frac{\alpha}{\mu}\right)^{\sigma-1} \Delta_H \Delta(\chi) \Phi\left(\chi'\right) \tag{2.26}$$

$$x\left(\chi, \chi'\right) = \frac{\alpha^{\sigma-1}}{\mu^{\sigma}} \Delta_H \Delta(\chi) \Phi\left(\chi'\right)^{\frac{\sigma}{\sigma-1}} \tag{2.27}$$

2.2.3 Household welfare and demand

To complete characterization of the static market equilibrium, it remains to determine the scale factor Δ_H. From the labor market clearing condition (2.11) and the firm variable employment equation (2.24), this is given by:

$$\Delta_H = \frac{L - L_f}{\int_{S_\chi} \Delta(\chi) \phi^{\sigma-1} dG_\chi(\chi)} \tag{2.28}$$

Equations (2.3) and (2.4) then give the CPI and household welfare respectively as:

$$P_H = \mu \left[\int_{S_\chi} \Phi(\chi) \delta^{\sigma-1} dG_\chi(\chi)\right]^{\frac{1}{1-\sigma}} \tag{2.29}$$

$$U = \mu^{-\sigma} (L - L_f) \frac{\left[\int_{S_\chi} \Phi(\chi) \delta^{\sigma-1} dG_\chi(\chi)\right]^{\frac{\sigma}{\sigma-1}}}{\int_{S_\chi} \Delta(\chi) \phi^{\sigma-1} dG_\chi(\chi)} \tag{2.30}$$

while household demand is given by:

$$x_H(\chi) = \mu^{-\sigma} \Delta_H \delta^{\sigma-1} \Phi(\chi)^{\frac{\sigma}{\sigma-1}} \tag{2.31}$$

Using equations (2.17) and (2.18) to substitute for $\Phi(\chi)$ and $\Delta(\chi)$ respectively in equation (2.30), we can also express household welfare as:

$$U = (L - L_f)\frac{\left[\int_{S_\chi}\int_{S_\chi}\left[\sum_{d=0}^{\infty}\left(\frac{\alpha}{\mu}\right)^{d(\sigma-1)} m^{(d)}\left(\chi,\chi^{'}\right)\right]\left(\delta\phi^{'}\right)^{\sigma-1} dG_\chi(\chi)\, dG_\chi\left(\chi^{'}\right)\right]^{\frac{\sigma}{\sigma-1}}}{\int_{S_\chi}\int_{S_\chi}\left[\sum_{d=0}^{\infty}\frac{1}{\mu^d}\left(\frac{\alpha}{\mu}\right)^{d(\sigma-1)} m^{(d)}(\chi,\chi')\right](\delta\phi')^{\sigma-1} dG_\chi(\chi)\, dG_\chi(\chi')} \tag{2.32}$$

Note that the integrands in the numerator and denominator of (2.32) are identical except for the term μ^{-d}. An intuitive approximation to the value of household welfare is therefore:

$$U \approx (L - L_f)\,\mathcal{C} \tag{2.33}$$

where $\mathcal{C}$ is a measure of the total connectivity between firms in the economy:

$$\mathcal{C} \equiv \left[\int_{S_\chi}\int_{S_\chi}\left[\sum_{d=0}^{\infty}\left(\frac{\alpha}{\mu}\right)^{d(\sigma-1)} m^{(d)}\left(\chi,\chi^{'}\right)\right]\left(\delta\phi^{'}\right)^{\sigma-1} dG_\chi(\chi)\, dG_\chi\left(\chi^{'}\right)\right]^{\frac{1}{\sigma-1}} \tag{2.34}$$

Equation (2.33) shows how household welfare is greater when buyers of greater fundamental quality δ are better connected with sellers of greater fundamental productivity $\phi^{'}$, with the welfare cost of additional relationships captured by the term $L - L_f$. The approximation (2.33) is exact only in the limit as $\mu \to 1$ (perfect competition), but when $\mu > 1$, the same general intuition applies.

2.2.4 Static market equilibrium definition

Given the matching function m and the associated quantity of labor L_f used for relationship costs, we can now define a static market equilibrium of the economy as follows.

Definition 1. A static market equilibrium of the economy is a pair of firm network characteristic functions $\Phi : S_\chi \to \mathbb{R}_+$ and $\Delta : S_\chi \to \mathbb{R}_+$ satisfying equations (2.15) and (2.16), a scalar household demand shifter Δ_H satisfying (2.28), and allocation functions $\{l(\cdot), X(\cdot), x(\cdot,\cdot), x_H(\cdot)\}$ given respectively as side equations by (2.24), (2.25), (2.27), and (2.31).

The computational algorithm used to solve for the static market equilibrium is described in detail in section A.1 of the online appendix. Since Proposition 1 guarantees that the network characteristic functions Φ and Δ are uniquely determined, uniqueness of the static market equilibrium follows immediately.

Proposition 2. *The static market equilibrium exists and is unique.*

2.2.5 Static market equilibrium efficiency

To characterize the efficiency of the static market equilibrium, one can compare the resulting allocation with the allocation that would be chosen by a social planner seeking to maximize household welfare subject to the same exogenous matching function, production technology, and resource constraints. The following proposition (proved in section B.1 of the online appendix) summarizes the solution to the planner's problem.

Proposition 3. *Given a matching function* $m : S_\chi \times S_\chi \to [0,1]$*, the network characteristic functions under the social planner's allocation satisfy:*

$$\Phi^{SP}(\chi) = \phi^{\sigma-1} + \alpha^{\sigma-1} \int_{S_\chi} m\left(\chi, \chi^{'}\right) \Phi^{SP}\left(\chi^{'}\right) dG_\chi\left(\chi^{'}\right) \tag{2.35}$$

$$\Delta^{SP}(\chi) = \delta^{\sigma-1} + \alpha^{\sigma-1} \int_{S_\chi} m\left(\chi^{'}, \chi\right) \Delta^{SP}\left(\chi^{'}\right) dG_\chi\left(\chi^{'}\right) \tag{2.36}$$

and the allocations of output and labor are given by equations (2.24), (2.25), (2.27), and (2.31) with μ *set equal to* 1.

This result implies that that any static market equilibrium allocation coincides with the corresponding planner's allocation if and only if all firms in the decentralized equilibrium charge zero markups. With monopolistically-competitive firms, the static market equilibrium allocation is therefore inefficient because of the distortion arising from double marginalization. Note that the introduction of relationship frictions into the model through the exogenous matching function m imposes no additional inefficiency beyond this standard distortion. Once the matching function is endogeneized in section 3, this will no longer be true, as firm's decisions about which relationships to keep active generate an additional dynamic source of inefficiency.

3 Dynamics and Endogenous Network Formation

As discussed above, solution of the model is straightforward given an arbitrary matching function m. It is the determination of the matching function, however, that encapsulates the decisions of firms regarding which relationships to form with one another. In this section, I introduce a dynamic process of firm matching to study how the production network is determined and how it evolves over time.

3.1 Model environment

3.1.1 Households

Time is discrete and the representative household has preferences at date t defined by:

$$V_t = \sum_{s=t}^{\infty} \beta^{s-t} U_s \tag{3.1}$$

where U_t is given by the date t equivalent of (2.1). Since the household's value function is linear in per-period utility, household decisions in every period are characterized exactly as in the static model, and the discount factor β only affects how firms (which are owned by the household) discount the future.

3.1.2 Costly relationships

Observe that the CES production technology (2.5) implies two things. First, access to additional suppliers always lowers the marginal cost of a firm, which follows from the love of variety feature of the production function.[8] Second, access to additional customers always increases a firm's variable profit, which follows from production being constant returns to scale. These forces generate incentives for firms to form as many upstream and downstream trading relationships as possible. To counterbalance these incentives and thereby model the endogenous selection of firm-to-firm relationships, I therefore assume that relationships are costly.[9]

In particular, it is assumed that in order for any buyer-seller relationship to be active at date t, a fixed quantity of labor must be hired by the selling firm, given by:

$$f_t = \psi \xi_t \tag{3.2}$$

The first term ψ is time-invariant, and captures the overall level of relationship costs in the economy.[10] The second term ξ_t, which I refer to as the *cost shock*, is a random variable that is independent and identically distributed across firm pairs and time, with cumulative

[8]Note that love of variety in the production technology can be reinterpreted as a firm facing convex costs of producing intermediate inputs using goods from any one supplier, which leads to the same demand functions and marginal costs.

[9]As a practical example of the form that this cost might take, market analysts estimate that US firms spent more than $10bn in 2014 on relationship management software systems alone (Gartner, Inc. (2014a, 2014b)).

[10]Here ψ is assumed to be constant across all firm pairs, but allowing dependence of this parameter on the fundamental characteristics of the buying and selling firms can easily be accommodated without increasing the computational complexity of the model.

distribution function G_ξ and unit mean. The stochastic nature of ξ_t is what generates the creation of new firm-to-firm linkages and the destruction of existing ones, even in steady-state, and allows the model to address relationship dynamics.

Note that the assumption that the selling firm always pays the full share of the relationship cost is necessary to ensure that the constant-markup pricing described in section 2 remains optimal in the dynamic setting.[11] Furthermore, this assumption implies that firms are always willing to form upstream relationships, which simplifies analysis of the network formation process, as this can then be considered solely from the perspective of potential sellers.

In addition, the assumption that ξ_t exhibits no serial correlation is also made primarily for tractability. While one might expect relationship costs to be persistent, allowing for ξ_t to be serially correlated greatly increases the computational complexity of the model, as it then becomes necessary to keep track of a state variable that varies across firm pairs in each period. Even with *iid* relationship cost shocks, however, the model generates non-trivial predictions about the persistence of relationships via assumptions about how often firms can adjust relationships, described next.

3.1.3 Sticky relationships

It is assumed that firm-to-firm trading relationships are also temporally sticky in the following sense: at each date, every relationship receives with probability $1-\nu$ the opportunity to be altered along the extensive margin.[12] I refer to this as the *reset shock*, and assume that it is independent across all firm pairs. The assumption that firms can only sell to new customers with probability less than one is intended to model the fact that potential trading partners take time to meet and learn about the suitability of their output for each other's production processes or to negotiate new trading arrangements. Similarly, the assumption that firms face frictions in terminating existing relationships may be interpreted as either legal barriers to reneging on pre-negotiated contractual obligations, or more simply as capturing the idea that winding down trading relationships also takes time.[13]

Although the model can easily accommodate differences in the probabilities with which a firm can create and destroy relationships, it is assumed for parsimony that these probabilities

[11] If a buying firm had to pay a positive fixed cost and found a relationship undesirable given CES markup pricing by the seller, the seller might then find it optimal to reduce its markup so as to incentivize the buyer to form the relationship.

[12] That is, to be activated if previously inactive, and to be terminated if previously active.

[13] Surveys of US firms show that the average business-to-business (B2B) deal requires approval from more than five decision-makers (Schmidt et al (2015)), while data from Google reveal that employees tasked with researching B2B purchases typically perform more than twelve online searches before engaging with a potential business partner's website (Snyder and Hilal (2015)).

are the same. Furthermore, note that regardless of whether a reset shock is received, selling firms can costlessly adjust prices every period, so that firm-to-firm relationships are sticky only along the extensive margin.

3.2 Dynamic market equilibrium

3.2.1 Law of motion for the matching function

Under the assumptions described above, the matching function evolves according to the following law of motion:

$$
\begin{aligned}
m_t\left(\chi,\chi^{'}\right) &= m_{t-1}\left(\chi,\chi^{'}\right) \\
&\quad + (1-\nu)\left[1-m_{t-1}\left(\chi,\chi^{'}\right)\right] a_t\left(\chi,\chi^{'}\right) \\
&\quad - (1-\nu)\, m_{t-1}\left(\chi,\chi^{'}\right)\left[1-a_t\left(\chi,\chi^{'}\right)\right] \\
&= \nu m_{t-1}\left(\chi,\chi^{'}\right) + (1-\nu)\, a_t\left(\chi,\chi^{'}\right)
\end{aligned}
\tag{3.3}
$$

where $a_t\left(\chi,\chi^{'}\right)$ is the endogenous probability that a $\chi^{'}$-firm sells to a χ-firm in period t conditional on being given the opportunity to reset that relationship. The first term on the right-hand side of (3.3) is the mass of relationships that were active in the previous period, the second term is the mass of relationships that are newly created in period t, and the third term is the mass of relationships that are terminated in period t. In any steady-state of the model, the matching function is then simply given by:

$$
m\left(\chi,\chi'\right) = a\left(\chi,\chi^{'}\right) \tag{3.4}
$$

Note that the acceptance probability a_t completely summarizes the dynamic strategic behavior of firms regarding which relationships to form and which to terminate. I refer to a_t as the *acceptance function*, and turn now to its characterization.

3.2.2 Dynamic relationship activation decisions

As discussed above, the assumption that buying firms pay none of the fixed relationship cost implies that the desirability of a relationship depends only the profit that can be generated for the seller. For a $\chi^{'}$-firm selling to a χ-firm at date t, this profit value is the same as in the static market equilibrium, given by equations (2.16) and (2.23) as:

$$
\pi_t\left(\chi,\chi^{'}\right) = \mu^{-\sigma}(\mu-1)\,\alpha^{\sigma-1}\Delta_{H,t}\Delta_t(\chi)\,\Phi_t\left(\chi^{'}\right) \tag{3.5}
$$

where Φ_t, Δ_t, and $\Delta_{H,t}$ are defined by the date t equivalents of equations (2.15), (2.16), and (2.28).

Now, let $V_t^+\left(\chi,\chi^{'}|\xi_t\right)$ denote the value to a $\chi^{'}$-firm of selling to a χ-firm in period t conditional on the realization of the relationship cost shock ξ_t, and let $V_t^-\left(\chi,\chi^{'}\right)$ denote the value to the firm of not selling.[14] These value functions are given by the following Bellman equations:

$$
\begin{aligned}
V_t^+\left(\chi,\chi^{'}|\xi_t\right) &= \pi_t\left(\chi,\chi^{'}\right) - \psi\xi_t \\
&\quad + \beta\left(1-\nu\right)\mathbb{E}_t\left[V_{t+1}^O\left(\chi,\chi^{'}|\xi_{t+1}\right)\right] + \beta\nu\mathbb{E}_t\left[V_{t+1}^+\left(\chi,\chi^{'}|\xi_{t+1}\right)\right] && (3.6)\\
V_t^-\left(\chi,\chi^{'}\right) &= \beta\left(1-\nu\right)\mathbb{E}_t\left[V_{t+1}^O\left(\chi,\chi^{'}|\xi_{t+1}\right)\right] + \beta\nu V_{t+1}^-\left(\chi,\chi^{'}\right) && (3.7)
\end{aligned}
$$

where $V_t^O\left(\chi,\chi^{'}|\xi_t\right)$ denotes the value to a $\chi^{'}$-firm of having the option to reset its relationship with a χ-firm customer given the relationship cost shock ξ_t:

$$
V_t^O\left(\chi,\chi^{'}|\xi_t\right) = \max\left\{V_t^+\left(\chi,\chi^{'}|\xi_t\right), V_t^-\left(\chi,\chi^{'}\right)\right\} \tag{3.8}
$$

Observe that if relationships are not sticky ($\nu = 0$) or firms are completely myopic ($\beta = 0$) , then $V_t^+\left(\chi,\chi^{'}|\xi_t\right) \geq V_t^-\left(\chi,\chi^{'}\right)$ if and only if $\pi_t\left(\chi,\chi^{'}\right) \geq \psi\xi_t$. In these two special cases, relationships are activated as long as the static profits accruing to selling firms cover the relationship cost in each period. The probability that a $\chi^{'}$-firm sells to a χ-firm at date t once it has the chance to do so is then given by:

$$
\tilde{a}_t\left(\chi,\chi^{'}\right) = G_\xi\left[\frac{\pi_t\left(\chi,\chi^{'}\right)}{\psi}\right] \tag{3.9}
$$

The assumption of sticky relationships, however, makes the activation and termination decisions facing a given firm forward-looking. If a firm chooses not to sell to a potential customer despite having the chance to do so, it may be forced to wait several periods before being able to activate the relationship. Similarly, if a firm chooses not to terminate a relationship given the chance to do so, it may find itself wishing to terminate the relationship in the future but lacking the opportunity to do so.

To solve the dynamic activation decision problem of a firm, it is instructive to first consider a steady-state of the model in which the functions π_t, V_t^+, V_t^-, and V_t^O are all

[14] Note that since the relationship cost shocks are i.i.d. over time, the value of not selling at date t does not depend on ξ_t. Furthermore, since there is no aggregate uncertainty in the model, this implies that there is no uncertainty over the value of V_t^- at any date for any pair of firms.

constant. From equations (3.6) and (3.7), one can verify that:

$$\mathbb{E}\left[V^{O}\left(\chi,\chi'|\xi\right)\right]=\begin{cases}\frac{\pi\left(\chi,\chi'\right)-\psi}{1-\beta}, & \forall\left(\chi,\chi'\right)\in S_{+}^{2}\\ 0, & \forall\left(\chi,\chi'\right)\notin S_{+}^{2}\end{cases} \tag{3.10}$$

where $S_{+}^{2}\equiv\left\{\left(\chi,\chi'\right)\subset S_{\chi}^{2}|\pi\left(\chi,\chi'\right)-\psi\geq 0\right\}$. That is, the option value of a relationship is positive if and only if the profit from that relationship exceeds the relationship cost *on average*. Substituting (3.10) into (3.6) and (3.7), we then find:

$$V^{+}\left(\chi,\chi'|\xi\right)-V^{-}\left(\chi,\chi'\right)=\frac{\pi\left(\chi,\chi'\right)-\beta\nu\psi}{1-\beta\nu}-\psi\xi \tag{3.11}$$

and therefore the probability that a χ'-firm sells to a χ-firm conditional on having the chance to do so is given by:

$$a\left(\chi,\chi'\right)=G_{\xi}\left[\frac{\pi\left(\chi,\chi'\right)/\psi}{1-\beta\nu}-\frac{\beta\nu}{1-\beta\nu}\right] \tag{3.12}$$

Comparing this expression with equation (3.9), we again see that a relationship with a greater ratio of profits to the average relationship cost is more likely to form. Once the option value of the relationship is taken into account, however, this effect becomes more pronounced, with the profit-cost ratio scaled by a factor $\frac{1}{1-\beta\nu}$. Note that relationships with $\pi\left(\chi,\chi'\right)>\psi$ have positive option values, and there is a positive probability that temporarily-unprofitable relationships of this kind will still be activated because the relationship is profitable enough on average. Conversely, relationships with $\pi\left(\chi,\chi'\right)<\psi$ have zero option value, and there is a positive probability that temporarily-profitable relationships will not be activated because the relationship is not profitable enough on average. Furthermore, observe that (3.12) implies that firm pairs with $\pi\left(\chi,\chi'\right)<\beta\nu\psi$ will never form trading relationships in steady-state.

To characterize the activation and termination decisions of firms outside the steady-state, one can then iterate forward on equations (3.6), (3.7), and (3.8), which yields the following expression for the selling premium:

$$\begin{aligned}V_{t}^{+}\left(\chi,\chi'|\xi_{t}\right)-V_{t}^{-}\left(\chi,\chi'\right)=&\ \pi_{t}\left(\chi,\chi'\right)-\psi\xi_{t}\\ &+\sum_{s=1}^{\infty}\left(\beta\nu\right)^{s}\left[\pi_{t+s}\left(\chi,\chi'\right)-\psi\right]\end{aligned} \tag{3.13}$$

Note that the right-hand side of (3.13) is simply the expected future stream of profits net of fixed costs until the relationship can be reset. The acceptance function at date t is therefore given by:

$$a_t\left(\chi,\chi^{'}\right)=G_\xi\left[\frac{\pi_t\left(\chi,\chi^{'}\right)}{\psi}+\sum_{s=1}^{\infty}\left(\beta\nu\right)^s\left[\frac{\pi_{t+s}\left(\chi,\chi^{'}\right)}{\psi}-1\right]\right] \tag{3.14}$$

Evidently, solving for the acceptance function at date t outside of the steady-state requires solving for the profit functions π_{t+s} for all $s\geq 1$. In section A.2 of the appendix, I describe the computational algorithm that I employ to accomplish this, which involves iterating on the path of profit functions $\{\pi_{t+s}\}_{s=1}^{T}$ for some value of T large enough such that m_{t+T} is close to the eventual steady-state matching function. This allows solution of the model's transition dynamics between steady-states in about one hour on a standard personal computer.

3.2.3 Aggregate relationship costs

To close the model, it remains to determine the aggregate quantity of labor $L_{f,t}$ used to pay for relationship costs at date t, which enters into the labor market clearing condition (2.11). Note that even though ξ_t is assumed to have a unit mean, firms in the dynamic market equilibrium select relationships based on the realized values of the relationship cost shocks. Therefore, the total mass of labor used to pay for relationship fixed costs is given by:

$$L_{f,t}=\int_{S_\chi}\int_{S_\chi}\left[\nu m_{t-1}\left(\chi,\chi^{'}\right)\psi+\left(1-\nu\right)\psi\bar{\xi}_t\left(\chi,\chi^{'}\right)\right]dG_\chi\left(\chi\right)dG_\chi\left(\chi^{'}\right) \tag{3.15}$$

The first term in the integral reflects the cost of relationships that cannot be reset (and hence for which there is no selection on ξ_t), while the second term reflects the cost of relationships that are voluntarily selected by firms. The term $\bar{\xi}_t\left(\chi,\chi^{'}\right)$ denotes the average value of the idiosyncratic component of the cost shock amongst $\chi-\chi^{'}$ firm pairs that receive the reset shock:

$$\bar{\xi}_t\left(\chi,\chi^{'}\right)=\int_0^{\xi_{max,t}\left(\chi,\chi^{'}\right)}\xi dG_\xi\left(\xi\right) \tag{3.16}$$

and $\xi_{max,t}\left(\chi,\chi^{'}\right)$ is the maximum value of the cost shock for which $\chi-\chi^{'}$ relationships are voluntarily selected:

$$\xi_{max,t}\left(\chi,\chi^{'}\right)=\max\left\{\frac{\pi_t\left(\chi,\chi^{'}\right)}{\psi}+\sum_{s=1}^{\infty}\left(\beta\nu\right)^s\left[\frac{\pi_{t+s}\left(\chi,\chi^{'}\right)}{\psi}-1\right],0\right\} \tag{3.17}$$

3.2.4 Dynamic market equilibrium definition

Having characterized the dynamics of firm matching, we can now define a dynamic market equilibrium as follows.

Definition 2. Given an initial matching function $m_{-1} : S_\chi \times S_\chi \to [0, 1]$, a dynamic market equilibrium of the model is a list of sequences of matching functions $\{m_t\}_{t=0}^{\infty}$, acceptance functions $\{a_t\}_{t=0}^{\infty}$, profit functions $\{\pi_t\}_{t=0}^{\infty}$, and network characteristic functions $\{\Phi_t, \Delta_t\}_{t=0}^{\infty}$, as well as a list of scalars $\{\Delta_{Ht}\}_{t=0}^{\infty}$, all of which satisfy equations (2.15), (2.16), (2.28), (3.3), (3.5), and (3.14). Given the matching function m_t, the allocation at date t in a dynamic equilibrium is as defined in the static model.

Similarly, we can define a steady-state of the dynamic model as a dynamic market equilibrium in which all variables in Definition 2 are constant.

Definition 3. A steady-state equilibrium of the dynamic model is a matching function m, an acceptance function a, a profit function π, network characteristic functions $\{\Phi, \Delta\}$, as well as a scalar Δ_H, all of which satisfy equations (2.15), (2.16), (2.28), (3.4), (3.5), and (3.12). Given the steady-state matching function m, the allocation in a steady-state equilibrium is as defined in the static model.

The computational algorithms used to solve for both the steady-state and transition dynamics of the dynamic market equilibrium are described in detail in section A.2 of the online appendix. Note that once the matching function is endogeneized, Blackwell's conditions can no longer be applied to establish the contraction mapping property of the network characteristic equations (2.15) and (2.16). Therefore, establishing uniqueness of the solution to these equations and hence of the dynamic market equilibrium is not trivial. Nonetheless, numerical solution of the steady-state of the dynamic market equilibrium is only marginally more computationally demanding than solving for the static market equilibrium, and numerical simulations reveal no counterexample to the supposition of uniqueness.

3.2.5 Dynamic market equilibrium efficiency

To characterize the efficiency of the dynamic market equilibrium, we can again compare the resulting allocation with the allocation that would be chosen by a social planner subject to the same static and dynamic constraints faced by firms. Recall from Proposition 3 that the static market equilibrium is inefficient relative to the social planner's allocation because of the monopoly markups charged by firms. The same static inefficiency characterizes the market equilibrium allocation in each period of the dynamic model.

In the dynamic setting, however, an additional potential source of inefficiency arises because the criterion by which firms select relationships may differ from that employed by the social planner. To study this, one can thus compare the cutoff value for the relationship cost shock chosen by firms, given by equation (3.17), to the cutoff value that would be chosen by the planner. In section B.2 of the appendix, I show that the planner's solution is characterized by the following proposition.

Proposition 4. *The cutoff value for the cost shock at date t chosen by the social planner is given by:*

$$\xi_{max,t}^{SP}\left(\chi,\chi^{'}\right)=\max\left\{\frac{\pi_t^{SP}\left(\chi,\chi^{'}\right)}{\psi}+\sum_{s=1}^{\infty}\left(\beta\nu\right)^s\left(\frac{\mathcal{C}_{t+s}}{\mathcal{C}_t}\right)\left[\frac{\pi_{t+s}^{SP}\left(\chi,\chi^{'}\right)}{\psi}-1\right],0\right\} \tag{3.18}$$

where π_t^{SP} is the planner's analog of the profit function:

$$\pi_t^{SP}\left(\chi,\chi^{'}\right)\equiv\left(\frac{\alpha^{\sigma-1}}{\sigma-1}\right)\Delta_{H,t}^{SP}\Delta_t^{SP}\left(\chi\right)\Phi_t^{SP}\left(\chi^{'}\right) \tag{3.19}$$

and $\mathcal{C}_t$ is a measure of the total connectivity between firms in the economy:

$$\mathcal{C}_t\equiv\left[\int_{S_\chi}\int_{S_\chi}\left[\sum_{d=0}^{\infty}\alpha^{d(\sigma-1)}m_t^{SP,(d)}\left(\chi,\chi^{'}\right)\right]\left(\delta\phi^{'}\right)^{\sigma-1}dG_\chi\left(\chi\right)dG_\chi\left(\chi^{'}\right)\right]^{\frac{1}{\sigma-1}} \tag{3.20}$$

Comparing equations (3.17) and (3.18), we see that the criterion by which firms select relationships in the market equilibrium differs from the socially-optimal criterion in two ways. First, because of the monopoly markup distortion discussed in section 2.2.5, the static social value of a given relationship (measured by π^{SP}) differs from the value of profits by which selling firms value relationships in the market equilibrium. Note that holding fixed the network productivity of the selling firm and the network quality of the buying firm, the functions π_t^{SP} and π_t differ only by a constant term $\mu^{-\sigma}$.

Second, the planner internalizes the effect of each relationship on all other firms in the production network whereas firms in the market equilibrium do not. To better understand the nature of this network externality, it is useful to consider the social value of a given relationship at date t, which can be characterized by the static marginal change in household utility resulting from a marginal increase in the mass of active relationships between firms of given states. In the proof of Proposition 4, I show that this is given by:

$$\frac{dU_t}{d\bar{m}_t\left(\chi,\chi^{'}\right)}=\mathcal{C}_t\left[\pi_t^{SP}\left(\chi,\chi^{'}\right)-\psi\right] \tag{3.21}$$

where $\bar{m}_t\left(\chi,\chi^{'}\right)\equiv m_t\left(\chi,\chi^{'}\right)g_\chi\left(\chi\right)g_\chi\left(\chi^{'}\right)$ denotes the total mass of connections between χ-firm buyers and $\chi^{'}$-firm sellers. From equation (3.21), we see that the social value of each relationship is equal to the difference $\pi_t^{SP}-\psi$ amplified by the aggregate connectivity measure $\mathcal{C}_t$. Intuitively, when firms are more connected to each other ($\mathcal{C}_t$ is larger), the activation or termination of a single relationship has larger aggregate effects. Since the amplification term $\mathcal{C}_t$ potentially varies across time, the planner values changes in the extensive margin of firm relationships accordingly. This effect appears through the term $\frac{\mathcal{C}_{t+s}}{\mathcal{C}_t}$ in equation (3.18) but is absent in firms' decision making processes about which relationships to activate and terminate at each date.

4 Data and Structural Estimation

4.1 Data

The data used for structural estimation of the model's parameters are sourced from two overlapping datasets. The first is provided by Standard and Poor's Capital IQ platform, which collects fundamental data on a large set of companies worldwide, covering over 99% of global market capitalization. For a subset of these firms, both public and private but located mostly in the US, the database also records supplier and customer relationships based on a variety of sources, such as publicly available financial forms, company reports, and press announcements. From this database, I select all firms in the continental US for which relationship data is available and average revenue from 2003-2007 is positive. This gives me a dataset comprising 8,592 firms with \$16.3 trillion in total revenue, accounting for 54% of total non-farm US business revenue.

The second dataset is based on information from the Compustat platform, which is also operated by Standard and Poor. The Compustat database contains fundamental information for publicly-listed firms in the US, compiled solely from financial disclosure forms, and includes firms' own reports of who their major customers are. In accordance with Financial Accounting Standards No. 131, a major customer is defined as a firm that accounts for at least 10% of the reporting seller's revenue. The Compustat relationship data has been processed and studied by Atalay et al (2011), and contains 103,379 firm-year observations from 1979 to 2007.

Both the Capital IQ and Compustat datasets have their advantages and disadvantages. The Capital IQ platform offers greater coverage of firms with relationship data, as the database includes both public and private firms and records relationships based on sources other than financial disclosure forms. However, the main drawback of the dataset is that it is not possible to tell whether a particular relationship reported in a given year is still active at a later date. The Compustat data, on the other hand, is in panel form and therefore allows one to track the creation and destruction of trading relationships across time. The main weakness of the Compustat data is the 10% truncation level, which implies that a firm cannot have more than 10 customers reported in a given year, although there is still substantial variation in the number of recorded suppliers a firm has. For these reasons, I treat the capital IQ data as cross-sectional and primarily use it to estimate the steady-state of the model. I use the Compustat data to measure dynamic moments that are also used in the estimation.

4.2 Parametric assumptions

To proceed with the structural estimation, I first impose two sets of parametric assumptions, one concerning the distribution of fundamental firm states G_χ, and the other regarding the distribution of the stochastic component of the relationship cost G_ξ.

First, given that the empirical firm size distribution has a log-normal shape (see Figure 2 below), I assume that the log of fundamental firm productivities and demands, ϕ and δ, are also jointly Gaussian. Note that in the empty network with $m\left(\chi,\chi^{'}\right)=0$ for all $\chi,\chi^{'}\in S_\chi$, this assumption would imply that firm revenue is exactly log-normally distributed. Furthermore, one can easily verify that the model is invariant to jointly scaling the population size L, the mean relationship cost ψ, and the mean of firm fundamental characteristics. As such, the mean of the distribution of firm fundamental characteristics is normalized to zero. In addition, I adopt a sparse parameterization of the model by assuming that ϕ and δ are uncorrelated, and that their marginal distributions share the same variance parameter v^2.

Parameterization of the fixed relationship cost is as follows. First, I assume that the stochastic component of the cost shock ξ_t has a Weibull distribution with shape parameter s_ξ. The Weibull distribution has a simple economic interpretation as the minimum amongst a series of cost draws for a given relationship. The scale parameter of the distribution is then chosen so that the mean of ξ_t is equal to one.

4.3 Estimation procedure

With these parametric assumptions, the model developed in sections 2 and 3 has 8 parameters: (1) the variance of fundamental firm characteristics, v^2; (2) the mean relationship cost, ψ; (3) the shape of the relationship cost distribution, s_ξ; (4) the reset friction, ν; (5) the household discount factor, β; (6) the labor supply, L; (7) the elasticity of substitution, σ; and (8) the input suitability parameter, α.

I first describe the set of parameters for which values are not estimated from data. First, observe from either equation (3.12) or (3.14) that the parameters β and ν cannot be separately identified, as it is only the product $\beta\nu$ that matters for the dynamic optimization problem of the firm. Since the Compustat data is of annual frequency, I therefore set $\beta=.95$

and estimate ν from data. Second, since the model is scale invariant, the labor supply L is normalized to one. Third, since the Capital IQ and Compustat data do not contain trade transaction values from which substitution elasticities are typically estimated, I set the value of σ to 4, which is a typical value estimated in the literature.[15]

Finally, note that the input suitability parameter α can in principle be estimated from the available data, as it has an intuitive connection to a moment that can be empirically observed: when α is larger, more firm-to-firm relationships are likely to form. A potential indeterminacy arises, however, from the fact that a large number of active relationships can also be rationalized in the model by a low value of the mean relationship cost ψ. To avoid this indeterminacy in the estimation procedure, I therefore normalize α to a value arbitrarily close to but less than one, and rely on data to estimate the magnitude of the mean relationship cost instead. This approach can be interpreted as assuming that the cost of forming a relationship embodies not only the resources that need to be devoted to managing that relationship, but also the costs of technological innovation - design of prototypes and customization of products, for example - that are required for the seller's good to be used in the buyer's production process.

The remaining 4 parameters of the model - υ, ψ, s_ξ, and ν - are then estimated from the Capital IQ and Compustat data using a simulated method of moments approach, targeting the following four sets of moments. First, the distribution of firm revenue normalized by its mean. Second, the distributions of in-degree (number of suppliers) and out-degree (number of customers). Third, the joint distributions of firm size and relationship retention rates (the fractions of suppliers and customers that are retained year-to-year). Fourth, the joint distribution of firm size and relationship creation rates (the fractions of suppliers and customers that are new year-to-year). The first two sets of moments (static) are computed from the Capital IQ data, while the remaining two sets of moments (dynamic) are computed from the Compustat data.

Note that the dispersion of the firm and degree distributions are directly influenced by the dispersion of firm fundamental characteristics υ, while the dynamic moments are directly impacted by the volatility of the relationship cost shock s_ξ and the reset friction ν. These three parameters therefore have clear and intuitive connections to the data. The mean relationship cost ψ, which controls the overall level of connectivity in the production network, also has in principle a direct connection to the empirical average degree count. However, this is complicated by the fact that the degree count is continuous in the model but discrete in the data. To deal with this problem, I adopt a slight modification to the standard simulated method of moments approach in order to estimate ψ. Specifically, given a set of values for

[15]See for example Broda and Weinstein (2006).

(υ, s_ξ, ν), the estimation algorithm searches for the value of ψ that generates a labor share of 0.7 in the model instead. In addition, the targeted degree distributions are normalized by their mean. For a detailed description of the estimation algorithm, refer to section C of the online appendix.

4.4 Estimation results

4.4.1 Parameter values

The estimated parameter values are shown in Table 1, together with standard errors that are computed using a bootstrapping technique described in section C of the online appendix. Figure 1 also shows contour plots of the objective function minimized by the estimation algorithm in (υ, s_ξ, ν) space. The objective function is well-behaved within a neighborhood of the estimated parameter value set, and starting the estimation algorithm from different initial parameter values yields almost identical parameter estimates.

Parameter		Value	Standard Error
standard deviation of (ϕ, δ) distributions	υ	.887	.029
mean relationship cost	ψ	.216	.027
shape of ξ distribution	s_ξ	.957	.135
reset friction	ν	.164	.011

Table 1: Estimated parameter values

4.4.2 Model fit

Targeted moments Figures 2-5 show the model's fit of the firm size distribution, degree distributions, relationship retention rates, and relationship creation rates respectively. First, note that even though the distribution of fundamental firm characteristics (ϕ, δ) in the model is assumed to be log-normal, the resulting firm size distribution deviates from log-normality because the distribution of linkages is heterogeneous across firms. Nonetheless, the model generates a reasonably close approximation to the empirical firm size distribution, with the deviation growing larger only in the lower tail.

Second, the model matches well the shape of the normalized degree distributions, although it underpredicts the extent of connectivity of the most connected firms. Third, the model's predictions are consistent with the empirical observation that larger firms tend to retain larger fractions of their customers and suppliers year-to-year, and also have smaller fractions of new relationships relative to the size of their existing customer and supplier bases. The model's fit of the joint distributions of firm size and relationship retention and

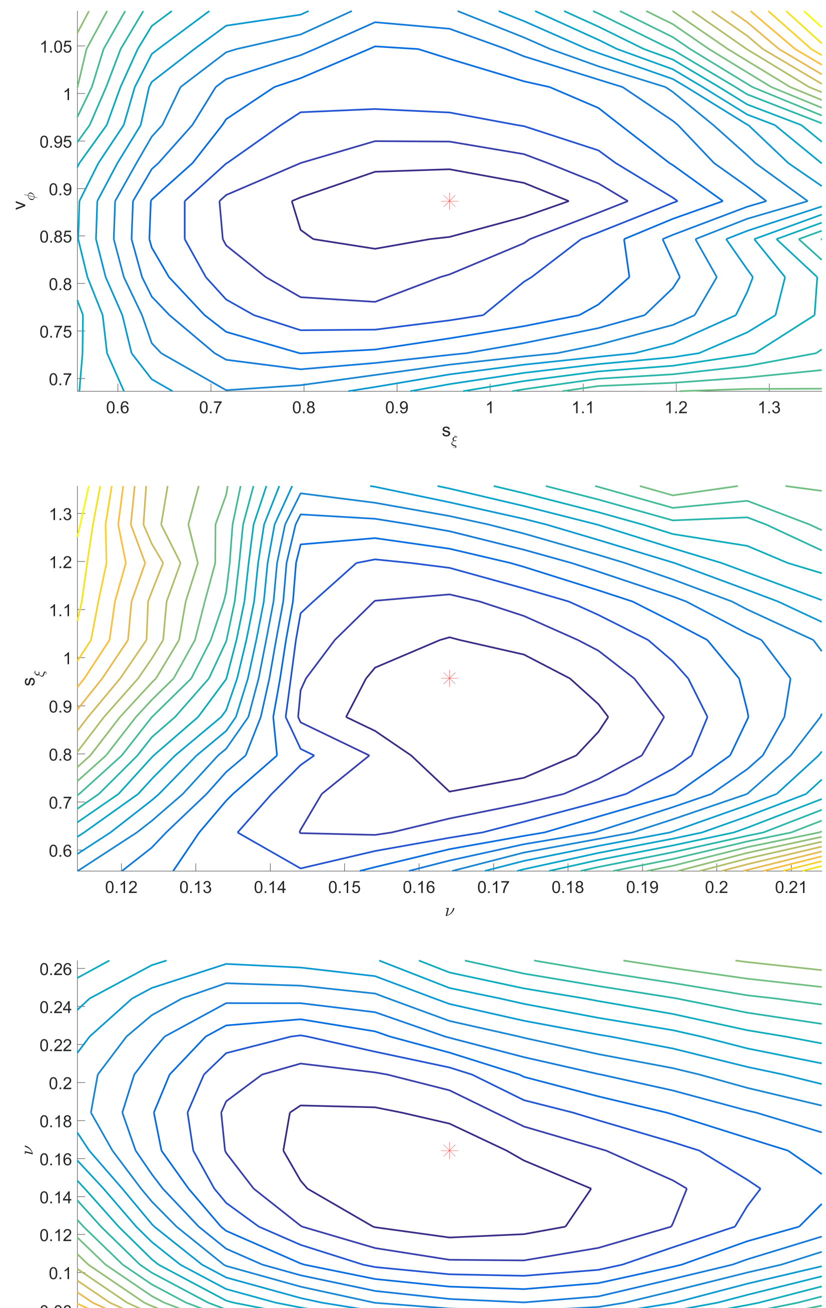

Figure 1: Contour plots of objective function in (v, s_ξ, ν) space
Asterisks indicate estimated parameter values

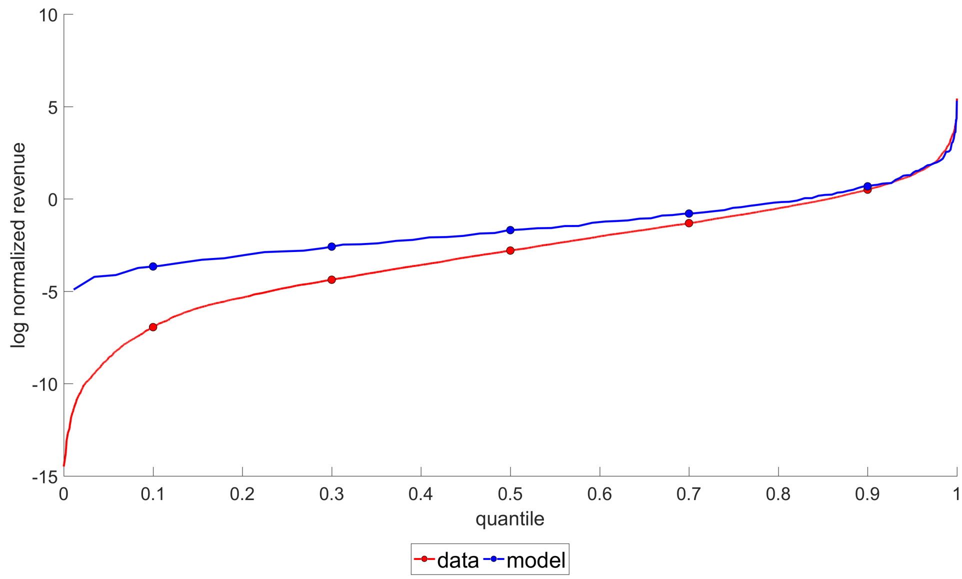

Figure 2: Firm revenue distribution

creation rates is not exact, but it matches well the average relationship retention and creation rates implied by these dynamics[16]

Untargeted moments As an overidentification check on the model's predictions, I examine here the fit to moments that are untargeted in the estimation procedure. First, Figure 6 shows the fit of the firm employment distribution. As with the empirical distribution of firm revenue, the distribution of employment in the dataset is well-approximated by a log-normal distribution, which the model approximately replicates. Second, Figure 7 shows the model's fit of the joint distributions of firm size and degree. While the fit is not exact, the model is nonetheless consistent with the empirical pattern that larger firms tend to have more customers and suppliers, as might be expected.

Finally, Figures 8 and 9 show the model's fit of the matching assortativity between firms, which characterizes whether larger and more connected firms are connected to firms that are also larger and more connected (positive matching), or to firms that are smaller and less connected (negative matching). Here, the empirical pattern of matching assortativity differs depending on whether matching is characterized by firm size or by connectedness. Measured by revenue, larger firms tend to have larger customers and suppliers than smaller firms, so that matching is positive. However, measured by degree, firms with more suppliers

[16]Mean retention rates are .59 for suppliers and .65 for customers in the data, .64 for suppliers and .66 for customers in the model. Mean creation rates are .44 for suppliers and .36 for customers in the data, and .36 and .34 in the data.

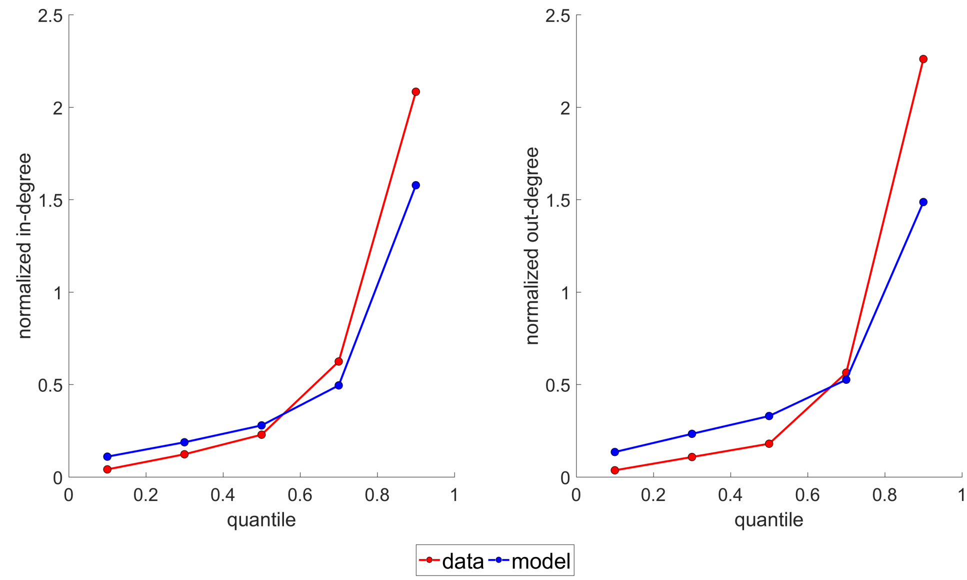

Figure 3: Firm degree distributions

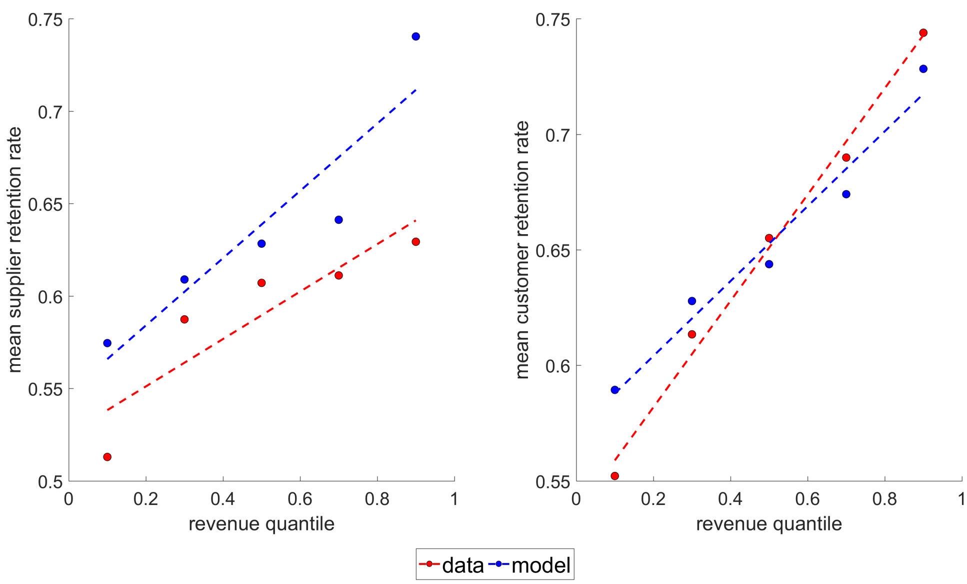

Figure 4: Joint distribution of firm size and supplier/customer retention rates

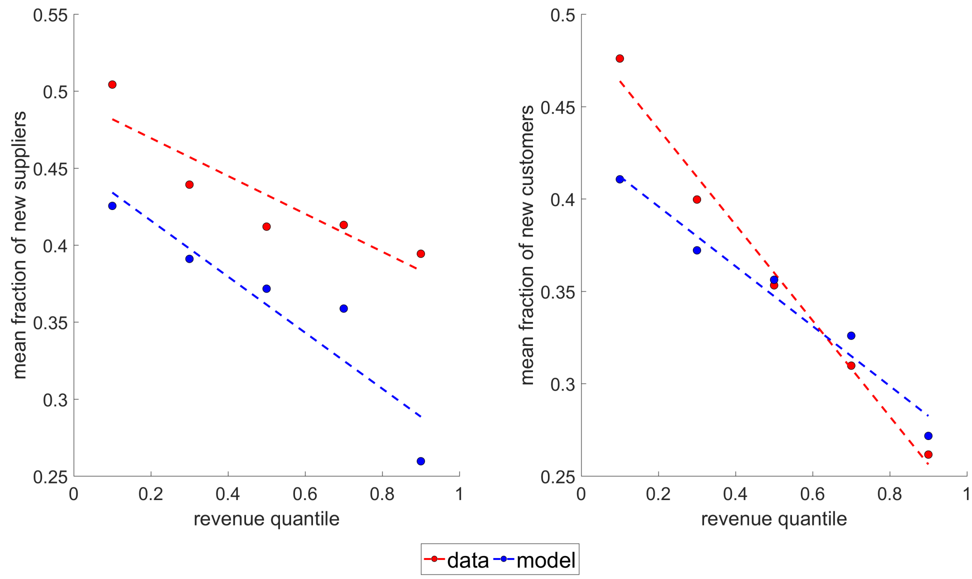

Figure 5: Joint distribution of firm size and fraction of new suppliers/customers

tend to have suppliers that have *fewer* suppliers themselves, and similarly with matching to customers, so that matching assortativity in this case is negative. With the assumed parametric form for the distribution of the relationship cost, the model replicates the latter pattern but not the former.[17]

5 Counterfactual Exercises

I now employ the model to study the aggregate effects of firm-level supply and demand shocks, with the goal of understanding how accounting for relationship heterogeneity, supply chain heterogeneity, and relationship dynamics in the production network affects the magnitudes of these effects. In what follows, I focus on a particular set of counterfactual exercises. Starting from the steady-state of the model corresponding to the parameter values estimated above, I first group the set of firms in the economy according to deciles of the firm size distribution. The model is then simulated to study the effects on household welfare of permanent and unanticipated changes in the fundamental productivities ϕ (supply shocks) or fundamental demands δ (demand shocks) of each group of firms.

As a standardization, the magnitude of the shock in each simulation is equal to one

[17]With alternative parametric choices for G_ξ, the model can generate positive revenue matching. An example is the Gompertz or log-Weibull distribution, although economic interpretation of this functional form is less straightforward.

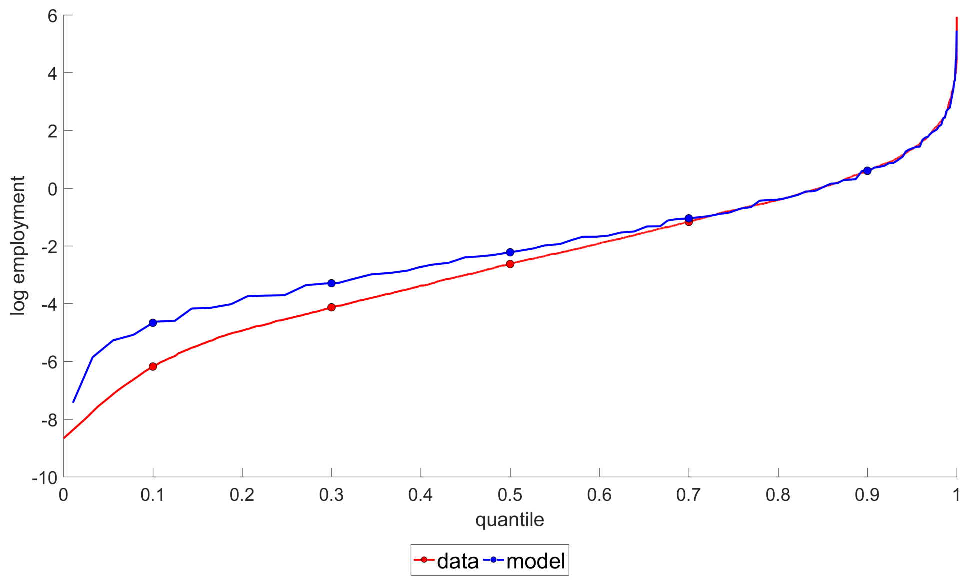

Figure 6: Firm employment distribution

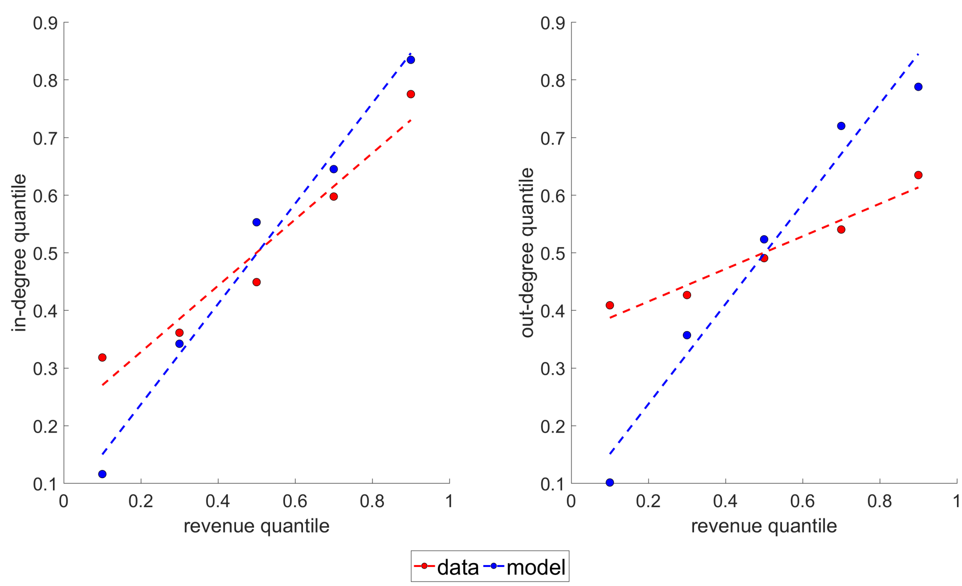

Figure 7: Joint distributions of firm revenue and degree

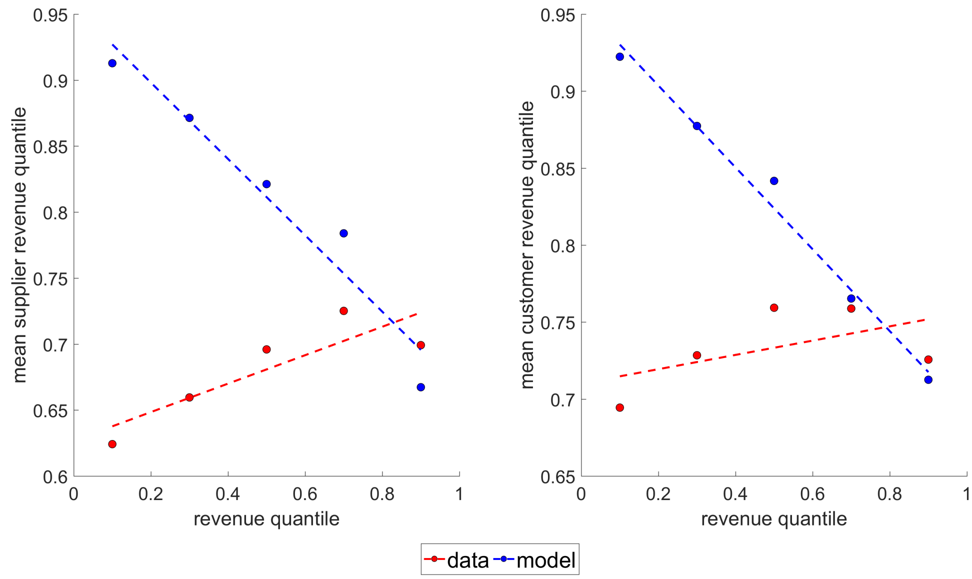

Figure 8: Firm matching assortativity (revenue)

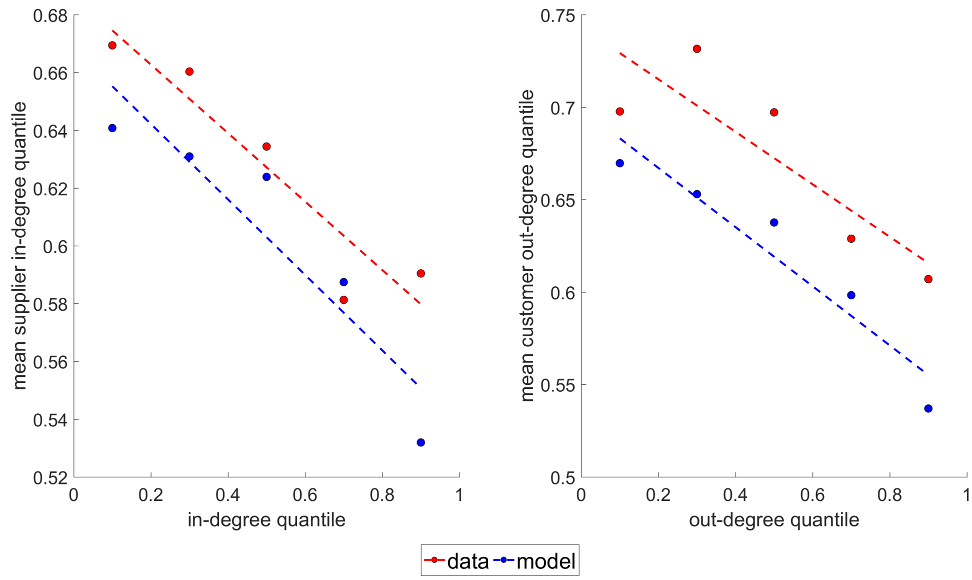

Figure 9: Firm matching assortativity (degree)

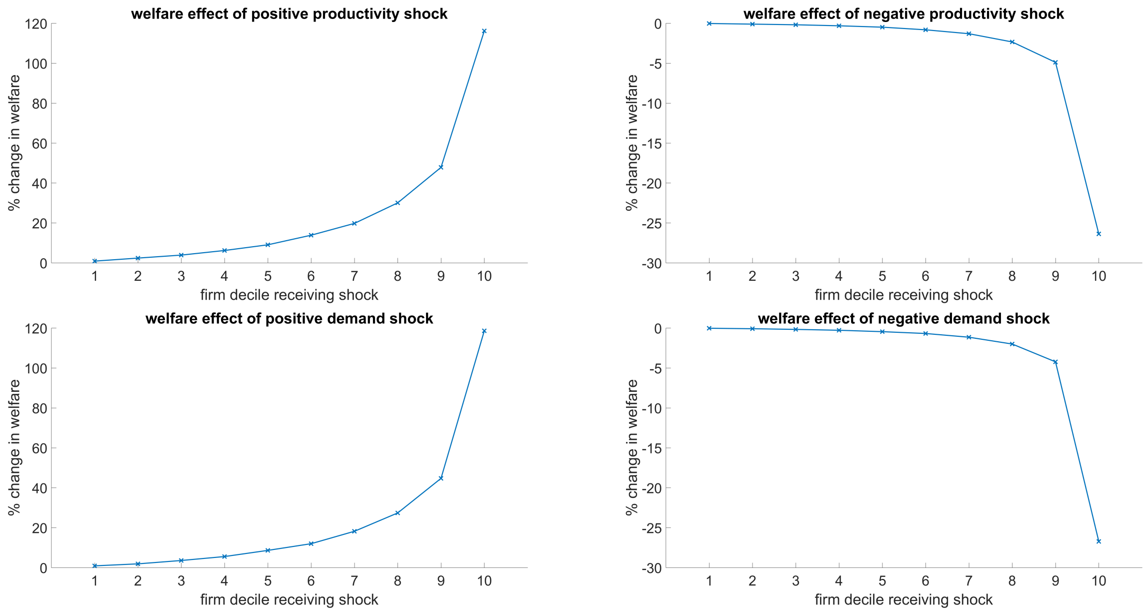

Figure 10: Baseline welfare effects of one standard deviation supply and demand shocks

standard deviation of the log of the relevant firm characteristic distribution.[18] Each shock is also assumed to occur at $t = 0$ after all relationships in that period have been set, with firms allowed to adjust their sets of active relationships only from $t = 1$ onwards. This allows us to disentangle the short-run effects of the shock (with the production network taken as fixed) from its long-run effects (once endogenous relationship adjustment is taken into account).

The baseline results of these simulations are summarized in Figure 10, which shows the percentage changes in the present value of welfare (integrated over the corresponding transition paths) resulting from each counterfactual shock. As might be expected, shocks to large firms have much greater effects on household welfare than small firms. We now examine how the structure and dynamics of the production network matter for these welfare responses.

5.1 Relationship heterogeneity

In both the model and the data, relationships are distributed heterogeneously across firms, with larger firms connected to more buyers and suppliers than smaller firms. To examine the quantitative importance of accounting for this feature of the production network, I consider an alternative model of production where the matching function is exogenously given as $m\left(\chi, \chi'\right) = \bar{m}$ for all $\left\{\chi, \chi'\right\} \in S_{\chi}^{2}$, so that firms are identical in their connectivity to other firms regardless of their characteristics. Note that this is equivalent to the assumption

[18]For example, a decline in ϕ for the set of firms considered by a factor of e^{-v}.

that all firms produce using labor and a common composite intermediate input, which is a standard assumption in many models featuring intermediate input trade. One can therefore interpret this as a "market model" of production instead of the "network model" developed in this paper.

Taking this market model, I then re-estimate the relevant structural parameters using the same simulated method of moments approach, targeting the firm size distribution in the data (which pins down the variance of firm fundamental characteristics v^2) and an aggregate labor share of 0.7 (which pins down the common level of connectivity $\bar{m}$). Simulations of the market model are then used to compute the welfare effects of the same counterfactual firm-level shocks described above.

The results of these simulations are summarized in Figure 11, which shows for each counterfactual shock the percentage changes in welfare in both the network and market models. The key takeaway from this analysis is that accounting for the heterogeneous distribution of relationships across firms leads to lower predicted effects of shocks to small firms and larger predicted effects of shocks to large firms. This is intuitive, as large firms are central to the production network not only because they have the best fundamental characteristics, but also because shocks to these firms affect a larger number of other firms either upstream or downstream. In terms of magnitudes, the deviations of the predicted welfare effects in the network and market models can be large. The market model under-predicts the welfare effects of shocks to firms in the largest decile by between $10\% - 20\%$, while over-predicting the welfare effects of shocks to small firms by even greater percentages.

5.2 Supply chain heterogeneity

Not only are relationships heterogeneously distributed across firms, but firms also occupy different positions in supply chains of varying lengths. A decline in a firm's labor productivity, for example, leads to an increase in its marginal cost and induces it to raise the price of its output, which in turn leads to an increase in the marginal costs of firms that it supplies and hence to further changes in prices downstream. The structure of the model developed above offers a simple way of decomposing the aggregate effects of such shocks into changes along each stage of the relevant supply chains.

To illustrate, consider the short-run effects (fixed m) of a shock to firm fundamental productivities, whereby a firm with original state $\{\phi, \delta\}$ has post-shock state $\left\{\hat{\phi}\left(\phi\right), \delta\right\}$. The immediate consequence of this shock is to change the marginal costs of the firms directly affected by the shock. If no firms change their intermediate input prices in response, however,

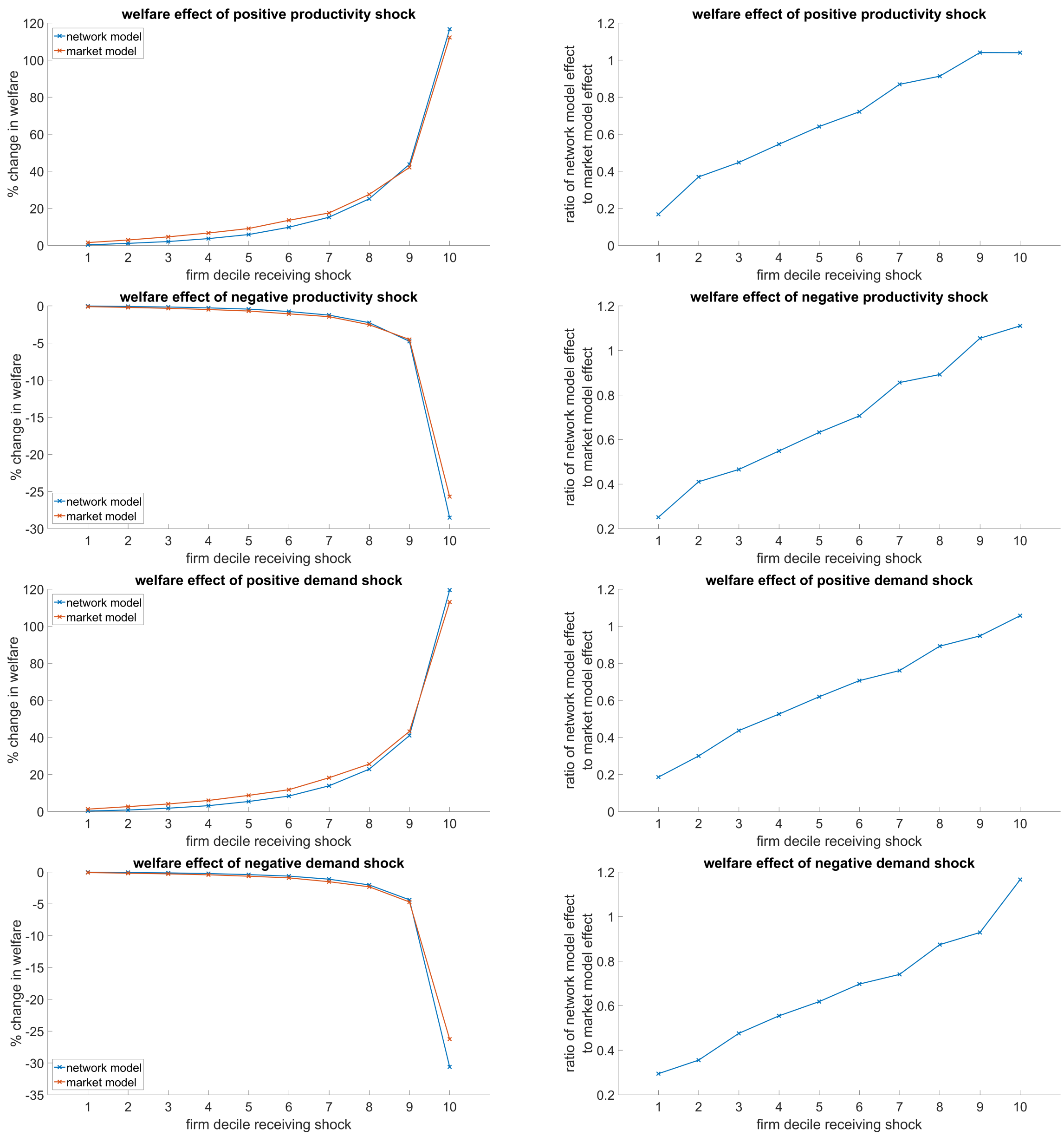

Figure 11: Market model versus network model welfare effects

the change in network productivities is then given by:

$$\hat{\Phi}^{(0)}(\chi) = \hat{\phi}(\phi)^{\sigma-1} + \left(\frac{\alpha}{\mu}\right)^{\sigma-1} \int_{S_\chi} m\left(\chi, \chi'\right) \Phi(\chi)\, dG_\chi\left(\chi'\right) \tag{5.1}$$

where Φ is the pre-shock network productivity function and the superscript (0) denotes the zeroth-order effect of the shock. Now, if only firms that are directly affected by the shock change their intermediate input prices, the first-order change in network productivities is then given by:

$$\hat{\Phi}^{(1)}(\chi) = \hat{\phi}(\phi)^{\sigma-1} + \left(\frac{\alpha}{\mu}\right)^{\sigma-1} \int_{S_\chi} m\left(\chi, \chi'\right) \hat{\Phi}^{(0)}(\chi)\, dG_\chi\left(\chi'\right) \tag{5.2}$$

Extending this logic, the effect on firm network productivities due to price changes occurring up to n stages downstream of the set of firms directly affected by the shock is given recursively by:

$$\hat{\Phi}^{(n)}(\chi) = \hat{\phi}(\phi)^{\sigma-1} + \left(\frac{\alpha}{\mu}\right)^{\sigma-1} \int_{S_\chi} m\left(\chi, \chi'\right) \hat{\Phi}^{(n-1)}(\chi)\, dG_\chi\left(\chi'\right) \tag{5.3}$$

with initial condition $\hat{\Phi}^{(-1)} = \Phi$. Analogously, following a demand shock in which a firm with original state $\{\phi, \delta\}$ has post-shock state $\left\{\phi, \hat{\delta}(\delta)\right\}$, the effect on firm network demands due to propagation of the shock up to n stages upstream of the set of firms directly affected by the shock is given recursively by:

$$\hat{\Delta}^{(n)}(\chi) = \mu^{-\sigma}\hat{\delta}(\delta)^{\sigma-1} + \mu^{-\sigma}\alpha^{\sigma-1} \int_{S_\chi} m\left(\chi', \chi\right) \hat{\Delta}^{(n-1)}(\chi)\, dG_\chi\left(\chi'\right) \tag{5.4}$$

with initial condition $\hat{\Delta}^{(-1)} = \Delta$.

In other words, each value function iteration (which is already employed to solve the model in the first place) captures successively higher-order effects of shock propagation downstream or upstream. To quantify the importance of firms' heterogeneous positions in their respective supply chains, one can therefore simply study the effects of a shock at each stage in this iterative process. The results of this exercise are summarized in Figure 12, which shows the zeroth- and first-order effects relative to the overall short-run effect.

As might be expected, the zeroth-order effect provides a poor approximation to any supply or demand shock. Perhaps somewhat surprisingly, however, the analysis also suggests that the higher-order propagation of firm-level productivity and demand shocks is quantitatively unimportant. In particular, first-order approximations that account for shock propagation only one stage upstream or downstream typically deviate from the overall general equilibrium short-run effect by less than 5%, and effects of subsequently higher order quickly approach the total effect.

Intuition for this finding is provided by the expressions for the firm network characteristics in equations (2.17) and (2.18) or for welfare in equation (2.32). From this, one observes that the rate at which shocks to firm fundamental characteristics decay downstream and upstream of a supply chain are governed by the values of $\mu^{1-\sigma}$ and $\mu^{-\sigma}$ respectively.[19] The downstream decay parameter $\mu^{1-\sigma}$ is strictly decreasing in σ, and even for a value of σ as low as 2, the decay parameter is only as large as 0.5. The upstream decay parameter $\mu^{-\sigma}$, on the other hand, is strictly increasing in σ, but even for a value of σ as large as 30, the decay parameter is only as large as 0.36. Consequently, for reasonable values of σ, higher-order effects diminish rapidly relative to the direct effect of the shock.

5.3 Relationship dynamics

While higher-order propagation of shocks upstream or downstream of supply chains appears to be quantitatively unimportant with the network held fixed, the same need not be true of the dynamic propagation of shocks once the endogenous response of the production network is taken into account. To examine this, I compare for each counterfactual simulation the percentage changes in the present value of welfare following adjustment of the network (the long-run effect) with the corresponding changes in the initial period of the shock with the matching function held fixed (the short-run effect).

The results of this analysis are summarized in Figure 13. Here, we see that the predicted welfare effects of firm-level shocks can differ greatly once the dynamic response of the network is taken into account. For example, the welfare gains from positive supply and demand shocks to the smallest decile of firms are more than three times as large in the long-run versus the short run (although in absolute terms both effects are small), and are 50% larger in the long-run for similar shocks to firms in the middle of the size distribution. The simulations suggest that the discrepancy between the long- and short-run effects are greater in percentage terms for smaller firms, but since the absolute magnitude of the welfare response following shocks to large firms is typically also large, taking into account the dynamic network response can be important in this case as well. For instance, the short-run welfare loss following a negative supply shock to the largest decile of firms leads to a 32% welfare loss in the short-run, but adjustment of the network leads to a long-run welfare loss that is five percentage points lower.

[19] The decay rate also depends on the value of $\alpha^{\sigma-1}$, but here we normalized $\alpha \approx 1$, which would make the decay rate as slow as possible given other model parameters.

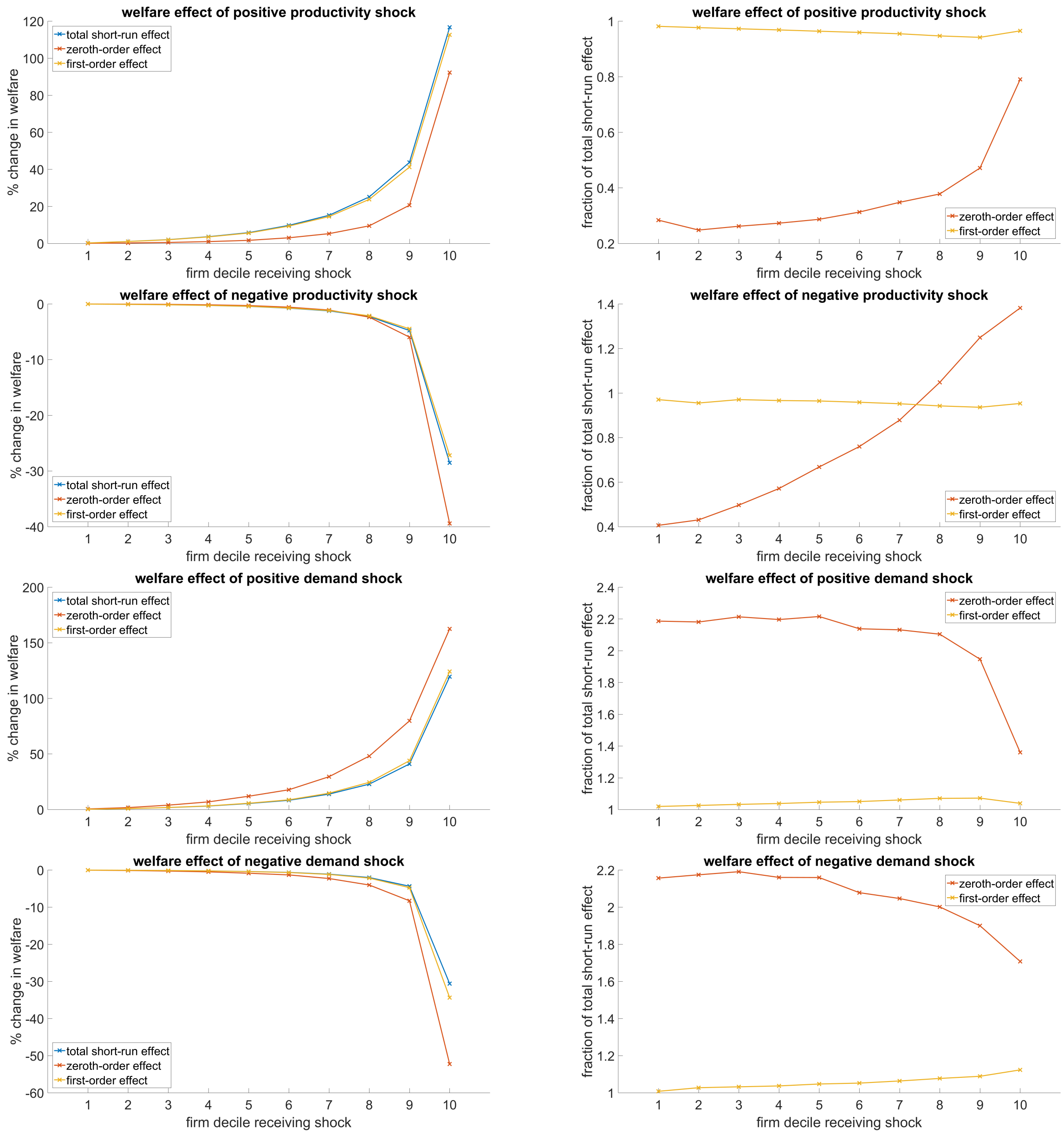

Figure 12: Propagation of shock effects holding the production network fixed

Interestingly, one also observes that the ability of firms to adjust trading relationships in response to shocks need not imply that the welfare losses following negative shocks are smaller in the short-run than in the long-run. In fact, we see from the simulations that for negative shocks to smaller firms in the economy, the converse is true. This follows from the fact that the market equilibrium is inefficient, as discussed above, and therefore there is no guarantee that removing the constraint of a fixed network will lead to greater welfare. In simulations of the planner's solution to the same supply and demand shocks, short-run welfare is always weakly lower than long-run welfare.

6 Conclusion

This paper offers a new theory of how heterogeneous firms create and destroy trading relationships with one another, and how these firm-level decisions influence the structure of the production network and its evolution over time. Despite the rich heterogeneity in relationships and endogenous dynamics, tractability is preserved, which enables structural estimation of the model and flexibility in simulating a range of counterfactual exercises.

The numerical analysis highlights how the structure and dynamics of the production network matter for the propagation of firm-level supply and demand shocks, with three key takeaways. First, the largest firms are also the most connected, and taking this relationship heterogeneity into account implies stronger effects of shocks to these firms. Second, although firms are heterogeneous in their supply chains, supply and demand shocks dissipate quickly upstream and downstream, and first-order approximations capturing effects only one stage along a supply chain account for a large fraction of the short-run effects. Third, the dynamic propagation of shocks is quantitatively important, as the aggregate effects of firm-level shocks can differ markedly once the endogenous adjustment of the production network is taken into account.

The issues discussed in this paper also provide scope for future research, with two areas in particular warranting further investigation. First, given that the market equilibrium of the model is shown to be inefficient, a natural question is whether there are market structures which lead to efficient outcomes. In this paper, the assumption of monopolistic competition and the associated constant markups is essential for tractability. Nonetheless, one must wonder whether tractable bargaining games between a large number of firms in a network can be developed. Moving away from constant markups would also allow the study of competition effects in production networks, which has not been addressed in depth in the literature.

Second, the modeling of relationship stickiness in this paper is a reduced-form approach

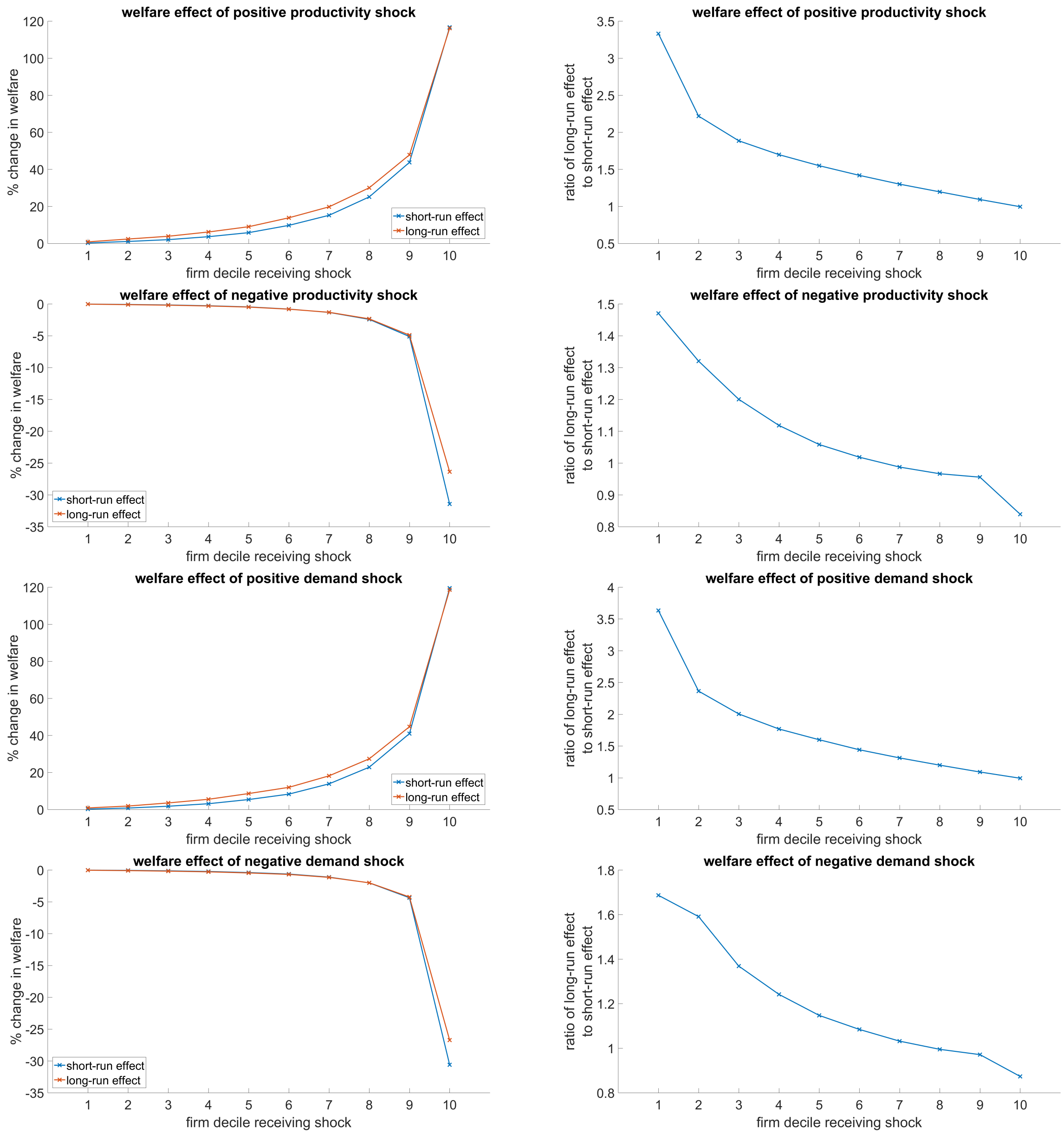

Figure 13: Long-run versus short-run welfare effects

towards capturing the idea that various frictions impede the creation and destruction of trading relationships. Understanding the microfoundations of these frictions requires further work and would likely yield new insights. For example, if these frictions have to do with the availability of information about potential buyers and sellers, then the frictions themselves must be endogenous, since surely information propagates through the network in a way that depends on its structure and dynamics.

References

Acemoglu, Daron, Vasco M Carvalho, Asuman Ozdagalar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica*, 80(5): 1977-2016.

Atalay, Enghin, Ali Hortaçsu, James Roberts, and Chad Syverson. 2011. "Network Structure of Production." *Proceedings of the National Academy of Sciences*, 108(13): 5199-5202.

Aumann, Robert J, and Roger Myerson. 1988. "Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value," in Roth, Al (ed.), *The Shapley Value*, Cambridge University Press.

Bala, Venkatesh, and Sanjeev Goyal. 2000. "A Noncooperative Model of Network Formation." *Econometrica*, 68(5): 1181-1229.

Barabási, Albert-László, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science*, 286 (5439): 509–512.

Bernard, Andrew, Andreas Moxnes, and Yukiko Saito. 2015. "Production Networks, Geography, and Firm Performance." Mimeo: Dartmouth College.

Bernard, Andrew, Andreas Moxnes, and Karen Helene Ulltveit-Moe. 2015. "Two-sided Heterogeneity and Trade." Mimeo: Dartmouth College.

Broda, Christian, and David E Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics*, 121(2): 541-585.

Chaney, Thomas. 2015. "The Gravity Equation in International Trade: An Explanation." Mimeo: Toulouse School of Economics.

Chaney, Thomas. 2014. "The Network Structure of International Trade." *American Economic Review*, 104(11): 3600-3634.

Currarini, Sergio, Matthew O Jackson, and Paolo Pin. 2010. "Identifying the Roles of Race-based Choice and Chance in High School Friendship Network Formation" *Proceedings of the National Academy of Science*, 107: 4857-4861.

Erdös, Paul, and Alfréd Rényi. 1959. "On Random Graphs." *Publicationes Mathematicae*, 6: 290–297.

Gartner, Inc. 2014a. "Market Share Analysis: Customer Relationship Management Software, Worldwide, 2014." http://www.gartner.com/document/3045522.

———. 2014b. "Market Share Analysis: Supply Chain Management and Procurement Software, Worldwide, 2014." http://www.gartner.com/document/3044217.

Jackson, Matthew O. 2011. "An Overview of Social Networks and Economic Applications," in Benhabib, Jess, Alberto Bisin, and Matthew O Jackson (eds.), *The Handbook of Social Economics*, North Holland Press: Amsterdam.

Jackson, Matthew O. 2005. "A Survey of Models of Network Formation: Stability and Efficiency," in Demange, Gabrielle, and Myrna Wooders (eds.), *Group Formation in Economics: Networks, Clubs, and Coalitions*, Cambridge University Press: Cambridge, Chapter 1.

Jackson, Matthew O, and Alison Watts. 2002. "The Evolution of Social and Economic Networks." *Journal of Economic Theory*, 106(2): 265-295.

Jackson, Matthew O, and Asher Wolinksky. 1996. "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory*, 71(1): 44-74.

Kranton, Rachel E, and Deborah F Minehart. 2001. "A Theory of Buyer-Seller Networks." *American Economic Review*, 91(3): 485-508.

Magerman, Glenn, Karolien De Bruyne, Emmanuel Dhyne, and Kan Van Hove. 2016. "Heterogeneous Firms and the Micro Origins of Aggregate Fluctuations." Mimeo: University of Leuven.

Myerson, Roger. 1991. *Game Theory: Analysis of Conflict*, Harvard University Press: Cambridge, MA.

Oberfield, Ezra. 2015. "Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture." Mimeo: Princeton University.

Rauch, James E. 1999. "Networks Versus Markets in International Trade." *Journal of International Economics*, 48: 7-35.

Schmidt, Karl, Brent Adamson, and Anna Bird. 2015. "Making the Consensus Sale." *Harvard Business Review*, March Issue.

Snyder, Kelsey, and Pashmeena Hilal. 2015. "The Changing Face of B2B Marketing." Web blog post. *Think With Google*, Google Inc..

Watts, Alison. 2001. "A Dynamic Model of Network Formation." *Games and Economic Behavior*, 34: 331-341.

Watts, Duncan J, and Steven H Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature*, 393(6684): 440-442.

Online Appendix to "Firm-to-Firm Trade in Sticky Production Networks"

A Computational Algorithms

A.1 Static algorithm

Given the matching function m and the associated quantity of labor L_f used for relationship costs, the static market equilibrium specified in Definition 1 can be solved for using the following algorithm.

1. Make initial guesses $\hat{\Phi}$ and $\hat{\Delta}$ for the network productivity and quality functions, and iterate on equations (2.15) and (2.16) until convergence.

2. Solve for Δ_H using equation (2.28).

3. Compute the allocation $\left\{l\left(\chi\right), X\left(\chi\right), x\left(\chi, \chi'\right), x_H\left(\chi\right)\right\}_{\chi \in S_\chi}$ using (2.24), (2.25), (2.27), and (2.31) respectively.

Since the functional equations (2.15) and (2.16) constitute contraction mappings with Lipschitz constants $\left(\frac{\alpha}{\mu}\right)^{\sigma-1}$ and $\frac{\alpha^{\sigma-1}}{\mu^{\sigma}}$ respectively, the iteration procedure in step 1 of the algorithm is guaranteed to converge at those rates. In practice, numerical solution of the model requires discretization of the state space S_χ into a mesh grid, of say $N_{grid} \times N_{grid}$ points. One can then solve for the functions $\Phi\left(\cdot\right)$ and $\Delta\left(\cdot\right)$ in step 1 at each point in the mesh grid, and then use bilinear interpolation to obtain numerical approximations of these functions for any desired value of $\chi \in S_\chi$.

A.2 Dynamic algorithm

I first describe the computational algorithm used to solve for the steady-state equilibrium specified in Defintion 3, which is as follows.

1. Make initial guesses $\hat{\Phi}$ and $\hat{\Delta_H \Delta}$ for the network productivity function and the network quality function scaled by the household demand shifter.

2. Compute the implied profit function $\tilde{\pi}$ from equation (3.5).

3. Compute the implied matching and acceptance functions, $\tilde{m}$ and $\tilde{a}$, from equations (3.4) and (3.12).

4. Compute the implied network productivity and quality functions, $\tilde{\Phi}$ and $\tilde{\Delta}$, from equations (2.15) and (2.16).

5. Compute the implied household demand shifter $\tilde{\Delta}_H$ from equations (2.28), (3.15), (3.16), and (3.17), and obtain the implied guess for the scaled network quality function, $\tilde{\Delta_H \Delta} = \tilde{\Delta}_H \tilde{\Delta}$.

6. Compute the residual $\mathcal{R} \equiv \max\left\{\mathcal{R}_\Phi, \mathcal{R}_\Delta\right\}$, where:

$$\mathcal{R}_\Phi \equiv \max_{\chi \in S_\chi} \left|\hat{\Phi}(\chi) - \tilde{\Phi}(\chi)\right|$$
$$\mathcal{R}_\Delta \equiv \max_{\chi \in S_\chi} \left|\hat{\Delta_H \Delta}(\chi) - \tilde{\Delta_H \Delta}(\chi)\right|$$

If $\mathcal{R} > \epsilon$ for some tolerance level ϵ, update the guesses for the network productivity and scaled quality functions according to $\hat{\Phi}' = \tilde{\Phi}$ and $\hat{\Delta_H \Delta}' = \tilde{\Delta_H \Delta}$, and repeat from step 1 until $\mathcal{R} \leq \epsilon$.

I now discuss the computational algorithm used to solve for the model's transition dynamics as specified in Definition 2. Suppose that the matching and profit functions at date 0 are given by m_0 and π_0 respectively, and that the economy is not in steady-state. The goal is to solve for the model's transition path to the eventual steady-state characterized by the matching function denoted by m_{ss}. Note that given the matching function m_t, it is straightforward to solve for the static market equilibrium at date t using the algorithm discussed in section A.1. The challenge in solving the model's transition dynamics therefore lies in computing the matching function at date t given the matching function at date $t-1$. As we see from equation (3.14), doing so requires solving for the profit functions $\left\{\pi_{t+s}\right\}_{s\geq 0}$. To accomplish this, I employ an algorithm that iterates on the path of profit functions $\left\{\pi_t\right\}_{t=1}^{T}$ for some value of T large enough such that the matching function at date T is close enough to the eventual steady-state matching function m_{ss}. Formally, the algorithm is as follows.

1. Make a guess $\hat{T}$ for the number of periods that it takes for convergence to the steady-state.

2. Make an initial guess for the profit functions $\left\{\hat{\pi}_t\right\}_{t=2}^{\hat{T}}$ (e.g. $\hat{\pi}_t = \frac{1}{2}(\pi_0 + \pi_{ss})$ for all $t \in \left\{2, \cdots, \hat{T}\right\}$).

3. At each date $t \in \left\{1, \cdots, \hat{T}\right\}$, given $\hat{m}_{t-1}$ (with $\hat{m}_0 = m_0$):

(a) Make initial guesses $\hat{\Phi}_t$ and $\Delta_H \hat{\Delta}_t$ for the network productivity function and the network quality function scaled by the household demand shifter.

(b) Compute the implied profit function $\tilde{\pi}_t$ from equation (3.5).

(c) Compute the implied acceptance function $\tilde{a}_t$ (3.12), setting $\pi_{t+s} = \hat{\pi}_{t+s}$ for $s \in \left\{1, \cdots, \hat{T} - t\right\}$ and $\pi_{t+s} = \pi_{ss}$ for $s > \hat{T} - t$.

(d) Compute the implied matching function $\tilde{m}_t$ from equation (3.3).

(e) Compute the implied network productivity and quality functions, $\tilde{\Phi}_t$ and $\tilde{\Delta}_t$, from equations (2.15) and (2.16).

(f) Compute the implied household demand shifter $\tilde{\Delta}_{H,t}$ from equations (2.28), (3.15), (3.16), and (3.17), and obtain the implied guess for the scaled network quality function, $\Delta_H \tilde{\Delta}_t = \tilde{\Delta}_{H,t}\tilde{\Delta}_t$.

(g) Compute the residual $\mathcal{R} \equiv \max\{\mathcal{R}_\Phi, \mathcal{R}.\}$, where:

$$\mathcal{R}_\Phi \equiv \max_{\chi \in S_\chi} \left| \hat{\Phi}_t(\chi) - \tilde{\Phi}_t(\chi) \right|$$
$$\mathcal{R}_\Delta \equiv \max_{\chi \in S_\chi} \left| \Delta_H \hat{\Delta}_t(\chi) - \Delta_H \tilde{\Delta}_t(\chi) \right|$$

If $\mathcal{R} > \epsilon$ for some tolerance level ϵ, update the guesses for the network productivity and scaled quality functions according to $\hat{\Phi}_t^{'} = \tilde{\Phi}_t$ and $\Delta_H \hat{\Delta}_t^{'} = \Delta_H \tilde{\Delta}_t$, and repeat from step (a) until $\mathcal{R} \leq \epsilon$, then set $\hat{m}_t = \tilde{m}_t$.

4. Compute the residual:

$$\mathcal{R}_\pi \equiv \max_{t \in \left\{2, \cdots, \hat{T}\right\}} \max_{\left(\chi, \chi^{'}\right) \in S_\chi^2} \left| \hat{\pi}_t\left(\chi, \chi^{'}\right) - \tilde{\pi}_t\left(\chi, \chi^{'}\right) \right|$$

If $\mathcal{R}_\pi > \epsilon_\pi$ for some tolerance level ϵ_π, update the guesses for the profit functions according to $\hat{\pi}_t^{'} = \tilde{\pi}_t$ for all $t \in \left\{2, \cdots, \hat{T}\right\}$, and repeat from step 2 until $\mathcal{R}_\pi \leq \epsilon$.

5. Compute the residual:

$$\mathcal{R}_m \equiv \max_{\left(\chi, \chi^{'}\right) \in S_\chi^2} \left| \hat{m}_{\hat{T}}\left(\chi, \chi^{'}\right) - m_{ss}\left(\chi, \chi^{'}\right) \right|$$

If $\mathcal{R}_m > \epsilon_m$ for some tolerance level ϵ_m, increment $\hat{T}$ and repeat from step 1.

As in solving for the static market equilibrium, numerical solution of the dynamic market equilibrium requires discretization of the state space S_χ into a mesh grid of $N_{grid} \times N_{grid}$

points, and bilinear interpolation can then be used to obtain numerical approximations of firm-level equilibrium variables off the grid points. Note that given the guess of future profit functions, step 3 of the algorithm has the same computational complexity as solving for the model's steady-state, and this part of the computation can be sped up by using the terminal guesses at the previous date when initializing the guesses for the network characteristic functions in step 3(a). Furthermore, upon increasing the guess for $\hat{T}$ to $\hat{T}+1$ in step 5, the new guess for the profit functions up to date $\hat{T}$ used in step 2 can be set at the previous terminal guesses for the profit functions up to that date, which also speeds up the computation.

With a grid size of $N_{grid} = 20$ and tolerance levels $\epsilon = \epsilon_\pi = \epsilon_m = 10^{-4}$, executing the steady-state algorithm typically takes around 30 seconds, while solving for a transition path such as those discussed in the main text typically takes about one hour on a standard computer. Since estimation of the model's parameters only requires solving for steady-state equilibria, the complexity of executing the dynamic algorithm does not factor into the tractability of estimating the model.

B Efficiency of the market equilbrium

B.1 Static efficiency

To chacaterize the efficiency of the static market equilibrium, I compare the resulting allocation with the allocation that would be chosen by a social planner whose goal is to maximize household welfare subject to the production technology and market clearing constraints. Given the matching function m, the social planner chooses the allocation $\mathcal{A} \equiv \left\{ l(\chi), X(\chi), \{x(\chi,\chi')\}_{\chi' \in S_\chi}, x_H(\chi) \right\}_{\chi \in S_\chi}$ according to :

$$U = \max_{\mathcal{A}} \left[\int_{S_\chi} \left[\delta x_H(\chi)\right]^{\frac{\sigma-1}{\sigma}} dG_\chi(\chi) \right]^{\frac{\sigma}{\sigma-1}}$$

subject to the following constraints:

$$X(\chi) = \left[\left[\phi l(\chi)\right]^{\frac{\sigma-1}{\sigma}} + \int_{S_\chi} m\left(\chi,\chi'\right) \left[\alpha x\left(\chi,\chi'\right)\right]^{\frac{\sigma-1}{\sigma}} dG_\chi\left(\chi'\right) \right]^{\frac{\sigma}{\sigma-1}} \tag{B.1}$$

$$X(\chi) = x_H(\chi) + \int_{S_\chi} m\left(\chi',\chi\right) x\left(\chi',\chi\right) dG_\chi(\chi') \tag{B.2}$$

$$\int_{S_\chi} l(\chi)\, dG_\chi(\chi) = L - L_f \tag{B.3}$$

where L_f is taken as given.

Denoting the Lagrange multipliers on constraints (B.2) and (B.3) by $\left(\frac{U}{\Delta_H}\right)^{\frac{1}{\sigma}} \eta(\chi) G_\chi(\chi)$ and $\left(\frac{U}{\Delta_H}\right)^{\frac{1}{\sigma}}$ respectively, the first-order conditions for the planner's problem can be expressed as:

$$x_H(\chi) = \Delta_H \delta^{\sigma-1} \eta(\chi)^{-\sigma} \tag{B.4}$$

$$l(\chi) = X(\chi)\,\eta(\chi)^{\sigma}\,\phi^{\sigma-1} \tag{B.5}$$

$$x\left(\chi, \chi'\right) = X(\chi)\,\eta(\chi)^{\sigma}\,\alpha^{\sigma-1}\eta\left(\chi'\right)^{-\sigma} \tag{B.6}$$

Substituting these equations into (B.1) and (B.2), one obtains:

$$\Phi(\chi) = \phi^{\sigma-1} + \alpha^{\sigma-1}\int_{S_\chi} m\left(\chi, \chi'\right)\Phi\left(\chi'\right) dG_\chi\left(\chi'\right) \tag{B.7}$$

$$\Delta(\chi) = \delta^{\sigma-1} + \alpha^{\sigma-1}\int_{S_\chi} m\left(\chi', \chi\right)\Delta\left(\chi'\right) dG_\chi\left(\chi'\right) \tag{B.8}$$

where $\Phi(\chi) \equiv \eta(\chi)^{1-\sigma}$and $\Delta(\chi) \equiv \frac{1}{\Delta_H} X(\chi)\,\eta(\chi)^{\sigma}$.

Note that equations (B.4)-(B.8) are identical to equations (2.2), (2.7), (2.8), (2.15), and (2.16) respectively only when $\mu = 1$. This tells us that the static market equilibrium allocation is identical to the planner's allocation if and only if the markups charged by all firms are equal to one. With a finite elasticity of substitution σ, the static market equilibrium is therefore inefficient relative to the planner's allocation because of the monopoly markup distortion.

B.2 Dynamic efficiency

To study the efficiency properties of the dynamic market equilibrium, we consider the problem of a social planner that chooses the set of relationships to activate and terminate at each date so as to maximize the present discounted value of household welfare, subject to the same dynamic frictions faced by firms in the market equilibrium. From the results in section B.1, we know that given the matching function m_t and the total mass of labor used to pay relationship costs $L_{f,t}$, household utility at date t under the planner's optimal allocation can be written as:

$$U_t = (L - L_{f,t})\,\mathcal{C}_t \tag{B.9}$$

where $\mathcal{C}_t$ measures the total connectivity of the static production network:

$$\mathcal{C}_t \equiv \left[\int_{S_\chi}\int_{S_\chi}\left[\sum_{d=0}^{\infty}\alpha^{d(\sigma-1)}m_t^{(d)}\left(\chi,\chi^{'}\right)\right]\left(\delta\phi^{'}\right)^{\sigma-1}dG_\chi\left(\chi\right)dG_\chi\left(\chi^{'}\right)\right]^{\frac{1}{\sigma-1}} \tag{B.10}$$

$$= \left[\int_{S_\chi}\Phi_t\left(\chi\right)\delta^{\sigma-1}dG_\chi\left(\chi\right)\right]^{\frac{1}{\sigma-1}} \tag{B.11}$$

$$= \left[\int_{S_\chi}\Delta_t\left(\chi\right)\phi^{\sigma-1}dG_\chi\left(\chi\right)\right]^{\frac{1}{\sigma-1}} \tag{B.12}$$

and Φ_t and Δ_t are given by the date t equivalents of equations (B.7) and (B.8) respectively.

To study the planner's dynamic optimization problem, let $V_t\left(m_{t-1}\right)$ denote the present value of discounted household utility at date t under the planner's optimal dynamic allocation when the matching function in the previous period is given by m_{t-1}. At each date t, the planner's choice about which relationships to activate and terminate is equivalent to a choice over the values$\{\xi_{max,t}\left(\chi,\chi'\right)\}_{\left(\chi,\chi'\right)\in S_\chi^2}$, where $\xi_{max,t}\left(\chi,\chi'\right)$ specifies the maximum value of the idiosyncratic relationship cost shock component for which $\chi-\chi^{'}$ firm pair relationships are accepted. The Bellman equation for the planner's problem can therefore be written as:

$$V_t\left(m_{t-1}\right) = \max_{\{\xi_{max,t}(\chi,\chi')\}_{\left(\chi,\chi^{'}\right)\in S_\chi^2}}\left[U_t+\beta V_{t+1}\left(m_t\right)\right] \tag{B.13}$$

where the maximization is subject to $\xi_{max,t}\left(\chi,\chi^{'}\right)\geq 0$ for all t and $\left(\chi,\chi^{'}\right)\in S_\chi^2$, as well as the following constraints:

$$U_t = \left(L-L_{f,t}\right)\mathcal{C}_t \tag{B.14}$$

$$\mathcal{C}_t = \left[\int_{S_\chi}\Phi_t\left(\chi\right)\delta^{\sigma-1}dG_\chi\left(\chi\right)\right]^{\frac{1}{\sigma-1}} \tag{B.15}$$

$$\Phi_t\left(\chi\right) = \phi^{\sigma-1}+\alpha^{\sigma-1}\int_{S_\chi}m_t\left(\chi,\chi^{'}\right)\Phi_t\left(\chi^{'}\right)dG_\chi\left(\chi^{'}\right) \tag{B.16}$$

$$L_{f,t} = \psi\int\int_{S_\chi}\left[\nu m_{t-1}\left(\chi,\chi^{'}\right)+\left(1-\nu\right)\int_0^{\xi_{max,t}\left(\chi,\chi^{'}\right)}\xi dG_\xi\left(\xi\right)\right]dG_\chi\left(\chi\right)dG_\chi\left(\chi^{'}\right) \tag{B.17}$$

$$m_t\left(\chi,\chi^{'}\right) = \nu m_{t-1}\left(\chi,\chi^{'}\right)+\left(1-\nu\right)G_\xi\left[\xi_{max,t}\left(\chi,\chi^{'}\right)\right] \tag{B.18}$$

For brevity, denote $\xi^*_{max,t}\equiv\xi_{max,t}\left(\chi^*,\chi^{*'}\right)$ and $m_t^*\equiv m_t\left(\chi^*,\chi^{*'}\right)$ for a given firm pair $\left(\chi^*,\chi^{*'}\right)$. The first step in solving the dynamic planner's problem is to find an expression for the derivative of U_t with respect to $\xi^*_{max,t}$. First, we differentiate (B.17) with respect to

$\xi^*_{max,t}$ to get:

$$\frac{dL_{f,t}}{d\xi^*_{max,t}} = (1-\nu) H\left(\chi^*, \chi^{*'}, \xi^*_{max,t}\right) \psi \xi^*_{max,t} \tag{B.19}$$

where $H\left(\chi, \chi^{'}, \xi\right) \equiv g_\chi(\chi) g_\chi\left(\chi^{'}\right) g_\xi(\xi)$ is the product of three probability densities. Next, differentiating (B.18) for $\left(\chi, \chi^{'}\right) = \left(\chi^*, \chi^{*'}\right)$ with respect to $\xi^*_{max,t}$ gives:

$$\frac{dm^*_t}{d\xi^*_{max,t}} = (1-\nu) G_\xi\left(\xi^*_{max,t}\right) \tag{B.20}$$

Differentiating the functional equation (B.8) with respect to $\xi^*_{max,t}$, we then obtain:

$$\frac{d\Phi_t(\chi)}{d\xi^*_{max,t}} = \frac{d\Phi_t(\chi)}{dm^*_t}\frac{dm^*_t}{d\xi^*_t} \tag{B.21}$$

$$= (1-\nu) G_\xi\left(\xi^*_{max,t}\right) \times \tag{B.22}$$

$$\left[\alpha^{\sigma-1}\Phi_t\left(\chi^{*'}\right)\mathbf{1}_{\chi^*}(\chi) + \alpha^{\sigma-1}\int_{S_\chi} m_t\left(\chi, \chi^{'}\right)\frac{d\Phi_t\left(\chi^{'}\right)}{d\xi^*_{max,t}} dG_\chi\left(\chi^{'}\right)\right] \tag{B.23}$$

$$= (1-\nu) H\left(\chi^*, \chi^{*'}, \xi^*_{max,t}\right)\left[\sum_{d=0}^{\infty}\alpha^{d(\sigma-1)} m_t^{(d)}\left(\chi, \chi^{*}\right)\right]\alpha^{\sigma-1}\Phi\left(\chi^{*'}\right) \tag{B.24}$$

where $\mathbf{1}_{\chi^*}(\chi)$ is the indicator function that equals 1 if $\chi = \chi^*$ and 0 otherwise. (Note that equation (B.24) summarizes the effect of a change in the mass of connections between $\chi^* - \chi^{*'}$ firm pairs on the network productivities of all firms that are downstream of χ^* firms.) Differentiating equation (B.14) with respect to $\xi^*_{max,t}$ and using (B.19) and (B.24), we then get:

$$\frac{dU_t}{d\xi^*_{max,t}} = (1-\nu) H\left(\chi^*, \chi^{*'}, \xi^*_{max,t}\right)\mathcal{C}_t\left[\tilde{\pi}_t\left(\chi^*, \chi^{*'}\right) - \psi\xi^*_{max,t}\right] \tag{B.25}$$

where we have defined:

$$\tilde{\pi}_t\left(\chi^*, \chi^{*'}\right) \equiv \frac{\alpha^{\sigma-1}}{\sigma-1}\Delta_{H,t}\Delta_t(\chi^*)\Phi_t\left(\chi^{*'}\right) \tag{B.26}$$

Note that conditional on the network characteristic functions, $\tilde{\pi}_t$ differs from the profit function π_t in the dynamic market equilibrium (given by equation (3.5)) only by a constant fraction $\mu^{-\sigma}$.

The next step in solving the planner's problem is to derive an expression for the derivative of the continuation value $V_{t+1}(m_t)$ with respect to $\xi^*_{max,t}$. First, we note that:

$$\frac{dV_{t+1}}{d\xi^*_{max,t}} = (1-\nu) G_\xi\left(\xi^*_{max,t}\right)\frac{dV_{t+1}}{dm^*_t} \tag{B.27}$$

The envelope condition then gives us:

$$\frac{dV_{t+1}}{dm_t^*} = \frac{dU_{t+1}}{dm_t^*} + \beta\nu \frac{dV_{t+2}}{dm_{t+1}^*} \tag{B.28}$$

Using the same approach as in solving for $\frac{dU_t}{d\xi^*_{max,t}}$, it is straightforward to show that:

$$\frac{dU_{t+1}}{dm_t^*} = \nu G_\chi\left(\chi^*\right) G_\chi\left(\chi^{*'}\right) \mathcal{C}_{t+1}\left[\tilde{\pi}_{t+1}\left(\chi^*, \chi^{*'}\right) - \psi\right] \tag{B.29}$$

Combining (B.27), (B.28) and (B.29), we then obtain:

$$\frac{dV_{t+1}}{d\xi^*_{max,t}} = \nu\left(1-\nu\right) H\left(\chi^*, \chi^{*'}, \xi^*_{max,t}\right) \sum_{s=0}^{\infty} \left(\beta\nu\right)^s \mathcal{C}_{t+1+s}\left[\tilde{\pi}_{t+1+s}\left(\chi^*, \chi^{*'}\right) - \psi\right] \tag{B.30}$$

Piecing together equations (B.25) and (B.30), we can finally write the first-order condition with respect to $\xi_{max,t}\left(\chi, \chi^{'}\right)$ in the planner's problem as:

$$\xi_{max,t}\left(\chi, \chi^{'}\right) = \max\left\{\frac{\tilde{\pi}_t\left(\chi, \chi^{'}\right)}{\psi} + \sum_{s=1}^{\infty}\left(\beta\nu\right)^s\left(\frac{\mathcal{C}_{t+s}}{\mathcal{C}_t}\right)\left[\frac{\tilde{\pi}_{t+s}\left(\chi, \chi^{'}\right)}{\psi} - 1\right], 0\right\} \tag{B.31}$$

C Estimation Procedure

To estimate the key parameters of the model, a simulated method of moments technique is employed. Targeted moments are computed from both the data and the model, as discussed in section C.1. A pattern search algorithm is then executed to search over the parameter space for the set of parameter values that minimizes a measure of distance between the empirical and simulated moments, as discussed in section C.2.

C.1 Calculation of moments

C.1.1 Firm revenue distribution

In the data, I first compute the log of firm revenue normalized by average revenue. This normalization is employed due to the scale invariance of the model. I then compute the empirical CDF of log normalized revenue, and use linear interpolation to obtain the value of the inverse CDF at the midpoints of N_{bin} evenly-space quantile bins. The same moments are calculated in the model, where firm size is given by equation (2.22). This yields a vector of moments $R^m \equiv \{R_b^m\}_{b=1}^{N_{bin}}$.

C.1.2 Firm degree distributions

In the data, I first normalize the empirical in- and out-degree distributions by their mean. As with the firm size distribution, I then compute the empirical CDF of normalized degree, and use linear interpolation to obtain the value of the inverse CDF at the midpoints of N_{bin} evenly-spaced quantile bins. The same moments are calculated in the model, where the in- and out-degree of a firm are given respectively by:

$$N_{sup}(\chi) = \int_{S_\chi} m\left(\chi, \chi'\right) dG_\chi\left(\chi'\right) \tag{C.1}$$

$$N_{cus}(\chi) = \int_{S_\chi} m\left(\chi', \chi\right) dG_\chi\left(\chi'\right) \tag{C.2}$$

This yields two vectors of moments $N_{sup}^m \equiv \left\{N_{sup,b}^m\right\}_{b=1}^{N_{bin}}$ and $N_{cus}^m \equiv \left\{N_{cus,b}^m\right\}_{b=1}^{N_{bin}}$.

C.1.3 Relationship retention and creation rates

In the data, I first consider the set of firms S_t^{in} with positive in-degree for each year $t \in \{1, \cdots, N_{year} - 1\}$, where N_{year} is the number of years of observations in the Compustat data. Within this sample, I compute the revenue quantile q_{it}^{in} of each firm $i \in S_t^{in}$. In addition, I also compute the fraction of each firm's suppliers $\rho_{sup,it}$ that are retained in year $t+1$. I then construct N_{bin} evenly-spaced quantile bins, and for each bin, I compute the average of the mean supplier retention rate $\rho_{sup,it}$ for all firms that have revenue quantile q_{it}^{in} falling within the bin, pooling the data across years. Analagous moments are computed for the customer retention rates. The same moments are then computed in the model, where the supplier and customer retention rates are given by:

$$\rho_{sup}(\chi) = \frac{\int_{S_\chi} \left[\nu m\left(\chi, \chi'\right) + (1-\nu)\, m\left(\chi, \chi'\right) a\left(\chi, \chi'\right)\right] dG_\chi\left(\chi'\right)}{N_{sup}(\chi)} \tag{C.3}$$

$$\rho_{cus}(\chi) = \frac{\int_{S_\chi} \left[\nu m\left(\chi', \chi\right) + (1-\nu)\, m\left(\chi', \chi\right) a\left(\chi', \chi\right)\right] dG_\chi\left(\chi'\right)}{N_{cus}(\chi)} \tag{C.4}$$

For the relationship creation rates, moments are calculated from the data in the same fashion. In the model, the relationship creation rates are given by:

$$\eta_{sup}(\chi) = \frac{\int_{S_\chi} (1-\nu)\left[1 - m\left(\chi, \chi'\right)\right] a\left(\chi, \chi'\right) dG_\chi\left(\chi'\right)}{N_{sup}(\chi)} \tag{C.5}$$

$$\eta_{cus}(\chi) = \frac{\int_{S_\chi} (1-\nu)\left[1 - m\left(\chi', \chi\right)\right] a\left(\chi', \chi\right) dG_\chi\left(\chi'\right)}{N_{cus}(\chi)} \tag{C.6}$$

This yields four vectors of moments $\rho^m_{sup} \equiv \left\{\rho^m_{sup,b}\right\}_{b=1}^{N_{bin}}$, $\rho^m_{cus} \equiv \left\{\rho^m_{cus,b}\right\}_{b=1}^{N_{bin}}$, $\eta^m_{sup} \equiv \left\{\eta^m_{sup,b}\right\}_{b=1}^{N_{bin}}$, and $\eta^m_{cus} \equiv \left\{\eta^m_{cus,b}\right\}_{b=1}^{N_{bin}}$.

C.2 Optimization algorithm

The procedure described above yields a vector of $N_{bin} \times 7$ moments computed from both the model and the data:

$$\mathscr{M} \equiv \begin{bmatrix} R^m & N^m_{sup} & N^m_{cus} & \rho^m_{sup} & \rho^m_{cus} & \eta^m_{sup} & \eta^m_{cus} \end{bmatrix}^T \tag{C.7}$$

I then compute the distance between the vector of empirical moments $\mathscr{M}_{emp}$ and its simulated counterpart $\mathscr{M}_{sim}$ according to:

$$\mathscr{D} = \left(|\mathscr{M}_{emp} - \mathscr{M}_{siml}|\right)^T \mathscr{W} \left(|\mathscr{M}_{sim} - \mathscr{M}_{sim}|\right) \tag{C.8}$$

The weighting matrix $\mathscr{W}$ is computed as the pseudo-inverse of the covariance matrix of the empirical moment vector, which is estimated by resampling with replacement 500 times from the set of firms for the Capital IQ dataset, and from the set of firm-years for the Compustat dataset.

Starting from an arbitrary initial choice of parameter values, I then execute a pattern search optimization algorithm to search for the set of parameter values that minimize $\mathscr{D}$. Standard errors are computed using a bootstrap procedure, in which I repeat the estimation procedure after replacing $\mathscr{M}_{data}$ by the corresponding moments from a resampling of the original data.