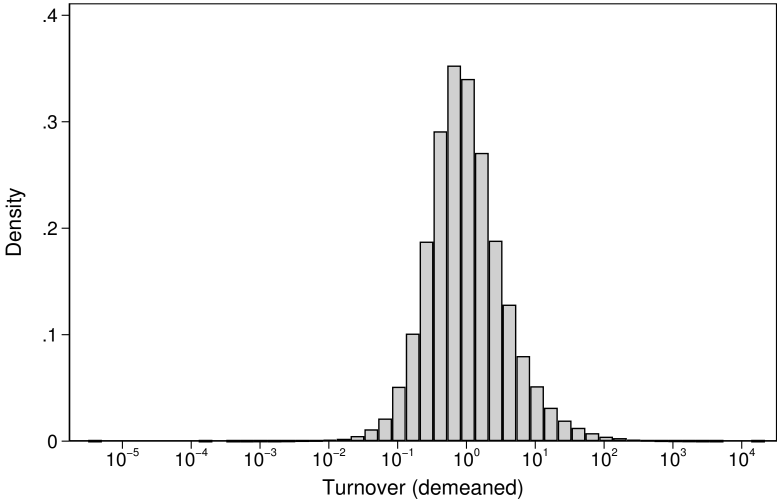
The Origins of Firm Heterogeneity: A Production Network Approach

A. Bernard	E. Dhyne	G. Magerman	K. Manova	A. Moxnes
Tuck@Dartmouth	NBB	ECARES, ULB	UCL	Uni Oslo
CEP, CEPR, NBER	UMons	CEPR, NBB	CEP, CEPR	CEPR

KU Leuven

June 3, 2022

Firm size distribution spans many orders of magnitude



Motivation

- **Why are firms big or small?**
  - random shocks
  - efficiency and demand explanations

- **This paper: firms are part of production networks**
  - buy inputs from firms, sell output to firms and final consumers
  - do these margins matter?

- **Understanding origins of firm heterogeneity is fundamental for**
  - micro: firm survival, innovation, trade participation
  - IO: market power, concentration
  - labor: sorting, skill premia
  - macro: granularity, allocative efficiency

What we do

1. **Present key facts on firm sales in networks**
   - larger firms have more customers, but lower sales per customer
   - seller, buyer and match characteristics matter
   - exact variance decomposition of firm sales to quantify components
   - these findings are inconsistent with canonical models

2. **Develop a model of endogenous production networks**
   - two firm primitives: efficiency and relationship costs
   - efficiency: lower marginal costs and prices
   - relationship costs: higher cost of matching with customers
   - links, input prices and sales all determined in equilibrium

3. **Quantitative predictions using SMM**
   - both primitives are strongly positively correlated
   - data rejects one-dimensional and uncorrelated models
   - counterfactual: higher efficiency gains from lowering matching costs

Related literature

- **Firm size heterogeneity and impact on outcomes**
  - skewness, granularity: Gibrat (1931), Syverson (2011), Gabaix (2011)
  - trade: Bernard et al. (2012), Arkolakis et al. (2012), Melitz & Redding (2015), Gaubert & Itskhoki (2021)
- **Origins of firm heterogeneity**
  - productivity: Jovanovic (1982), Hopenhayn (1992), Sutton (1997), Melitz (2003), Luttmer (2007), Arkolakis (2016), Bloom et al. (2016)
  - organizational capital: Prescott and Visscher (1980), Luttmer (2011)
  - upstream: Manova & Zhang (2011), Antras et al. (2017)
  - final demand: Fitzgerald et al. (2016)
  - firm-specific demand shocks: Foster et al. (2016)
  - supply vs demand: Hottman, Redding & Weinstein (2016)
- **Production networks**
  - Dhyne, Magerman, Rubinova (2015), Eaton et al. (2016), Magerman et al. (2017), Lim (2018), Bernard et al. (2018, 2019), Baqaee and Farhi (2019, 2020), Kikkawa, Magerman and Dhyne (2022)
- **Two-sided heterogeneity**
  - labor: Abowd et al. (1999), Card et al. (2015), Kramarz et al. (2016)
  - identification: Arellano and Bonhomme (2017), Bonhomme et al. (2017), Bonhomme (2020)

Agenda

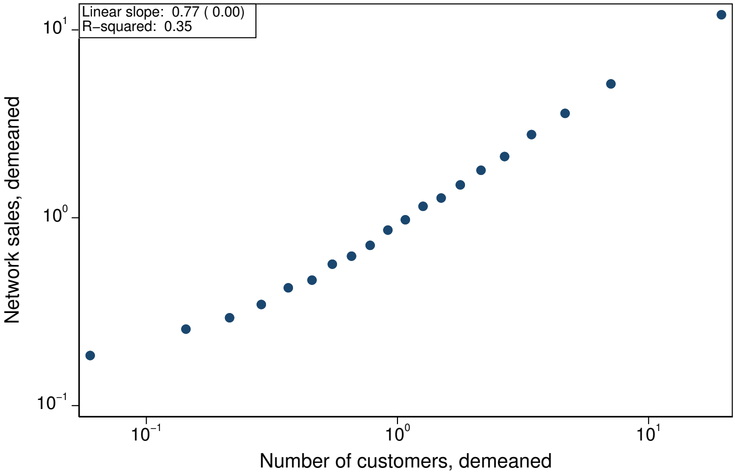
1. Stylized facts
2. Model of endogenous networks
3. Quantitative predictions

Data sources

- **B2B network (Dhyne, Magerman & Rubinova, 2015)**
  - sales value from firm i to j within Belgium
  - all VAT-liable firms in Belgium
  - all annual bilateral sales $\geq$ 250 euro
- **Firm characteristics**
  - annac: sales, input expenditures, employees, labor cost
  - VAT decl: sales and inputs for small firms missing in annac
  - CBE: NACE 4-digit sector, postal code
- **Final sample**
  - panel 2002-2014 (200 mln obs, 17mln in 2014)
  - firms with at least 1 FTE, with all their linkages
  - use 2014 for baseline analysis

1. Firms with more customers have higher sales...

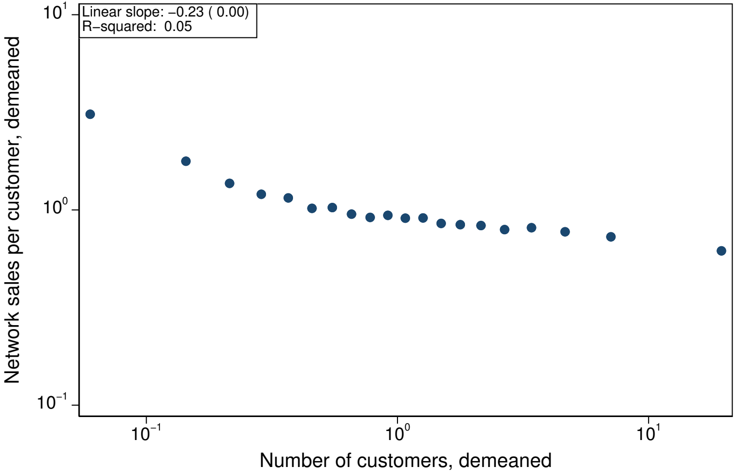
- **Sales ↑ in number of customers (slope: 0.77)**



Notes: Binned scatter plot, 20 quantiles by log number of customers. Each bin represents the mean of the variables on the x- and y-axes. Both axes demeaned by NACE 4-digit sectors. Regression estimated on full data, not bins ($N = 94,147$). Robust standard errors between parentheses.

2. ...but lower sales per customer...

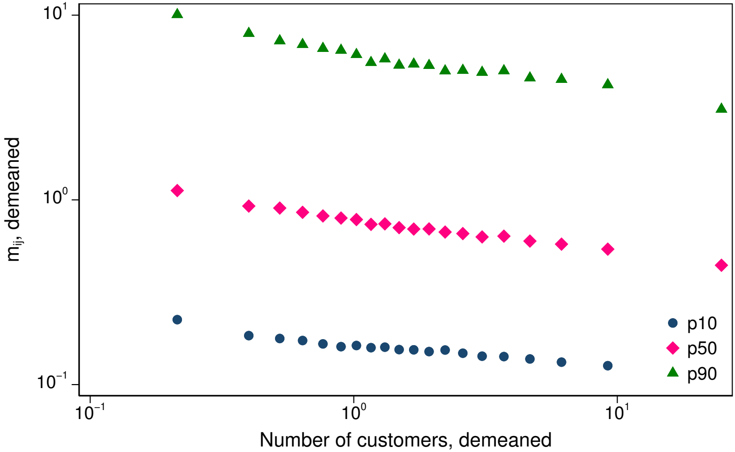
- **Avg. sales per customer ↓ (slope: –0.23); slopes sum to 1**



Notes: Binned scatter plot, 20 quantiles by log number of customers. Each bin represents the mean of the variables on the x- and y-axes. Both axes demeaned by NACE 4-digit sectors. Regression estimated on full data, not bins ($N = 94,147$). Robust standard errors between parentheses.

3. ...and lower sales across the customer distribution...

- **This is not driven by parts of the distribution**

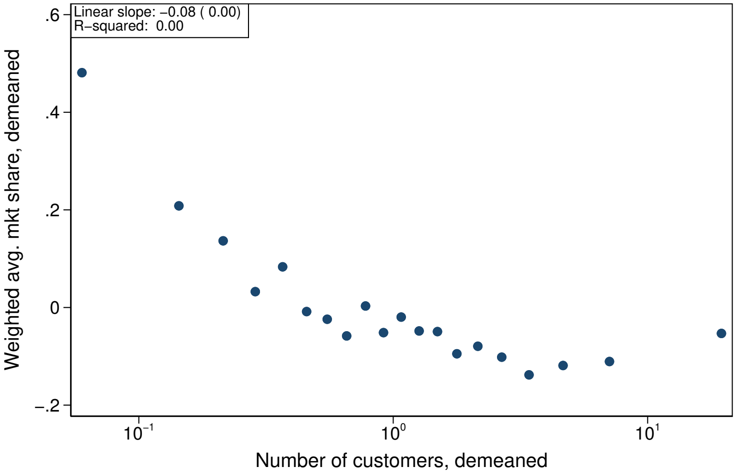


Linear slope: p10: –0.13 (0.00) p50: –0.18 (0.00) p90: –0.22 (0.00)

Notes: p90/p50/p10 refer to the 90*th*/50*th*/10*th* percentile of firm-to-firm sales m_{ij} across buyers j within firm i, demeaned by NACE 4-digit sectors.

4. ...and lower input shares within customers

- **Also not driven by selection on customers**



Notes: Average input share are weighted geometric means of firm i in the total network purchases of its buyers j, using sales shares as weights.

Decomposing bilateral sales

- **Exploit the firm-to-firm data to estimate the contribution of seller, buyer and match effects**

- **Estimate seller/buyer FE from** $\ln m_{ij}$

$$\ln m_{ij} = \ln G + \ln \psi_i + \ln \theta_j + \ln \omega_{ij}$$

  - seller FE $\ln \psi_i$: average sales of i to j, controlling for purchases of j
  - buyer FE $\ln \theta_j$: average purchases of j controlling for sales of i

- **Two-way fixed effects regression**
  - similar to Abowd et al. (1999), but in cross-section
  - estimation on giant connected component (“mobility group”)
  - our advantage: many obs per FE, no switching, no time variation
  - exogenous mobility checks

Two worlds

$$\ln m_{ij} = \ln G + \ln \psi_i + \ln \theta_j + \ln \omega_{ij}$$

- **Variance only in seller effects ψ_i**
  - firm i is bigger since it sells more to every given customer j
  - there is no variation in how much j buys from every i
  - who you are means everything: **sales is driven by sales capability**

- **Variance only in buyers effects θ_j**
  - firm i is bigger since it sells to bigger customers
  - there is no variation in how much i sells to common j
  - who you meet means everything: **sales is driven by matching ability**

- **Without data on firm-to-firm sales, we cannot disentangle these two**

An exact variance decomposition

- **Given estimates for $\Psi = \{\ln\psi_i, \ln\theta_j, \ln\omega_{ij}\}$, decompose total firm sales as**

$$S_i = S_i^{net} s_i / s_i^{net} \equiv S_i^{net}\beta$$

and

$$\begin{aligned} S_i^{net} &\equiv \sum_{j\in\mathcal{C}_i} m_{ij} = \sum_{j\in\mathcal{C}_i} G\psi_i\theta_j\omega_{ij} \\ &= G\ \psi_i\ n_i^c\ \bar{\theta}_i\ \bar{\omega}_i\ \underbrace{\frac{1}{n_i^c}\sum_j \frac{\theta_j}{\bar{\theta}_i}\frac{\omega_{ij}}{\bar{\omega}_i}}_{\equiv\Omega_i^c} \end{aligned}$$

- **So that (identity)**

$$\ln S_i = \ln G + \ln\psi_i + \ln n_i^c + \ln\bar{\theta}_i + \ln\Omega_i^c + \ln\beta_i$$

Total variation in sales decomposed

- **Variance decomposition**
  - demean all variables (NACE 4-digit)
  - regress each component on $\ln S_i$

Table: Variance Decomposition ($\ln S_i$).

Component			**Var. share**	**SE**
Upstream		$\ln \psi_i$	.18	(.00)
Downstream	# Customers	$\ln n_i^c$	.51	(.00)
	Avg Customer Size	$\ln \bar{\theta}_i$	.05	(.00)
	Customer Interaction	$\ln \Omega_i^c$	.25	(.00)
Final Demand		$\ln \beta_i$	.01	(.00)

Notes: Robust standard errors in parentheses.

- **Results**
  - network is key, FD component tiny (1%)
  - size dispersion mostly explained by downstream component (81%)
  - within downstream, number of customers matters most (51%)
  - robustness: by year, by sector

- **Need model on how firms match and why some are good at it**

Agenda

1. Stylized facts
2. Model of endogenous networks
3. Quantitative predictions

Model summary

- **Parsimonious model that can explain**
  - firms with more customers have higher sales...
  - ...but lower sales and market shares per customer
- **Two dimensions of firm heterogeneity**
  - efficiency (z) and bilateral relationship costs (F)
  - possibly $Cov(z, F) \neq 0$
- **Endogenous production network on both int. and ext. margins**
  - sellers match with buyers if profits larger than F
  - given matches, prices and sales are determined across network

Technology and demand

- **Production:** unit continuum of firms; $\sigma > 1$

$$y_i = \kappa z_i l_i^{\alpha} v_i^{1-\alpha}$$
$$v_i = \left(\int_{\mathcal{S}_i} \nu_{k,i}^{(\sigma-1)/\sigma} dk \right)^{\sigma/(\sigma-1)}$$

- **Marginal costs** (wages as numeraire)

$$c_i = \frac{P_i^{1-\alpha}}{z_i} \text{ where } P_i = \left(\int_{\mathcal{S}_i} p_k^{1-\sigma} dk \right)^{\frac{1}{1-\sigma}}$$

- **Final consumers:** inelastic labor supply
  CES, income from labor and profits: $X = wL + \Pi$

Firm-to-firm sales

- **Environment:** monopolistic competition with CES prefs

$$p_i = \underbrace{\frac{\sigma}{\sigma - 1}}_{\equiv \mu} c_i$$

- **Firm-to-firm sales:** conditional on a match

$$\begin{aligned} m_{ij} &= p_i^{1-\sigma} P_j^{\sigma-1} M_j \\ &= \left[\frac{z_i}{\mu P_i^{1-\alpha}} P_j \right]^{\sigma-1} M_j \end{aligned}$$

- All standard, with P_i and P_j endogenous through supplier-buyer linkages

Equilibrium conditional on network

- **Equilibrium is characterized by**
  - firm primitives $\lambda = (z, F)$
  - link function $l(\lambda, \lambda')$: share of seller-buyer pairs that match

- **Backward fixed point**
  Input price depends on $P(\lambda')$ and $z(\lambda')$ of all its suppliers λ'

$$P(\lambda)^{1-\sigma} = \mu^{1-\sigma} \int P(\lambda')^{(1-\sigma)(1-\alpha)} z(\lambda')^{\sigma-1} l(\lambda', \lambda) dG(\lambda')$$

- **Forward fixed point**
  Sales depends on own $z(\lambda)$ and $P(\lambda)$, and customer $S(\lambda')$ and $P(\lambda')$

$$S(\lambda) = \mu^{1-\sigma} z(\lambda)^{\sigma-1} P(\lambda)^{(1-\sigma)(1-\alpha)} \times \left(\frac{X}{\mathcal{P}^{1-\sigma}} + \frac{1-\alpha}{\mu} \int \frac{S(\lambda')}{P(\lambda')^{1-\sigma}} l(\lambda, \lambda') dG(\lambda') \right)$$

Firm-to-firm matching

- **Matching**
  - seller matches with buyer if positive profits: $\pi(\lambda, \lambda') > F(\lambda)\epsilon$
  - $F(\lambda)\epsilon$: bilateral relationship cost (seller pays in units of labor)

- **Linkage fixed point**
  - share of λ firms selling to λ' is

$$l(\lambda, \lambda') = \int \mathbb{I}\left[\ln \epsilon < \ln \pi(\lambda, \lambda') - \ln F\right] dH(\epsilon)$$

  - where $\pi(\lambda, \lambda') = \frac{m(\lambda, \lambda')}{\sigma}$
  - given profits $\pi(\lambda, \lambda')$ determine succesful linkages $l(\lambda, \lambda')$
  - linkages $l(\lambda, \lambda')$ determine $S(\lambda)$ and $P(\lambda)$
  - which in turn determine $\pi(\lambda, \lambda')$ for a given link

General equilibrium

- **Solve 3 fixed points**
- **Algorithm**
  - initial guess of $I(\lambda, \lambda')$
  - solve $P(\lambda)$ and $S(\lambda)$
  - calculate gross profits $\pi(\lambda, \lambda')$ for all matches
  - calculate share of pairs that match $I(\lambda, \lambda')$
  - iterate to convergence

Agenda

1. Stylized facts
2. Model of endogenous networks
3. Quantitative predictions

Estimation of equilibrium

- **Distributional assumptions**
  - $\lambda = (z, F)$ jointly lognormal
  - i.e. $\lambda \sim N(\mu_{\ln z}, \mu_{\ln F}, \Sigma)$ with $\Sigma = \left(\left(\sigma^2_{\ln z}, \rho\right), \left(\rho, \sigma^2_{\ln F}\right)\right)$

- **Parameters to be estimated**

$$\Upsilon = \{\sigma_{\ln z}, \mu_{\ln F}, \sigma_{\ln F}, \rho\}$$

- **Simulated method of moments**

$$\arg\min_{\Upsilon} \left(x - x^s\left(\Upsilon\right)\right)' \left(x - x^s\left(\Upsilon\right)\right)$$

  - choose data moments x, simulate moments from model $x^s(\Upsilon)$
  - generates multivariate distribution of $\lambda = (z, F)$ that best fits the data

Moments

- **Model moments**

$$\Upsilon = \{\sigma_{\ln z}, \mu_{\ln F}, \sigma_{\ln F}, \rho\}$$

Data moment(s)	Identifies
$mean\left(\ln n_i^c\right)$	$\mu_{\ln F}$
$var\left(\ln n_i^c\right),\ var\left(\ln S_i^{net}\right)$	$\sigma_{\ln z},\ \sigma_{\ln F}$
β from $\ln \bar{\delta}_i = \alpha + \beta \ln n_i^c + \varepsilon_i$	ρ
$\ln n_i^c, \ln \bar{\theta}_i$ and $\ln \Omega_i^c$ from decomp	$\sigma_{\ln z},\ \sigma_{\ln F}$

- **Calibrated parameters**

Parameter	Definition	Value	Source
α	Labor cost share	.24	Mean of $\frac{w_i L_i}{w_i L_i + M_i}$
μ	Markup	1.24	Mean of $\frac{S_i}{w_i L_i + M_i}$
X	Aggregate final demand	€470bln	$\sum_i S_i - \sum_i \sum_{j \in \mathcal{C}_i} m_{ij}$
$\ln \epsilon$	Idiosync matching cost	$N\left(0, var\left(\ln z + \ln F\right)\right)$	smoothness obj. funct.
$\sigma_{\ln \epsilon}$	Pair matching cost dispersion	4	smoothness obj. funct.

Estimated model parameters

$\mu_{\ln F}$	$\sigma_{\ln z}$	$\sigma_{\ln F}$	ρ
18.1	0.24	2.23	0.86
(0.02)	(0.00)	(0.01)	(0.00)

Notes: Bootstrapped standard errors in parentheses.

- **Strong positive correlation between z and F**
  - **More efficient firms have on average higher relationship costs F**

Targeted/untargeted moments & other models

- Model matches untargeted moments
- One-dimensional and no-correlation models fail to match negative slope in avg. input shares

	Data	Models			
	(1)	(2) Baseline	(3) No F	(4) No Z	(5) $\rho = 0$
Targeted moments:					
mean $(\ln n_i^c)$	-8.12	-8.12	-8.27	-8.14	-8.12
var $(\ln n_i^c)$	1.87	1.86	0.81	2.20	1.92
var $(\ln S_i^{net})$	3.12	3.12	3.37	2.78	3.08
β from mkt. share	-.11	-.10	1.10	.15	.28
Decomp.: $\ln n_L^c$	.51	.52	.49	.89	.76
Decomp.: $\ln \theta_i$	.05	.01	.01	.03	.02
Decomp.: $\ln \Omega_i^c$	.25	-.04	-.04	-.05	-.06
Non-targeted moments:					
Downstream					
var $(\ln S_i)$	1.73	2.08	1.61	.62	.90
var $(\ln \text{VA/Worker}_i)$	.62	.71	.15	.54	.42
β from p10/p50/p90	-.3/-.3/-.3	-.1/-.2/-.2	1.1/1.1/1	.2/.1/.1	.3/.3/.2
Downstream assort.	-.05	-.04	-.07	-.02	-.03
Upstream					
var $(\ln M_i^{net})$	2.12	2.08	1.61	.62	.90
var $(\ln n_i^s)$	.60	.41	.38	.12	.18
Upstream assort.	-.18	-.18	-.07	-.15	-.15

A counterfactual

- **What is the role of firm heterogeneity on aggregate outcomes?**
  - 50% reduction in relationship costs across the board (lower $\mu_{\ln F}$)
  - calculate change in welfare (real wages)
  - compare baseline and $\rho = 0$ model

- **Results**
  - both models generate more customers and higher welfare
  - baseline model: 17% increase
  - no-correlation model: 12% increase
  - 42% difference in welfare gains across models

Conclusions

- **Big firms have**
  - more customers, but lower sales per customer
  - size variance driven by downstream component
- **Develop model with endogenous production networks**
  - large positive correlation between efficiency and relationship costs
  - other models fail to match the data
- **Underlying mechanisms?**
  - span of control (Lucas 1978, Eeckhout and Kircher 2018)
  - multi-tasking (Holmstrom and Milgrom 1991)

Appendix

Customer heterogeneity

- **Concern**
  - firms with many customers might sell different types of products than firms with few customers, even within seller sector
  - then avg. input shares depend on bilateral characteristics
  - e.g. broad market coffee roasters sell to grocery stores, niche roasters to baristas
- **Exercise**
  - demean all variables by $NACE4_i \times NACE4_j$
  - also accounts for differences in input requirements across sectors
  - then we calculate avg. input share by customer industry k, and number of customers n_{ik}^c
- **Results**
  - estimated slope (SE): - 0.03 (0.00) $\rightarrow$ still negative!

Fringe buyers

- **Concern**
  - sellers with many customers have relatively more fringe buyers
  - i.e. very low m_{ij} to these customers
  - confirms stylized facts, but rejects model, as i sells more to each customer at each customer rank

- **Exercise**
  - drop fringe buyers (1st quartile m_{ij} by i; 1st quartile m_{ij} in globo)
  - recalculate objects and run $\ln S_i^{net} = \beta \ln n_i^c$

- **Results**
  - slopes remain below 1

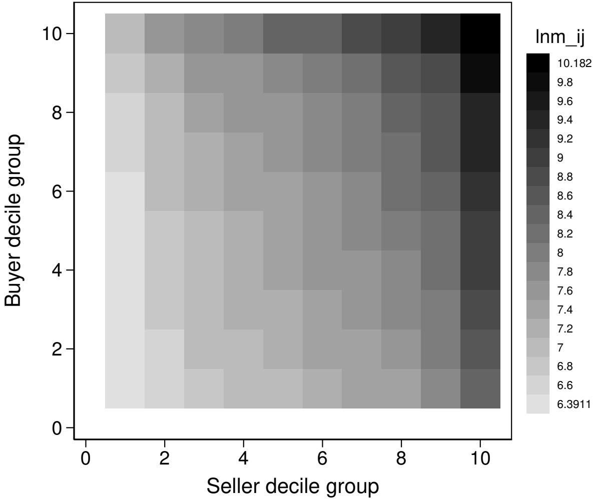
Dep. var: $\ln S_i^{net}$	(1) Base	(2) no fringe (firm-level)	(3) global
$\ln n_i^c$	$.77^{***}$	$.81^{***}$	$.91^{***}$
	(.00)	(.01)	(.00)
$NACE4_i$ FE	Yes	Yes	Yes
N	94,147	80,224	80,156

Buyer and seller FEs

- **The log-linear relationship** $\ln m_{ij} = \ln \psi_i + \ln \theta_j + \ln \omega_{ij}$ **predicts the following**
  - expected sales from i to j increase in avg. sales of i to other customers k
  - expected purchases by j from i increase in avg. purchases by j of other suppliers k

- **Non-parametric test**
  - 1. calculate leave-out means of log sales of i and purchases by j
  - 2. sort firms into decile groups
  - 3. calculate avg. $\ln m_{ij}$ for each decile pair
  - 4. is avg. $\ln m_{ij}$ increasing in these pairs?

Buyer and seller FEs



Note: The figure shows the average of $\ln m_{ij}$ in all decile group pairs $(q_{\bar{s}}, q_{\bar{m}})$.

Conditional exogenous mobility

- **Threats to identification**
  - estimating $\ln m_{ij} = \ln \psi_i + \ln \theta_j + \ln \omega_{ij}$ with OLS?
  - consistent if following moment conditions hold:

$$\begin{cases} \mathbb{E}(s_i' r) = 0 & \forall i \\ \mathbb{E}(b_j' r) = 0 & \forall j \end{cases}$$

    i.e. for all i, $\mathbb{E}(\ln \omega_{ij}) = 0$ across all j, and for all j $\mathbb{E}(\ln \omega_{ij}) = 0$ across all i.
  - i.e. that the assignment of suppliers to customers is exogenous with respect to $\ln \omega_{ij}$ (conditional exo. mobility)

- **Some scenarios where exo mobility holds**
  - matching on $\ln \psi_i$ and $\ln \theta_j$
  - matching on pairwise idiosync shocks, unrelated to $\ln \omega_{ij}$ (e.g.idio fixed costs for search and matching)

- **When does it fail?**
  - say matching on $\ln \psi_i$ and $\ln \omega_{ij}$
  - if $cov(\ln \psi_i, \ln \omega_{ij}) < 0$, OLS estimates are downward biased (since $\mathbb{E}(\ln \omega_{ij}) = 0$ imposed)

Conditional exogenous mobility

- **Can we evaluate endogenous mobility?**
  - consider firm i selling to customers 1 and 2
  - expected difference in bilateral sales is

$$\Delta \ln m_i \equiv E\left[\ln m_{i2} - \ln m_{i1} \mid (i,1),(i,2)\right] = \ln\theta_2 - \ln\theta_1 + E\left[\ln\omega_{i2} - \ln\right.$$

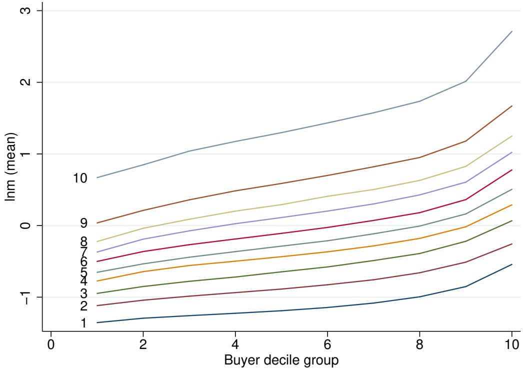
  - consider the case $\theta_2 > \theta_1$

- **Exo mobility**
  - $E\left[\ln\omega_{i2} - \ln\omega_{i1} \mid (i,1),(i,2)\right] = 0$
  - $\Delta \ln m_i$ is unrelated to firm i characteristics

- **Endo mobility**
  - say i would only want to match with customer 1 if $\ln\omega_{i1}$ is large enough
  - then $E\left[\ln\omega_{i2} - \ln\omega_{i1} \mid (i,1),(i,2)\right] < 0$
  - matching is determined by both ψ_i and ω_{ij}
  - for firms with low $\ln\psi_i$, contribution of $\ln\omega_{ij}$ is small to match
  - for high $\ln\psi_i$, contribution $\ln\omega_{ij}$ matters a lot
  - under endo mobility $E\left[\ln\omega_{i2} - \ln\omega_{i1} \mid (i,1),(i,2)\right]$ is less negative for high-ψ_i than low-ψ_i firms

Conditional exogenous mobility

- Moving from a small to a big customer, across seller groups, how does the average $\ln m_{ij}$ change?
- Each line shows avg. $\ln m_{ij}$ for a given seller decile, and its change across larger buyer deciles
- under exo mobility, these lines should be parallel (independent of seller/buyer characteristics)
- sufficient but not necessary condition: if DGP is not linear in logs, can have non-parallel lines even under exo mobility



Two-way fixed effects regression

$$\ln m_{ij} = \ln G + \ln \psi_i + \ln \theta_j + \ln \omega_{ij}$$

Table: Seller and Buyer Effects

	N	$\frac{var(\ln\psi_i)}{var(\ln\psi_i+\ln\theta_j)}$	$\frac{var(\ln\theta_j)}{var(\ln\psi_i+\ln\theta_j)}$	$\frac{2cov(\ln\psi_i,\ln\theta_j)}{var(\ln\psi_i+\ln\theta_j)}$	R^2	Adjusted R^2
$\ln m_{ij}$	17,054,274	0.66	0.32	0.02	0.43	0.39

Notes: The table reports the (co)variances of the estimated seller and buyer fixed effects. The estimation is based on the high-dimensional fixed effects estimator from Correia (2016).

Correlations components

- **Alternative to variance decomposition**

Table: Correlation Matrix

Firm Size Component	$\ln S_i$	$\ln \psi_i$	$\ln n_i^c$	$\ln \bar{\theta}_i$	$\ln \Omega_i^c$	$\ln \beta_i$
$\ln S_i$	1					
$\ln \psi_i$	0.23	1				
$\ln n_i^c$	0.49	-0.33	1			
$\ln \bar{\theta}_i$	0.20	0.22	-0.18	1		
$\ln \Omega_i^c$	0.45	0.16	0.09	0.23	1	
$\ln \beta_i$	0.02	-0.36	-0.33	-0.16	-0.42	1

Note: All correlations are significant at 5%. All variables are demeaned at the NACE 4-digit level.

- sales correlate positively with the components, little with final demand
- n_i^c and $\ln \psi_i$ are negatively correlated (firms with many customers have smaller input shares in these customers)

Decomposition by sector

- **Concern**
  - results might be driven by sector compositional effects (e.g. retail vs manufacturing)
- **Exercise**
  - perform decomposition by 2-digit and 4-digit sectors separately
- **Results**
  - summary statistics across 2-digit sectors
  - little variation in components across sectors, mostly β (confirms intuition)

NACE Sector	$\ln \psi_i$	$\ln n_i^c$	$\ln \bar{\theta}_i$	$\ln \Omega_i^c$	$\ln \beta_i$
mean	0.17	0.43	0.06	0.26	0.07
st. dev.	0.14	0.24	0.05	0.09	0.20
CV	0.81	0.55	0.82	0.34	2.70

Decomposition by year

- **Concern**
  - results might be varying across years
- **Exercise**
  - perform decomposition for other years
- **Results**
  - variance shares by year
  - little variation in components across years $\rightarrow$ underlying mechanism

Year	N	$\ln\psi_i$	$\ln n_i^c$	$\ln\bar{\theta}_i$	$\ln\Omega_i^c$	$\ln\beta_i$
2002	81,254	0.17	0.49	0.04	0.25	0.05
2003	83,678	0.17	0.49	0.04	0.25	0.05
2004	85,030	0.18	0.49	0.04	0.25	0.04
2005	86,474	0.17	0.49	0.04	0.25	0.04
2006	88,581	0.17	0.50	0.04	0.25	0.04
2007	91,027	0.18	0.50	0.04	0.25	0.03
2008	92,280	0.18	0.50	0.04	0.25	0.03
2009	92,333	0.17	0.50	0.04	0.25	0.04
2010	92,713	0.17	0.50	0.04	0.25	0.03
2011	94,093	0.18	0.50	0.05	0.25	0.03
2012	95,375	0.18	0.50	0.05	0.25	0.03
2013	94,135	0.18	0.50	0.05	0.25	0.02
2014	94,147	0.18	0.51	0.05	0.25	0.01

Agglomeration effects

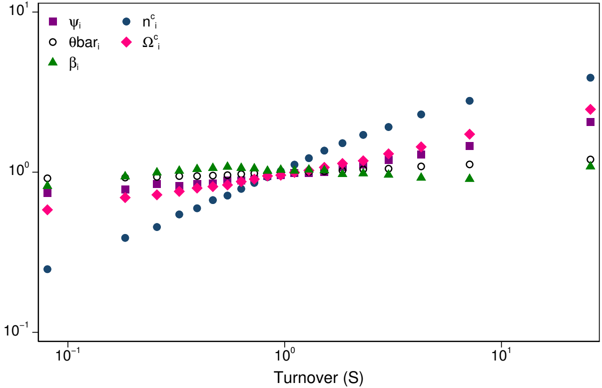
- **Concern**
  - large firms might be located in large cities, where there are more potential customers and suppliers
  - these firms might then have both more customers and lower market shares for reasons unrelated to our model
- **Exercise**
  - perform firm size decomposition, controlling for seller location
- **Results**
  - Components very stable after accounting for geography

	N	$\ln\psi_i$	$\ln n_i^c$	$\ln\bar{\theta}_i$	$\ln\Omega_i^c$	$\ln\beta_i$
$NACE + NUTS3$	94,147	0.17^{***}	0.51^{***}	0.05^{***}	0.25^{***}	0.02^{***}
$NACE \times NUTS3$	80,328	0.16^{***}	0.54^{***}	0.05^{***}	0.25^{***}	0.01^{***}

Note: Significance: * < 5%, ** < 1%, *** <0.1%.

Non-parametric results

- **Concern**
  - are the variance shares stable across the firm size distribution?
- **Exercise**
  - non-parametric visualization of the conditional expectation function
  - sum of components on y-axis = value $x-$axis
- **Results**
  - OLS decomposition fits both small and large firms



Bootstrapping SMM

- **Procedure**
  - draw random sample with replacement until same sample size as main dataset
  - for each sample, create empirical moments from data
  - create 200 bootstrap samples
  - estimate SMM 200 times to generate standard errors

Sensitiviy analysis (Andrews et al., 2017)

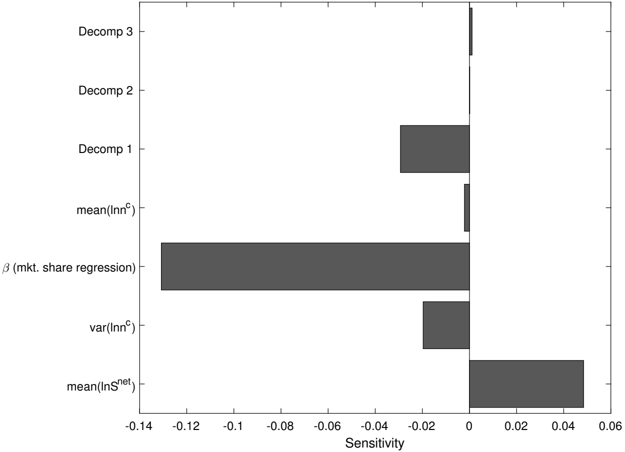
- **Concern**
  - intuitively, $\rho \to 1$ relies heavily on the data correlations
  - estimate of key parameter ρ might be sensitive to perturbations in data moments
- **Exercise**
  - consider perturbations that are additive shifts of moment functions
  - due to either misspecification of $x^s(\Upsilon)$, or measurement error in empirical moments x
  - Andrews et al. (2017) show that sensitivity can be summarized by the matrix

    $$\Lambda = (S'WS)^{-1}S'W$$

    S is the matrix of partial derivatives of $x^s(\Upsilon)$ evaluated at Υ_0
    W is the methods of moments weighting matrix, with is I in our case
  - this allows the reader to further evaluate changes in outcomes from changes in underlying modelling assumptions

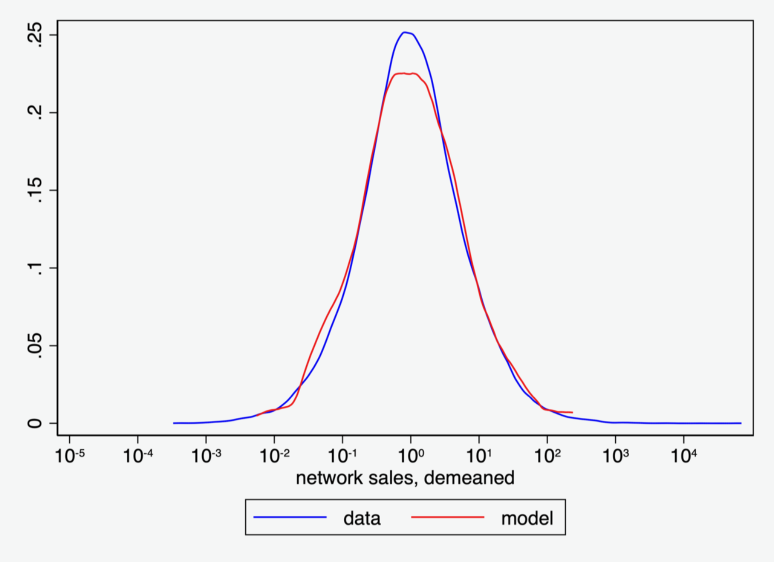
Sensitiviy analysis (Andrews et al., 2017)

- **Results**
  - figure plots the column of Λ corresponding to the estimate of ρ
  - sensitivity as the effect of a 1 stdev change in the moment, on ρ
  - confirms our intuition: the slope coefficient β from the avg. mkt share regression matters most
  - a steeper slope (higher β) has a negative impact on ρ



Model prediction on full distribution

- **The model predicts well the full distribution of firm sizes**



One-dimensional models

- **Fail to generate negative slope in avg. input shares**