Input Diffusion and the Evolution of Production Networks

Vasco Carvalho
Cambridge and CEPR

Nico Voigtländer
UCLA and NBER

NBER Meeting San Francisco
26 February 2015

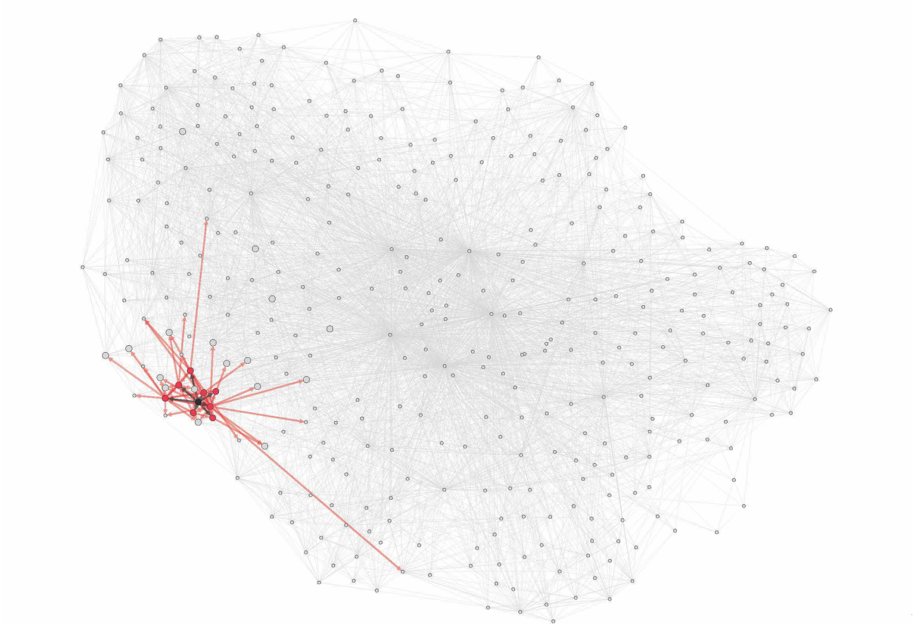
Motivation

- Adoption of inputs is an important dimension of technical progress
- Recent literature also stresses the role of input linkages:
  - for aggregate productivity outcomes
  - in propagating micro shocks and generating aggregate fluctuations

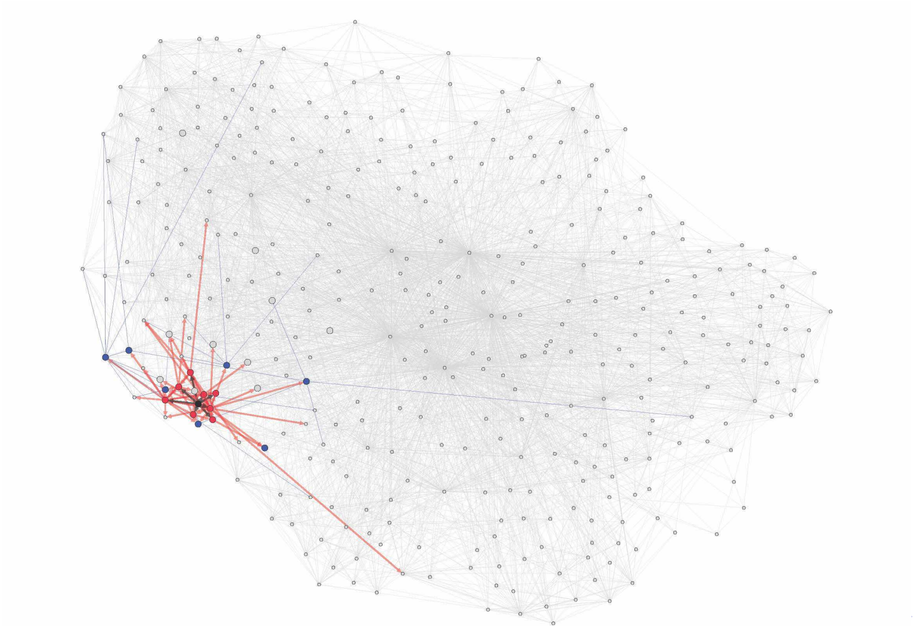
Motivation

- Adoption of inputs is an important dimension of technical progress
- Recent literature also stresses the role of input linkages:
  - for aggregate productivity outcomes
  - in propagating micro shocks and generating aggregate fluctuations
- This paper:
  - Analyze formation of input-output linkages through a network perspective
  - Empirics: document novel pattern in the data: Producers tend to adopt new inputs from the network neighborhood of their existing suppliers
  - Theory: stylized model of networked input search & adoption

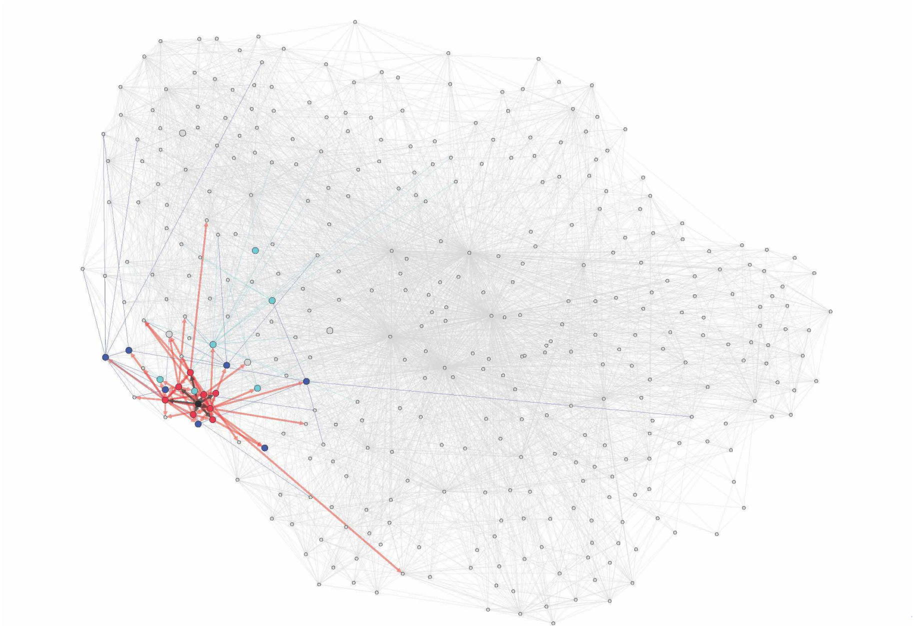
Diffusion of Semiconductors. I-O Network in 1967



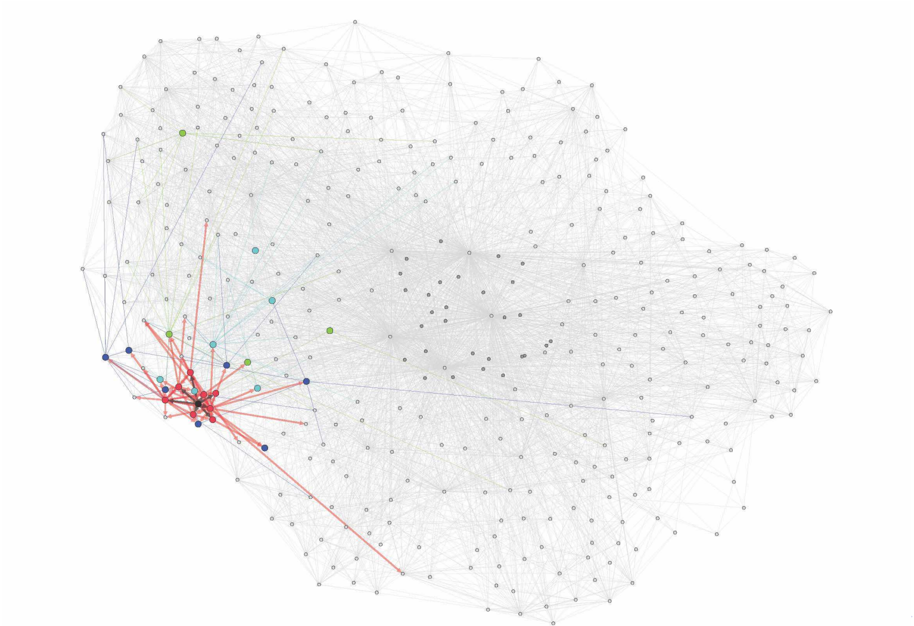
1972



1977



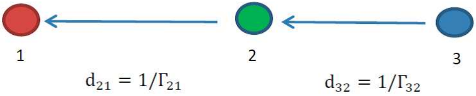
1982



Network Distance

- BEA Input-Output Tables, 4-digit 1967-2002. Define $i - j$ pairs:
    - i: potential input supplier
    - j: potential adopter
- Network distance d_{ij}: minimum-distance path linking sector j to potential supplier i (directed, weighted by input flows) ▸ Detail
- Focus on i-j pairs that are not (yet) directly connected

Distance between input $i = 3$ and sector $j = 1$

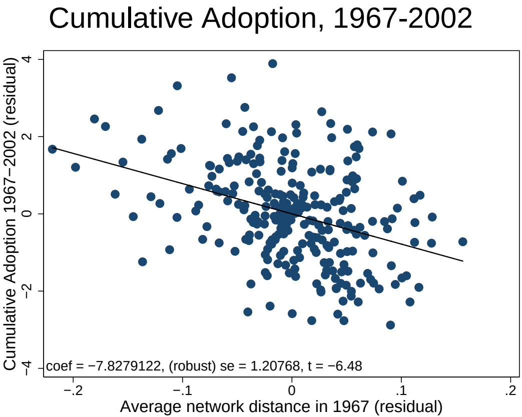


Main finding illustrated in a single graph

- For each sector *i* ("input supplier"), compute its *average* network distance to all other sectors *j* (potential adopters) in 1967
- "Cumulative adoption" = number of sectors *j* that adopt *i* until 2002

Main finding illustrated in a single graph

- For each sector *i* ("input supplier"), compute its *average* network distance to all other sectors *j* (potential adopters) in 1967
- "Cumulative adoption" = number of sectors *j* that adopt *i* until 2002



Findings and their Relationship to the Literature

- Role of networks in input and technology adoption
  - Diffusion of innovations in social networks (e.g., Conley & Udry, 2010; Banerjee et al., 2013)
- Growth by recombination of ideas (Weitzman, 1998)
- Evolution of input-output networks under random search (Oberfield, 2013).
- Generalized diffusion (e.g. GPT) is more likely when input is used by central producers in the network
  - Novel implication for GPT literature (e.g. Helpman and Trajtenberg, 1998; Jovanovic and Rousseau 2005)
- Out-degree distribution follows a power law
  - Consistent with data. Key for propagation of shocks as stressed in Acemoglu et al. (2012)

Plan for the Talk

1. Empirics:
    - 4-digit SIC sectors
    - Firm level, based on Compustat data
2. Theory: Sketch model of input search and adoption

Does network proximity between two sectors predict subsequent input adoption?

Probit, OLS, Hazard model:

$$Prob\left(A_{ij}(y) = 1\right) = g\Big(d_{ij}(y-5),\ X_i(y),\ X_j(y)\Big)$$

- $A_{ij}(y)$: indicator for sector j adopting input i in year y
- $d_{ij}(y-5)$: (directed) network distance b/w i and j, lagged by 5 years
- $X_i(y), X_j(y)$: controls for input-producing/adopting sector (e.g., TFP, fixed effects)

Panel Results on Input Adoption

Dep. Var.: Dummy for adoption of input i by sector j in year y

Estimation	(1) Probit	(2) Probit	(3) OLS	(4) OLS	(5) Hazard	(6) Hazard
Distance $d_{ij}(y-5)$	-0.1885***	-0.1882***	-0.0092***	-0.0092***	0.5947***	0.5955***
	(0.0041)	(0.0041)	(0.0002)	(0.0002)	(0.0059)	(0.0058)
	[-2.34%]	*[-2.34%]*	*[-1.45%]*	*[-1.45%]*	*[-4.14%]*	*[-4.20%]*
$\triangle_5 TFP_i$		0.1131***		0.0154***		1.5061***
		(0.0365)		(0.0025)		(0.1068)
		[0.07%]		*[0.12%]*		*[0.16%]*
$\triangle_5 TFP_j$		-0.0950		-0.0064		1.2088
		(0.1453)		(0.0113)		(0.3588)
Observations	577,498	577,498	577,498	577,498	577,498	577,498

Notes: Standard errors in parentheses, clustered at the adopting sector (j) level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Values in *[square brackets]* are standardized coefficients, reflecting the change in adoption probability (over a 5-year interval) due to a one standard deviation increase in the explanatory variable.

Robustness Check

Panel results are robust to

- Require use of new inputs for $\geq$15 years to qualify as adoption
- Exclude new links formed within 2-digit sectors
- Use only initial network distance in 1967
- Controls for i and j (employment, fixed effects, TFP level)
- Consider only links with $\geq$\$1mio purchase

Panel Results on Input Adoption: Placebo

No predictive power of forward distance

Dep. Var.: Dummy for adoption of input *i* by sector *j* in year *y*

Estimation	(1) Probit	(2) Probit	(3) OLS	(4) OLS	(5) Hazard	(6) Hazard
Forward Distance $d_{ji}(y-5)$	0.012 (0.012) *[0.11%]*	0.012 (0.012) *[0.11%]*	0.029^{**} (0.012) *[0.18%]*	-0.013 (0.013) *[-0.07%]*	-0.013 (0.013) *[-0.07%]*	0.016 (0.012) *[0.01%]*
Distance $d_{ij}(y-5)$		-0.205^{***} (0.011) *[-1.61%]*	-0.199^{***} (0.012) *[-1.57%]*		-0.367^{***} (0.024) *[-1.24%]*	-0.356^{***} (0.026) *[-1.24%]*
Controls	✓	✓	✓	✓	✓	✓
Observations	501,539	501,539	418,734	358,390	358,390	292,244

Notes: Standard errors in parentheses, clustered at the adopting sector (*j*) level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Values in *[square brackets]* are standardized coefficients, reflecting the change in adoption probability (over a 5-year interval) due to a one standard deviation increase in the explanatory variable.

Does network proximity lead to faster adoption?

Time-to-Adopt Regressions

$$T_{ij} = \beta \cdot d_{ij}^{67} + \gamma \cdot \triangle Efficiency_i + \delta_i + \eta_j + \varepsilon_{ij}$$

- T_{ij}: Years until sector j adopts input i after 1967 (not defined if no adoption by 2002)
- d_{ij}^{67}: network distance in 1967
- $\triangle Efficiency_i$ (average annual) change in efficiency in input-producing sector (TFP, price)
- δ_i and δ_j: input-producing and adopting sector fixed effects

Time to Adoption

Dep. Var.: Time to adoption of input *i* by sector *j* after 1967

	(1)	(2)	(3)	(4)	(5)	(6)
Years excluded	1997	1997	1972,97	none	1997	1997
Other remarks					2-digit†	narrow‡
Distance d_{ij} in 1967	0.937^{***} (0.196) *[0.64]*	3.112^{***} (0.341) *[2.14]*	1.778^{***} (0.360) *[1.15]*	3.104^{***} (0.311) *[2.04]*	3.307^{***} (0.354) *[2.28]*	1.228^{***} (0.290) *[0.72]*
$\triangle TFP_i(1967 - y_{adopt})$	-96.925^{***} (3.919) *[-1.78]*	-364.787^{***} (13.186) *[-6.70]*	-331.477^{***} (26.434) *[-3.95]*	-281.502^{***} (11.861) *[-4.37]*	-376.759^{***} (14.029) *[-6.97]*	-146.929^{***} (12.575) *[-3.06]*
Using Sector FE	✓	✓	✓	✓	✓	✓
Producing Sector FE		✓	✓	✓	✓	✓
R^2	0.19	0.73	0.72	0.67	0.73	0.66
Observations	14,849	14,849	8,604	24,312	13,856	6,421

Notes: Standard errors in parentheses, clustered at the adopting sector (*j*) level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Values in *[square brackets]* are standardized coefficients, reflecting the change in the dependent variable due to a one standard deviation increase in the explanatory variable.

† Column 5 excludes all *i*-*j* pairs that belong to the same 2-digit industry.

‡ The narrow definition of adoption requires new *i*-*j* pairs to be present for at least 15 years in order to qualify as adoption.

Firm-Level Results

Compustat data (customer segment file) 1977-2008

- Customers of a given firm that account for more than 10% of sales
- 43,506 firm-to-firm links
- Compute (binary) network distance

Dep. Var.: Dummy for firm *j* adopting inputs from firm *i* in a 5-year time interval *y*

	(1) OLS	(2) Probit	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
Sample					2-digit†	Manufacturing	Services
$I_{ij}(y-5)$	0.02854***	1.61359***	0.02161***	0.02140**	0.01834**	0.01966**	0.02367*
	(0.00745)	(0.11780)	(0.00779)	(0.00888)	(0.00806)	(0.00904)	(0.01348)
	[2.85%]	*[2.69%]*	*[2.16%]*	*[2.14%]*	*[1.83%]*	*[1.97%]*	*[2.37%]*
ln(geodistance)	-0.00006***	-0.05809***	-0.00007***	-0.00007***	-0.00006***	-0.00006***	-0.00007***
	(0.00001)	(0.00604)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00001)
	[-0.007%]	*[-0.005%]*	*[-0.007%]*	*[-0.007%]*	*[-0.006%]*	*[-0.006%]*	*[-0.007%]*
$\triangle_5 \ln(Y/L)_i$				0.00003***	0.00003***	0.00003**	0.00003
				(0.00001)	(0.00001)	(0.00001)	(0.00002)
Controls				✓	✓	✓	✓
Using Firm FE			✓	✓	✓	✓	✓
Producing Firm FE			✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Observations	14,634,939	14,634,939	14,634,939	8,895,481	8,461,685	4,906,536	3,381,959

Notes: The dependent variable is a dummy that takes on value 1 if firm *j* adopts input *i* in a given 5-year interval *y* between 1977 and 2006. $I_{ij}(y-5)$ is an indicator that equals one if firms *i* and *j* were indirectly linked (had a binary distance of 2) in the previous five-year interval. The variable geodistance is the geographical distance between *i* and *j*. $\triangle_5 \ln(Y/L)_i$ denotes the change in output per worker in the input-producing firm (*i*) over the previous (lagged) 5-year interval. Controls include the change in output per worker in the input-using firm over the previous 5 year interval ($\triangle_5 \ln(Y/L)_j$), as well as output per worker and ln(employment) for *both* input-producing and input-using firms.

Model – Overview

Model structure – variety level

- Build on models of dynamic network formation (Jackson and Rogers, 2007; Chaney, 2013)
- Every period t, a new variety arrives exogenously
- Variety production uses labor and intermediate inputs
- Input choice made in period t; fixed thereafter

Input adoption occurs in 2 steps:

1. Network Search: Identify *potential* inputs
2. Adoption decision

Aggregation from variety-level to sector-level

Step 1: Network Search for Potential Inputs

Producer of new variety t:

- Randomly draws a set K_t of "essential" input varieties
  - e.g. if t is a car: K_t includes wheels, body, engine
- Randomly chooses a set N_t of potentially useful input varieties from the network neighborhood of K_t
  - e.g. make car lighter: search among producers that supply body materials (BMW i3: ultra-light carbon fiber body)

▸ Outdegree Equation

Step 2: Input Adoption

Essential inputs

- No customization costs. All are used.

Network inputs

- Input-specific random customization costs
- Trade-off between:
  1. Gains from input variety à la Romer (1990)
  2. Input-specific (randomly drawn) customization costs
- Endogenous optimal number of network inputs is adopted
  - In expectation: identical across varieties

▸ Variety Production Function ▸ Graph

Main Implications

- Adoption of input i by variety t is more likely...
  - If i is in t's network neighborhood (search)
  - If the price of i is relatively low (adoption)
- Aggregation to Sector-Level
  - Use assignment rule based on essential inputs (also used by BEA)
  - Variety-level results hold

Main Implications

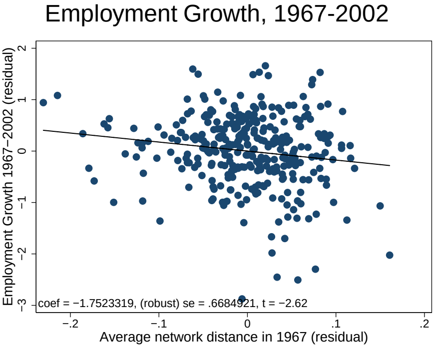
- Adoption of input i by variety t is more likely...
  - If i is in t's network neighborhood (search)
  - If the price of i is relatively low (adoption)
- Aggregation to Sector-Level
  - Use assignment rule based on essential inputs (also used by BEA)
  - Variety-level results hold
- The out-degree distribution follows a power law
  - Emergence of "star" varieties/sectors that serve as inputs to many other varieties

Conclusion

- Analyze input adoption from a network perspective theoretically and empirically
- Initial network proximity raises likelihood of input adoption. Interpretation:
  - Search for inputs along supplier relationships
  - Technological proximity: ‘Closer’ inputs are more useful and/or easier to integrate
- Important implications for growth
  - Emergence of GPTs
  - "Growth bottlenecks": distortions to gateways for adoption
  - Predicting sector-specific growth

Network Distance and Growth

Average network distance in 1967 is a strong predictor of subsequent growth



BACKUP

Computers Adopting Semiconductors

- Early computers: used vacuum tubes, no semiconductors

- 1960s: start using transistors (‘Electronic Components’)
  - Transistors in turn used semiconductors
  - Input flow:
    Semiconductors $\Rightarrow$ Electronic Components $\Rightarrow$ Computers
  - But: Semiconductors $\not\Rightarrow$ Computers in 1967 I-O Table

- Early 1970s: switch to integrated circuits/microprocessors
  - Integrated circuits: rely heavily on semiconductors
  - Adoption of semiconducting material in motherboard and other components
  - 1972 I-O Table: Semiconductors $\Rightarrow$ Computers

Evolution of Outdegree

Growth rate of variety *i*'s outdegree:

$$\frac{\partial d_i^{out}(t)}{\partial t} = p_K \frac{m_K}{t} + p_N \frac{m_K d_i^{out}(t)}{t} \frac{m_N}{m_K (p_K m_K + p_N m_N)}$$

- t: overall number of varieties in the economy at time t
- m_K: number of essential inputs that the new variety t draws
- m_N: number of network inputs that t identifies as potentially useful
- p_K, p_N: adoption probabilities

▸ Back to talk

Variety Production Function

Output of variety t:

$$y_t = \frac{A_t}{1 + C_t} \left(\mathbf{X}_t^K \right)^{\alpha} \left(\mathbf{X}_t^N \right)^{\beta} l_t^{1-\alpha-\beta}$$

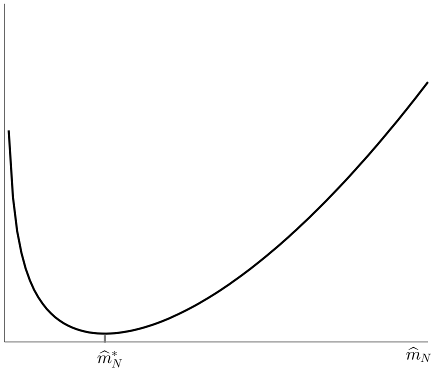
- $C_t = \sum_{n \in \widehat{N}_t} c_{t,n}$: (annualized) customization cost of adopted inputs
  $n \in \widehat{N}_t$; $c_{t,n} = b \cdot r_{t,n}$ with $b > 0$ and $r_{t,n}$ uniform random
- $\mathbf{X}_t^K = \left(\sum_{k \in K_t} x_{tk}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}$: composite of essential inputs
- $\mathbf{X}_t^N = \left(\sum_{n \in \widehat{N}_t} x_{tn}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}$: composite of adopted network inputs

Cost minimization: $\Rightarrow$ optimal choice of $\widehat{N}_t$:

$$\widehat{N}_t^* = \underset{\widehat{N}_t \subseteq N_t}{\arg\min} \left\{ \left(1 + \sum_{n \in \widehat{N}_t} c_{t,n} \right) \left(\sum_{n \in \widehat{N}_t} \phi_n^{\frac{1}{1-\epsilon}} \right)^{\frac{\beta}{1-\epsilon}} \right\}$$

Back to talk

Optimal number of adopted network inputs



Notes: The figure illustrates the optimal choice of input adoption. The x-axis shows the number of adopted network inputs, $\widehat{m}_N$. These are ranked by their customization cost. The y-axis shows the term from equation (8) that is proportional to marginal production cost, and that an input adopter seeks to minimize. For small $\widehat{m}_N$, the input variety effect à la Romer (1990) dominates, so that production costs are decreasing if more inputs are adopted. For higher $\widehat{m}_N$, customization costs for each additional adopted input are also high, outweighing the input variety effect. Thus, production cost become increasing in $\widehat{m}_N$. The optimal number of adopted network inputs is denoted by $\widehat{m}_N^*$.

▸ Back to talk

Towards Empirics: Measurement of Network Distance

- Direct-requirements input-output matrix Γ. Γ_{ij}: cost share of input i in the total intermediate input expenditures of sector j.
- If $\Gamma_{ij} > 0$: define distance from j to i as $d_{ij} = \frac{1}{\Gamma_{ij}}$
- If $\Gamma_{ij} = 0$ (i.e., j does not directly source inputs from i) but j is further downstream from i, then d_{ij} is the sum of the distances connecting i and j
  - If several such paths exist, d_{ij} is the minimum distance path linking i to j.

▸ Back to talk

1/1

Adoption of Inputs in the Data: Example

SIC Sector 3661 (Telephone and telegraph apparatus)

- 1972: adopts Adhesives and sealants (SIC 2891), Metal coating and allied services (SIC 3479)
- 1982: adopts Mechanical measuring devices (SIC 3820)
- 1987: adopts Electrometallurgical products (SIC 3313), Relays and industrial controls (3625)
- 1997: adopts Environmental controls (SIC 3822), Porcelain electrical supplies (3264)

Time to Adoption – Additional Results

Dep. Var.: Time to adoption of input *i* by sector *j* after 1967

	(1)	(2)	(3)	(4)	(5)	(6)
Remarks			2SLS[†]			narrow[‡]
Distance d_{ij} in 1967	1.620^{***} (0.181) *[1.11]*	0.968^{***} (0.212) *[0.66]*	0.976^{***} (0.182) *[0.67]*	3.464^{***} (0.323) *[2.37]*	3.148^{***} (0.327) *[2.17]*	0.889^{***} (0.229) *[0.52]*
$\triangle TFP_i(1967 - y_{adopt})$			-211.401^{***} (24.137) *[-3.88]*		-147.042^{***} (14.199) *[-2.70]*	
$\triangle P_i(1967 - y_{adopt})$	99.765^{***} (3.029) *[4.13]*			157.734^{***} (4.282) *[6.52]*	134.913^{***} (4.760) *[5.40]*	130.989^{***} (2.740) *[5.31]*
$\triangle TFP_i(1958 - 67)$		-18.341^{***} (6.189) *[-0.28]*				
Using Sector FE	✓	✓	✓	✓	✓	✓
Producing Sector FE				✓	✓	✓
R^2	0.26	0.17	0.16	0.76	0.77	0.82
Observations	15,072	15,072	14,849	15,072	14,849	6,456

Notes: Standard errors in parentheses, clustered at the adopting sector (*j*) level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Values in *[square brackets]* are standardized coefficients, reflecting the change in the dependent variable due to a one standard deviation increase in the explanatory variable.

[†] Two stage least square regression uses historical TFP growth in input-producing sectors ($\triangle TFP_i$ 1958-67) as in instrument for TFP growth after 1967 ($\triangle TFP_i$ since '67). The first stage has an F-statistic of 807.

[‡] The narrow definition of adoption requires new *i*-*j* pairs to be present for at least 15 years in order to qualify as adoption.