Community Networks, Entrepreneurship and the Process of Economic Development

Ruochen Dai
Central University of Finance and Economics

Dilip Mookherjee
Boston University

Kaivan Munshi
Yale University and Toulouse School of Economics

Xiaobo Zhang
Peking University and IPFRI

CESI, Aug 2024

Introduction

- The process of economic development is often characterized by an initial transition from agriculture to domestic production (industry), followed by a second transition to higher value exporting
- These two transitions have been described in the literature as decisions of entrepreneurship: farmers make the occupational decision about starting firms in industry (Lewis, 1954; Banerjee and Newman, 1993); and domestic companies make the exporting decision (Melitz, 2003)
- The conventional individual-specific view is that entrepreneurship is determined by talent (Murphy, Shleifer and Vishny, 1991), education (Levine and Rubinstein, 2017), and inherited wealth when credit is constrained (Banerjee and Newman, 1993)
- These factors have been seen to be relevant in the transition from agriculture to domestic production, as well as in the subsequent shift to exporting (Melitz, 2003; Atkin and Khandelwal, 2020)

Introduction

- Adding a new dimension to the analysis of entrepreneurship, this research documents the important role played by community networks, which emerge in response to market imperfections, at early stages of economic development in China.
- This research proves that the domestic community network can facilitate the entry of domestic producers in the first stage, and the same is true for export community network and the entry of exporters in the second stage.
- However, the overall effect of community network turns out to be more nuanced in the second stage, as our analysis indicates that the incumbent domestic networks created a disincentive to enter exporting that dominated the positive effect of newly emerging export networks
- This tension highlights the complex dynamics of the development process in economies where community networks, supporting occupational mobility at different stages, are active.

China's Economic Transitions

- China's first transition commenced in the 1980's with the establishment of TVE's and then accelerated with the entry of private firms in the 1990's
  - Starting with almost no private firms in 1990, there were 10 million registered private firms in 2012, accounting for 94% of all registered firms. Among them, 55% are established by individual entrepreneurs that are born in rural counties
- A decade after privatization commenced, China entered the WTO in 2001 and soon became the largest exporter in the world (Brandt et al., 2017)
- Previous studies have focused on the reallocation of resources in the manufacturing sector, between state and non-state sectors, in the perspects of firm entry and exit, and across different types of entrepreneurs, to explain China's rapid growth in the two transitions (Hsieh and Klenow, 2009; Brandt and Zhu, 2010; Song, Storesletten and Zilibotti, 2011; Brandt, Van Biesbroeck and Zhang, 2012; Brandt and Lim, 2024)
  - This research contributes to the literature by explaining how millions of rural-born entrepreneurs were able to acquire know-how and establish their businesses during the early stages of economic reforms when many markets were missing or incomplete

The origins of Entrepreneurship in China

- The central idea is that business networks are active in China and that they are organized around the birth county, which is not new. The famous chambers of commerce, such as the 'Jin Merchants' and 'Hui Merchants', are named after their place of origin.
  - A longstanding literature describes how firms respond to the difficulty in enforcing formal contracts in developing economies by establishing relational contracts (McMillan and Woodruff, 1999; Macchiavello and Morjaria, 2015, 2021).
  - Informal arrangements that are based on community, providing different forms of support, must have been at work in a developing economy characterized by weak market institutions and property rights (Peng, 2004; Greif and Tabellini, 2017)
  - The community is defined by the native place in China (Honig, 1992, 1996; Goodman, 1995)

The origins of Entrepreneurship in China

- Building on this past research, we posit that birth county networks allow firms to share inputs and information
  - Long-term relationship (LTR) is one key governance form to solve contracting problem and is costly to maintain (Macchiavello, 2022). Community networks can expand the scope of such bilateral arrangements: a firm in the LTR can provide a (credible) referral for another firm from its network who only requires that connection temporarily (Greif, 1993, 1994).
  - Members of a network can also provide information about new technologies and business opportunities to each other.
  - The existence of information frictions and the help of social networks are also true for exporting activities (Fernandes and Tang, 2014; Atkin et al., 2017)
- The same motivation for cross-firm spillovers has been proposed in the agglomeration literature (Combes et al., 2012; Duranton and Puga, 2020; Rosenthal and Strange, 2020)
  - We focus on a restricted set of firms, within which social ties can be used to increase cooperation, while allowing for conventional agglomeration effects
  - The productivity enhancing mutual help that members of a network provide to each other is inherently local, and we specify the domain of the network by the birth county-destination prefecture

Entrepreneurship in China: Data

- The SAIC registration database covers the universe of registered firms in China from the 1980's onward, and recorded their establishment date, industry, and location.
- Besides, these administrative data provide a list of key personnel in each firm, with their citizenship ID, which can be used to recover the county of birth
- Among these individuals, we designate the principal or legal representative as the "entrepreneur"
- There were 2000 rural counties in China, accounting for 74% of its population, when the economic reforms commenced and so this is an important group to study
  - We posit that networks organized around the hometown (birth county) played an instrumental role in supporting the business of rural-born entrepreneurs
  - The data allows us to track the evolution of all business networks at the birth county-destination level from the starting point
  - Our estimates indicate that birth county networks were active and important; in their absence, the number of rural-born entrepreneurs would have been 23% lower in 2012, the end point of our analysis

The Second Transition: Data

- We focus on relatively productive exporting firms who ship their products directly to foreign buyers.
- Merging the Customs database with the SAIC registration database, we also constructed the export networks at the birth county-destination level since 2000.
  - We find that export networks were still active.
  - 99% of the export network of a birth county-destination have previously established domestic network, and we finds that a large domestic networks discouraged entrepreneurs from moving into the new activity, by increasing the profitability of serving the domestic market.
  - Our estimates indicate that the number of rural-born exporters would have been 76% higher in the absence of the birth county networks in 2012

Relation to the Literature

- This research speaks to two influential literatures in economics:
  1. A literature going back to Galor and Zeira (1993), Banerjee and Newman (1993), studies how market imperfections constrain occupational mobility in developing economies, resulting in the persistence of inequality
     - Our analysis indicates that community-based networks can break these occupational traps
  2. A literature starting with Acemoglu, Johnson and Robinson (2001) documents the long shadow of historical institutions on the trajectory of economic development
     - Our analysis examines the informal institutions that emerge endogenously at early stages of economic development in response to market imperfections, showing that they also have important consequences for an economy's subsequent trajectory

Steps in the Analysis

- Previous research in developing and advanced economies has shown that community-based networks that have been in place for many generations can restrict the mobility of their members when economies restructure (Munshi, 2014)
  - The two unique administrative databases, coupled with the compressed nature of the Chinese development experience, allow us to document the positive and the negative role played by the same (domestic) network at different stages of the development process
- Our analysis proceeds in the following steps
  1. Outline the key features of community business networks in China and the importance of birth county ties (formal network)
  2. Develop a model that explains why networks can both support and dampen mobility
  3. Identify the network effects and overhang effects, and test the model
  4. Quantify the network effects and overhang effects

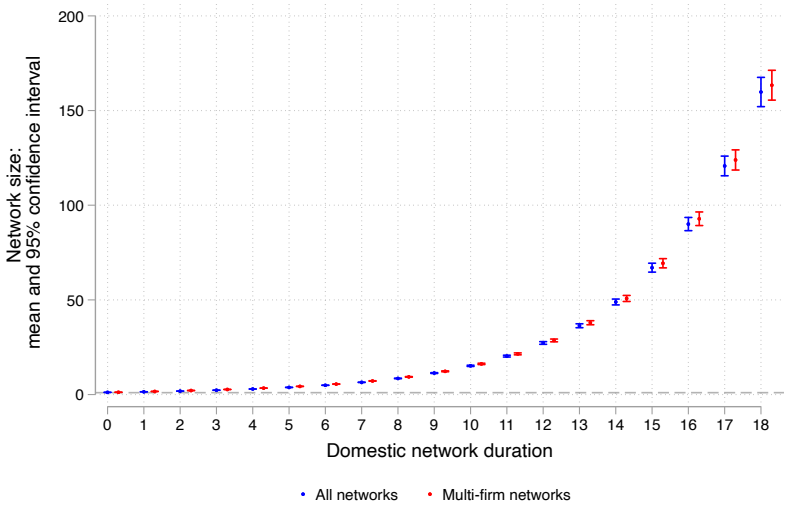
Descriptive Evidence

Domain of the Network

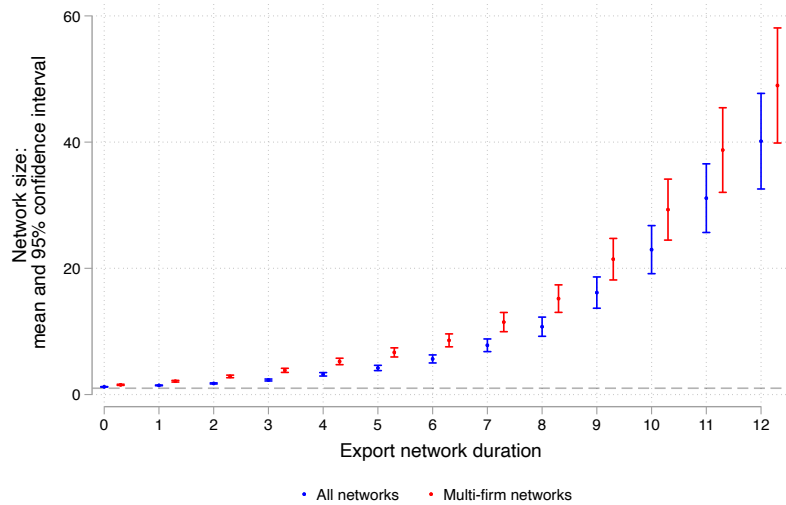
	Domestic network	Export network
Panel A: birth county-destination prefecture level network		
total number	140,924	7,874
fraction of networks that are multi-firm	71.5%	40.7%
fraction of firms in multi-firm networks	99.3%	90.7%
fraction of multi-firm networks that start with 1 firm	84.9%	78.5%
Panel B: characteristics of multi-firm networks in 2012		
average destination per community	63.6	2.7
fraction of firms in networks outside the birth county	66.2%	46.3%
average network size	49.0	12.0
average duration after the initial year	9.9	7.5
average duration after the initial year of domestic network	–	12.5

- More than 80% domestic firms and 70% exporters are in networks that started from a single firm and grew over time.
  - It echoes anecdotal evidence that there is typically an accidental aspect to business network formation (Munshi, 2014; Kerr and Mandorf, 2023)

Trajectory of the Network



(a) Domestic Network



(b) Exporter Network

Source: Registration database and Customs database

- Based on the administrative data, we observe the precise point in time at which each of the domestic networks and export networks in our data commenced, which, in turn, allows us to compute their durations in any subsequent year.
- Both domestic and export networks grew exponentially on the duration.
-

Birth County Homophily

- We provide descriptive support for the importance of birth county ties. While networks largely rely on informal interactions between socially connected firms, formal links can also be used to complement these interactions and increase cooperation.
- We restrict firms that are established outside the birth county, to avoid the natural homophily led by home effects.

Variable:	fraction of key personnel from the entrepreneur's birth county		fraction of linked firms that are linked to a firm from the same birth county	
Network:	domestic network	export network	domestic network	export network
	(1)	(2)	(3)	(4)
Mean	0.487	0.445	0.512	0.397
Counter-factual mean	0.012	0.055	0.014	0.208

Note: The counter-factual mean is based on the random assignment of key personnel and the random matching of linked firms in the prefectures where they are located.

- A more stringent test that networks are active is that these informal arrangements should improve the performance of their members (Munshi, 2003)
  - The model described next will allow us to derive such tests

The Model

Population and Technology

- For expositional convenience, the analytical model is based on a single (rural) birth county and a single destination prefecture where businesses are established
- Successive cohorts of agents, $t' = 1, \ldots, T$ of mass s
- Individual ability ω, with $\log \omega \sim U[A-1, A]$ in each cohort
- Cohort t' agents choose occupations at each date $t \geq t'$:
  - There are two occupations: a traditional occupation and entrepreneurship
  - Payoff in the traditional occupation: ω^{σ}, $\sigma \in (0, 1)$

Payoff from Entrepreneurship

- If an agent chooses to become an entrepreneur, he can produce for the domestic (d) market, the export (e) market, or both
  - Serving a market $b = d, e$ requires a capital investment K_{bt}, which is irreversible

$$R_{dt} = C_{dt}\omega^{1-\alpha}K_{dt}^{\alpha}, R_{et} = C_{et}\omega^{\delta(1-\alpha)}K_{et}^{\alpha}$$

  - $\alpha \in (0,1)$, $\delta > 1$, and C_{bt} is a productivity multiplier, described below

$$E_{dt} = rK_{dt}, E_{et} = r(1+I)K_{et}$$

  - r includes interest and material costs, and $I > 0$ is the incremental cost of operating an export plant
- The total cost of a *mixed* exporter equals $E_{dt} + E_{et} + \beta$, where β measures the diseconomies of scope that are needed to explain the presence of *pure* exporters
- Pure exporters specialize in that activity and are key to explaining why the domestic network can discourage entry into exporting
  - Pure exporting firms have been observed in many developing countries; e.g. Lu et al. (2014), Blum et al. (2020)
  - Based on data from the 2004 and 2008 economic censuses, 15% of exporting firms in China are pure exporters

Productivity Multiplier

- The productivity multiplier, C_{bt}, is determined by exogenous market-time effects, Q_{bt}, and the endogenously determined birth county network

$$C_{dt} = Q_{dt} \cdot [n_{t-1}]^{\theta_d}, C_{et} = Q_{et} \cdot [n_{e,t-1}]^{\theta_e}$$

- Q_{bt} incorporates conventional agglomeration effects and other exogenous business opportunities associated with product demand, government support and infrastructure that apply equally to firms from the different origins that are active in the prefecture
  - $q_{bt} \equiv \log Q_{bt}$ is growing over time
- The network term reflects the idea that mutual help is complementary and so larger networks are more effective at improving the outcomes of their members
  - We assume for the analysis in this section that exogenous entry shocks bring n_0 firms, n_{e0} of whom are exporters, into the market in period 0
  - Given the irreversibility of market entry decisions, network sizes cannot shrink

Occupational Choice in Equilibrium

- Assume that all agents are myopic (this can be relaxed) and that network sizes at past dates are observable by all agents
- Current plant sizes K_{bt} are then chosen to maximize:

$$[C_{dt}\omega^{1-\alpha}K_{dt}^{\alpha} - rK_{dt}] + [C_{et}\omega^{\delta(1-\alpha)}K_{et}^{\alpha} - r(1+I)K_{et}] - \beta\mathbb{I}(K_{dt}K_{et})$$

  - Subject to the irreversibility constraints
- Substituting the optimal decision back in the profit function:

$$\begin{aligned}
\Pi_{Tt}(\omega) &= \omega^{\sigma} \\
\Pi_{Dt}(\omega) &= \omega[\frac{1}{\zeta(r)}]C_{dt}^{\frac{1}{1-\alpha}} \\
\Pi_{Et}(\omega) &= \omega^{\delta}[\frac{1}{\zeta(r)\gamma(I)}]C_{et}^{\frac{1}{1-\alpha}} \\
\Pi_{Mt}(\omega) &= \Pi_{Dt}(\omega) + \Pi_{Et}(\omega) - \beta
\end{aligned}$$

  - Returns to ability and costs are increasing as we move up the occupational ladder, which implies positive selection on ability

Occupational Choice in Equilibrium

- Given suitable parameter restrictions, equilibrium involves sorting by ability thresholds for cohort t' agents at date $t \geq t'$:

$$A - 1 < \log \omega_{dt}^* < \log \omega_{et'}^* < \log \omega_{mt}^* < A$$

  - individuals below ω_{dt}^* stay in the traditional occupation
  - those between ω_{dt}^* and $\omega_{et'}^*$ specialize in domestic production
  - those between $\omega_{et'}^*$ and ω_{mt}^* specialize in exporting
  - and those above ω_{mt}^* become mixed exporters

- The condition $\log \omega_{et'}^* < \log \omega_{mt}^*$ for all t maintains the ordering of thresholds and ensures that some pure exporters in each cohort stay that way, which implies that the additional profit from operating a domestic plant never exceeds β for them
- This implies that domestic producers from cohort t' with ability less than $\omega_{et'}^*$ never transition to (mixed) exporting and, as a result, the export propensity of a given cohort does not change over time
- In contrast, the domestic production threshold ω_{dt}^* and the mixed exporting threshold ω_{mt}^* are independent of the cohort and decreasing over time

Entry into Business

- Individuals with ability $\omega \in [\omega^*_{dt}, A]$ become entrepreneurs
- Solving for the ability threshold and unpacking C_{dt}

$$n_t = ts[A - \omega^*_{dt}] = ts[A - \frac{\log \zeta}{1-\sigma} + \frac{q_{dt} + \theta_d \log n_{t-1}}{(1-\sigma)(1-\alpha)}]$$

  - The initial, exogenous, entry of firms in period 0, n_0, generates subsequent entry through the dynamic network multiplier effect
  - Solving the preceding equation recursively, the number of firms is increasing over time, independently of q_{dt}, when networks are active
  - The empirical analysis will identify and quantify this positive (domestic) network effect on the number of firms

Entry into Exporting

- Individuals from cohort t' with ability $\omega \in [\omega^*_{et'}, A]$ become exporters
  - There is no further entry into exporting from the t' cohort after that period
  - The stock of exporters at any period t is just the sum of exporters supplied by all preceding cohorts (pinned down by the intermediate ability thresholds)
- Solving for these thresholds and unpacking $C_{dt'}$, $C_{et'}$

$$n_{et} = ts[A - \frac{\log \gamma}{\delta - 1}] + \frac{s}{(\delta - 1)(1 - \alpha)} \sum_{t'=1}^{t} [q_{et'} - q_{dt'} + \theta_e \log n_{e,t'-1} - \theta_d \log n_{t'-1}]$$

  - The marginal (pure) exporter's ability is decreasing in the export network term and increasing in the domestic network term
  - The (domestic) network "overhang," which we identify and quantify in the empirical analysis, arises because the marginal exporter is a pure exporter who must choose between domestic production and exporting

Identifying Network Effects and Overhang Effects

Estimating Equations

- The equations that we estimate are derived directly from the analytical model, by extending it in the following ways
  1. We allow for multiple birth counties and multiple destination prefectures, indexed by j and k, respectively
  2. For a given birth county-destination prefecture, we allow the domestic network to start exogenously at time t_{djk} and for the export network to start exogenously at a later time t_{ejk}
  3. We add a stochastic term to the payoff in the traditional occupation, which is now specified as $U_{jkt}\omega^{\sigma}$; denote $u_{jkt} \equiv \log U_{jkt}$
  4. We add a birth county-destination prefecture specific term to the productivity multiplier, to further discuss the exogeneity condition.
     - $C_{djkt} = V_{djkt} Q_{dkt} [n_{jk,t-1}]^{\theta_d}$; denote $v_{djkt} \equiv \log V_{djkt}$
     - $C_{ejkt} = V_{ejkt} Q_{ekt} [n_{ejk,t-1}]^{\theta_e}$; denote $v_{ejkt} \equiv \log V_{ejkt}$

Estimating Equations: Revenue

- Domestic revenue equation:

$$\log R_{djkt} = \frac{\alpha}{1-\alpha}\log\left(\frac{\alpha}{r}\right) + \frac{q_{dkt}}{1-\alpha} + \frac{\theta_d \log n_{jk,t-1}}{1-\alpha} + \frac{[(1-\alpha)^2+1]}{1-\alpha}\log\omega + \frac{v_{djkt}}{1-\alpha}$$

- Export revenue equation:

$$\log R_{ejkt} = \frac{\alpha}{1-\alpha}\log\left(\frac{\alpha}{r(1+l)}\right) + \frac{q_{ekt}}{1-\alpha} + \frac{\theta_e \log n_{ejk,t-1}}{1-\alpha} + \frac{\delta[(1-\alpha)^2+1]}{1-\alpha}\log\omega + \frac{v_{ejkt}}{1-\alpha}$$

- Since we have rich firm-level panel data, we first-difference the revenue equations to purge firm fixed effect for $\log\omega$, and control destination-year fixed effects for $q_{dkt} - q_{dkt-1}, q_{ekt} - q_{ekt-1}$ in estimation.
- The endogeneity problem of $\log n_{jkt-1} - \log n_{jkt-2}$ and the corresponding IV construction can be directly illustrated in the propensity equation.

Estimating Equations: Propensity

- Propensity equations of domestic firms:

$$\frac{n_{jkt}}{s_{jt}} = A_{jk} + \frac{q_{dkt}}{(1-\sigma)(1-\alpha)} + \theta_d \frac{\log n_{jk,t-1}}{(1-\sigma)(1-\alpha)} + \frac{v_{djkt}}{(1-\sigma)(1-\alpha)} - \frac{u_{jkt}}{1-\sigma}$$

- For revenue equation estimation, unobserved v_{djkt} causes the endogeneity problem. The decline of u_{jkt} has a "push" effect on n_{jkt}, thus is a valid instrumental variable.
- Propensity equations of "fresh" exporting firms:

$$\frac{n_{fjkt}}{s_{jt}} = A_{jk} + \sum_{t'=t_{ejk}+1}^{t} \frac{[q_{ekt'} - q_{dkt'}] + [\theta_e \log n_{ejk,t'-1} - \theta_d \log n_{jk,t'-1}] + v_{ejkt'}}{(t - t_{ejk})(\delta - 1)(1-\alpha)}$$

- For revenue equation estimation, unobserved v_{ejkt} causes the endogeneity problem. The growth of n_{jkt} has an "overhang" effect on n_{fjkt}, thus is a valid instrumental variable.

Instrumental Variable Construction: Revenue Equation

- The first instrument that we construct for network size takes advantage of the fact that the birth counties are rural
- Following Imbert et al. (2022) we thus construct a shift-share instrument for network size that is based on agricultural income shocks in the birth county that *push* individuals into business (u_{jkt} in the model)
  1. Assume that world crop prices follow an AR1 process and construct a price shock for 11 crops, that account for 96% of cultivated area, in each year
  2. Weight each crop's price shock by a factor that reflects its contribution to county-level agricultural production (by value) to construct a composite agricultural income shock in each year
  3. Assume that the decision to establish a firm and, hence, firm entry in a given year is based on a three-year average of the income shocks
  4. "Distribute" the entering firms across destination prefectures by dividing the county-level income shocks by distance

Instrumental Variable Construction

- Based on the trajectory pattern of network and the model, there is be a deterministic relationship between a network's duration and its contemporaneous growth rate
- We observe the precise point in time at which each networks in our data commenced, which, in turn, allows us to compute their durations any subsequent year
  - It is common practice in the migration literature, going back to Card (2001), to assume that initial settlement from an origin in a particular destination is exogenously determined
  - When constructing our second instrument, we make the weaker assumption that the *timing* of network formation is exogenous
  - This assumption is especially supported when the data have thousands of networks that started from a single firm and grew over time.
- Once we first-difference the revenue equations to purge firm fixed effects, the instruments need to predict changes (growth) in network size, and we see that both domestic and export network duration have sufficient power to do this

First-Stage and Second-Stage Estimates

- Based on our model, any variable that determines the growth in domestic network size will also be a valid instrument for the growth in export network size, since a larger domestic network dampens entry into exporting through the overhang effect
- Birth county income shocks and domestic network duration can thus be used as instruments for the growth in domestic network size, while both these variables and export network duration can be used as instruments for the growth in export network size First-stage Estimates
- The second-stage estimates indicate that firm revenues and productivity are increasing in network size, for both domestic production and exporting Second-stage Estimates
- Results are robust when we keep only firms outside the birth county and control birth county-time trends Firms Outside Birth County Birth County Time Trends
    - Our analysis complements a well established literature that documents the positive effect of migrant labor networks on the outcomes of their members (Munshi, 2003; Beaman, 2012; Heath, 2018; Barwick et al., 2023; Tang, 2024)
    - It is also the first to provide causal evidence, at the economy-wide level, that networks of firms can improve the outcomes of their members

Validating the Instruments

- We pass the over-identification test with domestic revenue and productivity as the outcome.
  - We check the robustness of the shift-share instrument further, by (1) restricting the sample to firm outside the birth county (2) excluding agricultural processing (3) controlling agriculture income shock (4) controlling distance-time effects. Robustness Check of Shift-Share Instrument
  - We follow Goldsmith-Pinkham, Sorkin and Swift (2020) method to validate each component of the shift-share instruments. Robustness Check by Each Crop
  - This suggest that the agriculture income shock and network duration are both valid instrumental variables.
- We fail the over-identification test with export revenue as the outcome
- One interpretation of this finding is that the instruments have heterogeneous treatment effects (Mogstad et al., 2021) Second-Stage Estimates by Instrument
  - The domestic duration is a valid instrument when existing the "overhang" effects.
  - The consistency across individual instruments further increases our confidence in their validity

Testing the Model: Independence of Networks between Birth Counties

Estimation:	OLS		2SLS	
Dependent variable:	log domestic revenue	log export revenue	log domestic revenue	log export revenue
	(1)	(2)	(3)	(4)
Log network size	0.584***	0.629***	1.482***	1.221***
	(0.035)	(0.029)	(0.178)	(0.176)
Log network size of the nearest birth county	0.094***	0.011	0.110	0.070
	(0.016)	(0.018)	(0.169)	(0.143)
Prefecture-time effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Kleibergen-Paap F	–	–	9.397	23.26
Observations	3,388,312	126,693	3,388,312	126,693

Note: Including firms of a birth county-destination that have both domestic network and export network.
Firm fixed effects are purged by first-differencing prior to estimation.
Instruments for each birth county's network in columns (3): domestic network duration, agriculture income shock.
Instruments for each birth county's network in columns (4): domestic network duration, export network duration.

Testing the Model: Independence between Domestic Network and Export Network for a Birth County

Method:	OLS		2SLS		
Dependent variable	log domestic revenue	log export revenue	log domestic revenue	log export revenue	log export revenue
	(1)	(2)	(3)	(4)	(5)
Log domestic network size	0.613***	0.002	1.561***	-0.626**	-0.756***
	(0.037)	(0.042)	(0.068)	(0.271)	(0.202)
Log export network size	-0.002	0.630***	-0.023	1.184***	1.126***
	(0.007)	(0.028)	(0.130)	(0.162)	(0.124)
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F	–	–	15.87	21.56	20.81
Observations	3,388,312	126,693	3,388,312	126,693	126,693

Note: Including firms of a birth county-destination that have both domestic network and export network.
Instruments in columns (3)-(4): domestic network duration, export network duration.
Instruments in column (5): domestic network duration, export network duration, and their interactions with initial entry.

- We cannot use IV estimation to prove the non-existence of domestic network effects on exporting revenue, with the existence of both heterogeneous treatment effects and "overhang" effects.

Quantify the Network Effect and the Overhang Effect

Firm Entry

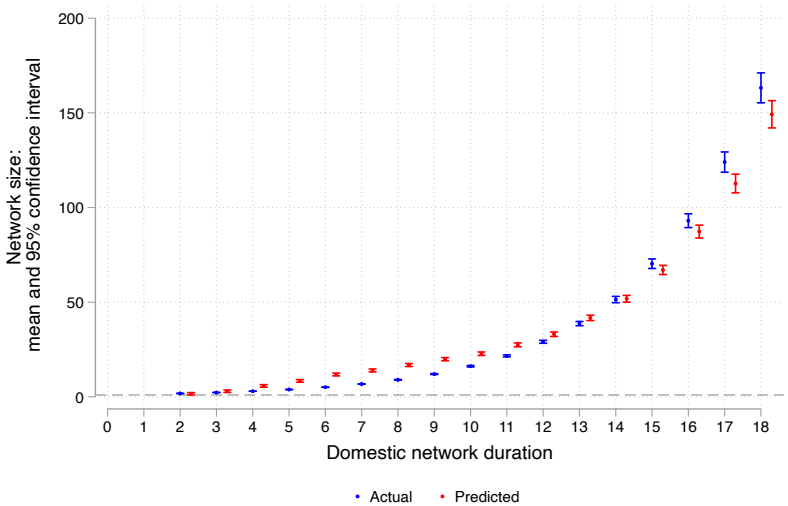
- Based on the model, with a slight abuse of notation, the propensity equations can be specified as follows

$$\frac{n_{jkt}}{s_{jt}} = A_{jk} + q_{dkt} + \theta_d \log n_{jk,t-1} + v_{djkt} - u_{jkt}$$

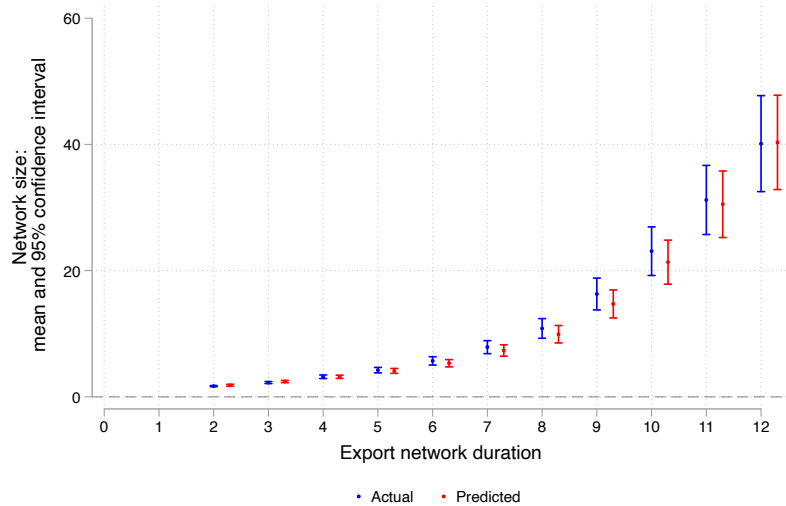
$$\frac{n_{ejkt}}{s_{jt}} = A_{jk} + \sum_{t'=t_{ejk}+1}^{t} \frac{[q_{ekt'} - q_{dkt'}] + [\theta_e \log n_{ejk,t'-1} - \theta_d \log n_{jkt'-1}] + v_{ejkt'}}{(t - t_{ejk})}$$

  - The income shock in the birth county, u_{jkt}, can no longer be used as an instrument in the entrepreneurial propensity equation
  - The network durations remain feasible, but they lack the power to predict the growth in average network size in the "fresh" export propensity equation
  - We thus include network duration interacted with initial entry as additional instruments (they pass the over-identification test in the cross-network revenue estimation)
  - We also estimate an "incumbent" export (Melitz-type exporter) propensity equation, where the size of the export network is the source of variation Propensity Equation Estimates

Non-targeted Moment Matching



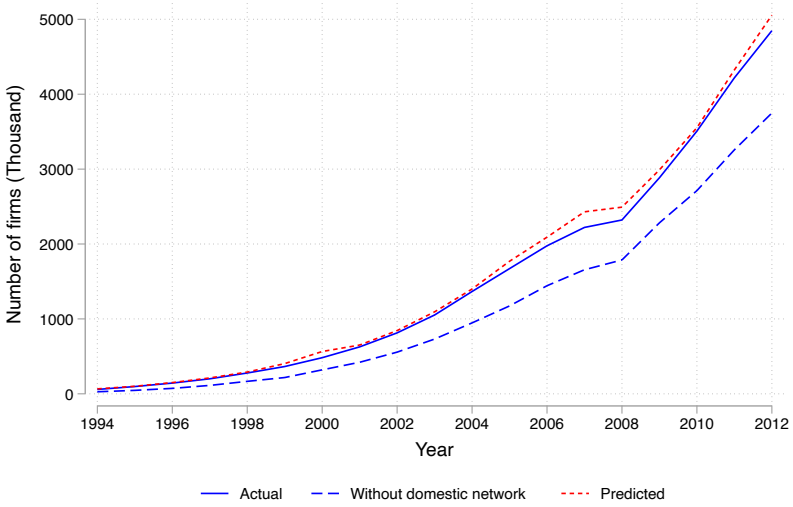
(a) Domestic Network



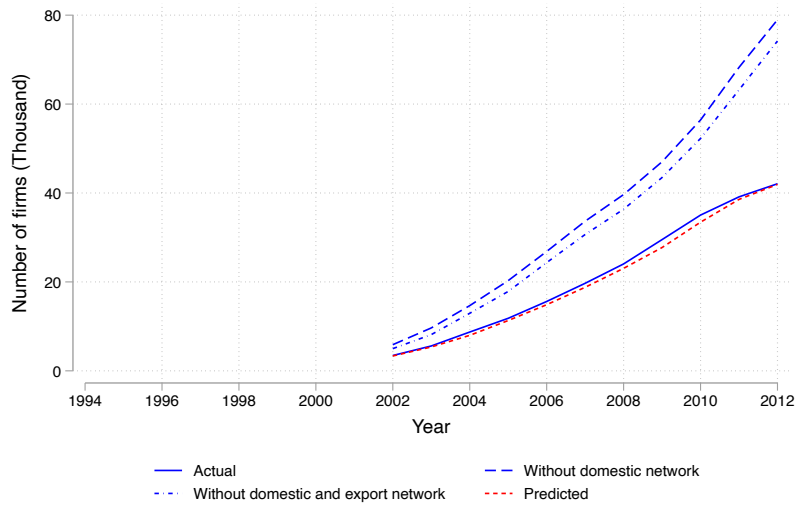
(b) Exporter Network

Source: Registration database and Customs database

Quantification Analysis



(a) Domestic Network



(b) Exporter Network

Source: Registration database and Customs database

- The total number of firms in 2012 would decline by 23 percent in the absence of the birth county networks
- The number of export firms in 2012 would have increased by 76 percent if the birth county networks were absent

Conclusion

- Our analysis examines the community networks that emerged in the first two decades after economic reforms in China, allowing *groups* of individuals to move into new activities
- While these networks played an important facilitating role in the initial transition to domestic production, the pre-existing domestic networks slowed the growth of newly emerging export networks and delayed the transition to the next stage of economic development
- Akcigit and Nicholas (2019) advocate for the use of historical data, theory and empirics to study economic growth, and our analysis exemplifies the value of this approach

Conclusion

- Entrepreneurs do not internalize their contribution to the networks and, hence, there is a role for entry and export subsidies
- While export subsidies are unambiguously efficiency enhancing, the entry subsidies must be attentive to their negative effect on export profits, due to the domestic network overhang
- The networks that we describe in this paper are not specific to business or to China, but their importance in other developing economies will depend on the underlying social structure and this will vary across regions of the world; e.g. Asia versus Africa

First-Stage Estimates of the Firm Performance Equations

Dependent variable:	growth in domestic network size	growth in export network size
	(1)	(2)
Birth county income shocks	-0.343***	0.516
	(0.094)	(0.770)
Domestic network duration	-0.004***	0.006***
	(0.000)	(0.001)
Export network duration		-0.012***
		(0.002)
Prefecture-time effects	Yes	Yes
Observations	5,211,514	126,971

Note: Network size is constructed from SAIC registration data and Customs data.
Growth in network size is measured by $\log n_{jk,t-1} - \log n_{jk,t-2}$ for the domestic network and $\log n_{ejk,t-1} - \log n_{ejk,t-2}$ for the export network.
Instruments include birth county income shocks, domestic network duration, and export network duration.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Second-Stage Estimates of the Firm Performance Equations

Estimation:	OLS			2SLS		
Dependent variable:	log domestic revenue	log domestic TFP	log export revenue	log domestic revenue	log domestic TFP	log export revenue
	(1)	(2)	(3)	(4)	(5)	(6)
Log network Size	0.504***	1.339***	0.630***	1.194***	2.932***	1.353***
	(0.023)	(0.072)	(0.029)	(0.073)	(0.194)	(0.141)
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F	–	–	–	118.2	118.2	42.80
Hansen J	–	–	–	3.558	1.178	4.760
Observations	5,211,514	5,211,514	126,971	5,211,514	5,211,514	126,971

Note: Network size is constructed from SAIC registration data and Customs data.
Revenue and TFP are constructed from SAIC inspection data and Customs data.
Firm fixed effects are purged by first-differencing prior to estimation.
The modified network variable is thus measured by the growth in its size: $\log n_{jk,t-1} - \log n_{jk,t-2}$ for the domestic network and $\log n_{ejk,t-1} - \log n_{ejk,t-2}$ for the export network.
Instruments for the growth in domestic network size: birth county income shocks, domestic network duration.
Instruments for the growth in export network size: export network duration, domestic network duration.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Second-Stage Estimates of Revenue Equations: Located Outside Birth County

Estimation:	OLS			2SLS		
Dependent variable:	log domestic revenue	log domestic TFP	log export revenue	log domestic revenue	log domestic TFP	log export revenue
	(1)	(2)	(3)	(4)	(5)	(6)
Log network Size	0.436***	1.244***	0.663***	0.342***	0.830***	1.388***
	(0.019)	(0.048)	(0.029)	(0.083)	(0.201)	(0.133)
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F	–	–	–	62.22	62.22	167.2
Observations	3,270,885	3,270,885	45,564	3,270,885	3,270,885	45,564

Note: Network size is constructed from SAIC registration data and Customs data.
Revenue and TFP are constructed from SAIC inspection data and Customs data.
Firm fixed effects are purged by first-differencing prior to estimation.
The modified network variable is thus measured by the growth in its size: $\log n_{jk,t-1} - \log n_{jk,t-2}$ for the domestic network and $\log n_{ejk,t-1} - \log n_{ejk,t-2}$ for the export network.
Instruments for the growth in domestic network size: birth county income shocks, domestic network duration.
Instruments for the growth in export network size: export network duration, domestic network duration.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Second-Stage Estimates of Revenue Equations: Birth County Fixed Effects

Estimation:	OLS			2SLS		
Dependent variable:	log domestic revenue	log domestic TFP	log export revenue	log domestic revenue	log domestic TFP	log export revenue
	(1)	(2)	(3)	(4)	(5)	(6)
Log network Size	0.423***	1.205***	0.608***	0.914***	2.348***	1.161***
	(0.017)	(0.043)	(0.028)	(0.115)	(0.246)	(0.126)
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Birth county fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F	–	–	–	74.63	74.63	175
Observations	5,211,514	5,211,514	126,901	5,211,514	5,211,514	126,901

Note: Network size is constructed from SAIC registration data and Customs data.
Revenue and TFP are constructed from SAIC inspection data and Customs data.
Firm fixed effects are purged by first-differencing prior to estimation.
The modified network variable is thus measured by the growth in its size: $\log n_{jk,t-1} - \log n_{jk,t-2}$ for the domestic network and $\log n_{ejk,t-1} - \log n_{ejk,t-2}$ for the export network.
Instruments for the growth in domestic network size: birth county income shocks, domestic network duration.
Instruments for the growth in export network size: export network duration, domestic network duration.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Robustness Check: Agricultural Income Shock

Sample:	all	outside birth county	excluding agricultural processing	all	all
Dependent variable:	log revenue				
	(1)	(2)	(3)	(4)	(5)
Log network size	1.314***	0.567***	1.302***	1.461***	1.344***
	(0.098)	(0.152)	(0.097)	(0.080)	(0.109)
Agriculture income shock	–	–	–	0.384**	–
				(0.149)	
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes
Distance-time effects	No	No	No	No	Yes
Kleibergen-Paap F	64.07	67.37	64.27	98.06	48.15
Observations	5,211,514	3,270,885	5,100,144	5,211,514	5,211,514

Note: Network size is constructed from SAIC registration data and Customs data.
Firm fixed effects are purged by first-differencing prior to estimation.
Instruments for the growth in network size: birth county income shocks.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Testing the Exogeneity of the Crop Shares: Agricultural Income Shock

Crop used to construct IV:	maize	potato	rapeseed	rice	wheat	soybean	sorghum
Dependent variable:	log domestic revenue						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log network size	1.669***	1.788***	1.104***	1.604***	1.239***	1.305***	1.646***
	(0.119)	(0.122)	(0.120)	(0.079)	(0.114)	(0.085)	(0.184)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F	19.01	13.39	6.822	17.17	5.314	9.262	3.760
Share	0.328	0.115	0.123	0.065	0.099	0.120	0.013
Rotemberg weight	0.432	0.140	0.121	0.099	0.091	0.073	0.044
Observations	5,211,514	5,211,514	5,211,514	5,211,514	5,211,514	5,211,514	5,211,514

Note: Network size is constructed from SAIC registration data and Customs data.
Firm fixed effects are purged by first-differencing prior to estimation.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Second-Stage Estimates of the Firm Performance Equations by Instrument

Instrument:	birth county income shocks	domestic network duration	export network duration	domestic network duration
Dependent variable:	log domestic revenue		log export revenue	
	(1)	(2)	(3)	(4)
Log network size	1.560***	1.157***	0.954***	1.363***
	(0.234)	(0.075)	(0.213)	(0.142)
Domestic network duration	0.002*	–	0.002**	–
	(0.001)		(0.001)	
Birth county income shocks	–	-0.138**	–	–
		(0.070)		
Export network duration	–	–	–	0.005**
				(0.002)
Prefecture-time effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Kleibergen-Paap F	13.30	224.4	27.12	82.31
Observations	5,211,514	5,211,514	126,971	126,971

Note: Instruments for the growth in domestic network size: birth county income shocks or domestic network duration.
Instruments for the growth in export network size: export network duration or domestic network duration.
The excluded instrument is included as a covariate in each case.
Standard errors clustered at the birth county level are reported in parentheses. * significant at 10%, ** at 5%, *** at 1%.

Go Back

Propensity Equation Estimates: Birth County - Destination Level

Method:	OLS		2SLS		OLS	2SLS
Dependent variable	entrepreneurial propensity	incumbent exporter propensity	entrepreneurial propensity	incumbent exporter propensity	fresh exporter propensity	
	(1)	(2)	(3)	(4)	(5)	(6)
Log network size	0.012***	0.001***	1.903***	0.005***	–	–
	(0.000)	(0.000)	(0.076)	(0.000)		
Average log export network size	–	–	–	–	0.023***	0.083***
					(0.001)	(0.010)
Average log domestic network size	–	–	–	–	-0.003**	-0.397***
					(0.001)	(0.068)
Prefecture-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Birth county-prefecture effects	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F	–	–	465.2	521.2	–	18.17
Observations	913,802	22,272	913,802	22,272	22,573	22,573

Note: Prefecture-time effects and birth county-prefecture effects controlled. Birth county-prefecture fixed effects are purged by first-differencing prior to estimation.
Instruments for the growth in domestic network size: domestic network duration and its interaction with initial entry.
Instruments for the growth in average export and domestic network size: export network duration and its interaction with initial entry, domestic network duration and its interaction with initial entry.
Instruments for the growth in export network size: export network duration and its interaction with initial entry, domestic network duration and its interaction with initial entry.
Standard errors clustered at the birth county level are reported in parentheses.

Go Back