Startup Location, Local Spillovers and Neighborhood Sorting

Rolando Campusano

University of Toronto (Rotman)

October 28, 2021

Research question and quick summary

How critical is economic concentration for the success of startup firms?

In a nutshell

- data on the universe of firms/workers in all industries in large Canadian cities
- broad definition of startup - using ownership info and labor tracking
- hyper-proximity at the level of city blocks and their surrounding neighborhoods
- the focus is on correcting for location sorting

Local spillovers apply to startups at micro-geographies

Causal evidence of **large** and **very-local** spillovers from incumbents to new firms

- Revenue consistent with productivity effects
  incumb average from p10-p90 pct
  $\Longrightarrow$ p50 startup revenue $\Uparrow$8.2%, employment $\Downarrow$8.6%
- Employment consistent with scale effects
  incumb average from p10-p90 pct
  $\Longrightarrow$ p50 startup revenue $\Uparrow$78%, employment $\Uparrow$121%

These effects are highly **heterogeneous** across industries

- employment-intensive startups better off with *larger* incumbents
- knowledge-intensive startups better off with *better* incumbents

And across type of exposure

Zoom IN/OUT approach

Identification Problem: location sorting

Strategy: zoom in/zoom out - distinction between blocks and neighborhoods

- Sorting into blocks corrected using *economic* neighborhoods
  - propensity score captures the likelihood that a block is suitable for a new startup Heblich et al. (2019) and Qian and Tan (2021)
  - industry specific neighborhoods using Campusano (2021)
  - exclusion restriction: conditional on neigh-year no sorting within neighborhoods
- Sorting into neighborhoods corrected using location choice model
  - sorting based on expected startup outcomes and preferences for location
  - exclusion restriction: owners consider location personal preferences

entrepreneurs choose neighborhoods and are randomly assigned to a block

Local Spillovers

Providing random allocation within neighborhoods, neighborhood-year fixed effects identify local spillovers θ within neighborhoods

$$y_{bnt}^{(i,j)} = x_t^{(i,j)}\beta + X_{bnt}^{(j)}\theta + \lambda_t^{(j)} + \delta_{nt}^{(j)} + \epsilon_{bnt}^{(i,j)} \tag{1}$$

(incumbents same industry)	Revenue	Employment	Alive $t+5$	Move Blocks $t+5$	Move Nbhds $t+5$
Log (Average Employment$_{t-1}$)	**0.120****	**0.218*****	0.018**	-0.030***	-0.034***
	(0.052)	**(0.044)**	(0.006)	(0.005)	(0.004)
Log (Average Revenue$_{t-1}$)	**0.027****	**-0.026***	0.000	0.003**	0.004***
	(0.012)	**(0.014)**	(0.002)	(0.001)	(0.001)
Number of Observations (Startups)	30,318	20,675	28,989	28,528	37,529

Controls: Ownership structure, number of incumbents, dummy for zero activity.
FE: neighborhood year, industry-year, city-year
Std. Errors clustered at Neighborhood-Year level
Estimation: HDFE and many zeroes in outcome variable => PPML

Local Spillovers

- These effects are very local ▸ Spatial Decay
  - decay rapidly from 75 meters
- Remarkable heterogeneity across industries ▸ Industry Heterogeneity
  - Employment intensive industries benefit more from higher employment
  - Knowledge intensive industries benefit more from higher average revenue
- And across type of exposure ▸ Cross-Industry Linkages
  - Similar firms or workers have higher effect than input-output linkages
- All of this is consistent with theories of agglomeration that highlight the role of knowledge flows between workers rather than productivity spillovers that arise from agglomeration

Local Spillovers and Neighborhood Sorting

Sorting into neighborhoods corrected using extended Roy model of neighborhood choice

Adapting Lee (1983) and Dahl (2002) leads to a two step procedure

1. (First stage) location choice probabilities for all possible neighborhoods
   - entrepreneurs' demographics + distance to first residential location
   - neural networks to estimate entrepreneur-specific probabilities
2. (Second stage) use probabilities to correct for self-selection into neighborhoods

$$y_{bnt}^{(i,j)} = x^i\beta + X_{bnt}^{(j)}\theta + \delta_{nt}^{(j)} \times \lambda_{nt}\left(\overset{star}{\mathrm{P}}{}_{nt}^{(i,j)}, \overset{home}{\mathrm{P}}{}_{nt}^{(i,j)} \right) + \omega_{bn}^{(i,j)} \tag{2}$$

Local Spillovers and Neighborhood Sorting

Accounting for sorting into neighborhoods

- increases the elasticities to average incumbent employment between 25 to 50%
- decreases the elasticities to average revenue between 30 to 40%

This indicates that, on average, scale effects are more important for new firms early stages.

	End of Year Revenue			End of Year Employment		
	$\delta_{nt}^{(j)}$	$\delta_{nt}^{(j)}$	$\delta_{nt}^{(j)} \times \mathrm{P}_{nt}^{home(i,j)}$	$\delta_{nt}^{(j)}$	$\delta_{nt}^{(j)}$	$\delta_{nt}^{(j)} \times \mathrm{P}_{nt}^{home(i,j)}$
Log (Average Employment$_{t-1}$)	0.120**	0.169***	0.176***	0.218***	0.316***	0.313***
	(0.0523)	(0.0197)	(0.0311)	(0.0442)	(0.0167)	(0.0396)
Log (Average Revenue$_{t-1}$)	0.0273**	0.0381***	0.0183**	-0.0257*	-0.0159**	-0.0223**
	(0.0129)	(0.0100)	(0.00699)	(0.0139)	(0.00532)	(0.0100)
Number of Observations (Startups)	30,318	247,594	125,676	20,675	249,938	87,760

Conclusions

- A substantial amount of evidence of productivity spillovers for regions and cities
- This paper uses variation within and across neighborhoods to provide causal evidence of very local spillovers for new firms while accounting for entrepreneurs location sorting
- Local spillovers are positive and very local and manifest in startups' short and medium-term outcomes
- Firms benefit more from exposure to larger firms than from exposure to firms with higher sales
- Accounting for sorting accentuates the patterns observed within neighborhoods

Thank YOU!

Comments, criticisms and suggestions are super welcome

Rolando.Campusano@Rotman.UToronto.Ca

References I

- Campusano, Rolando (2021). "Delineating Neighborhoods Using Location Choices."
- Dahl, Gordon (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." In: *Econometrica* 70.6, pp. 2367–2420. ISSN: 0012-9682.
- Heblich, Stephan et al. (May 21, 2019). *Industrial Clusters in the Long Run: Evidence from Million-Rouble Plants in China*. 19/712. School of Economics, University of Bristol, UK.
- Lee, Lung-Fei (1983). "Generalized Econometric Models with Selectivity." In: *Econometrica* 51.2, pp. 507–512. ISSN: 0012-9682.
- Qian, Franklin and Rose Tan (2021). "The Effects of High-Skilled Firm Entry on Incumbent Residents."

*Zoom IN: Local spillovers are positive, but very local

These effects are very local

- decay quickly one block away

	Same Block	1st Ring 150m	2nd Ring 225m	3rd Ring 300m
End of Year Revenue				
Log (Average Employment Same Industry)	0.118**	0.0621*	-0.0230	0.0154
	(0.0489)	(0.0350)	(0.0508)	(0.0280)
Log (Average Revenue Same Industry)	0.0259**	0.00739	-0.0168*	-0.00138
	(0.0123)	(0.0133)	(0.0101)	(0.00802)

Each panel is one regression. Coefficients correspond to measure of variable in the first column

Back

*Zoom IN: heterogenous across industries

- Employment intensive and industries benefit more from higher employment
- Knowledge intensive industries benefit more from higher average revenue

End of Year Revenue	Information / Financial Services	Manufacturing	Professional / Business Services	Retail, Leisure and Hospitality	Transport / Wholesale Trade
Log (Average Employment Same Industry)	-0.227**	0.812***	0.0739	0.209***	0.340**
	(0.106)	(0.189)	(0.114)	(0.0459)	(0.126)
Log (Average Revenue Same Industry)	0.0530**	-0.0802*	0.0293*	0.00145	0.0143
	(0.0187)	(0.0442)	(0.0159)	(0.0129)	(0.0185)

Each panel is one regression. Coefficients correspond to a dummy for a group of industries interacted with the variable in the first column.

▸ Back

Zoom IN: and type of industry exposure

Exposure not only to same industry

- Use of input-output weights (StatsCan 2001)
- Use of occupational similarity weights (BLS 2001)

	Same	All	Downstream	Upstream	Occ. Similarity
End of Year Revenue					
Log (Average Employment)	0.120**	0.186***	0.0437**	0.0407**	0.141***
	(0.0523)	(0.0296)	(0.0147)	(0.0149)	(0.0261)
Log (Average Revenue)	0.0273**	0.0319**	0.0907***	0.100***	0.0637***
	(0.0129)	(0.0145)	(0.0130)	(0.0135)	(0.0131)

Back