The Spatial Diffusion of Technology

Diego Comin
Harvard University

Mikhail Dmitriev
Boston College

and

Esteban Rossi-Hansberg
Princeton University

May 2012

Introduction

- Technological differences across countries are critical to explain differences in output per capita
- Empirical studies have treated adoption units independently
- Restrictive assumption because:
  - adoption requires acquiring knowledge which often comes from other agents.
  - return to adopting may increase if agents in other countries have adopted too
- In this paper we study empirically the diffusion of technology adoption over time and space
  - Uncover geographical co-movement patterns in adoption
    - ★ General patterns
    - ★ Unavoidable reflection problem (OVB); We (significantly) raise the barr for OV to explain the data
  - Develop a simple model of spatial technology diffusion
  - Use empirical findings to inform the parameters of the model

Literature

- Micro studies (Foster and Rosenzweig, 1995, Bandeira and Rasul, 2006) find role of social interactions and informational frictions in adoption of simple agricultural technologies.
  - How relevant for other technologies?
  - How relevant to explain cross-country differences in adoption?
- Macro literature (e.g., Keller, 2004)

$$TFP_c = F(R\&D_k, d_{ck})$$

  - Long adoption lags
  - International co-movement in output
  - Cyclicality of R&D and TFP
  - $->$ likely to reflect co-movement in demand rather than knowledge spillovers

Data

- Use direct measures of technology
- Data comes from CHAT dataset (Comin and Hobijn, 2004)
  - 161 countries (unbalanced)
  - Early ones starting in 1825 data up to 1990's
  - Use 20 significant technologies

Technologies and Sectors

Sector	Name
Transportation	Aviation Passengers
	Aviation Tons
	Cars
	Rail Line Km
	Rail Passengers
	Rail Tons
	Ships
	Trucks
Communication	Cellphone
	Computer
	Internet
	Radio
	Telegram
	Telephone
	TV
Industry	ATM
	Electricity
	Steel Blast Oxygen
	Steel Electric Arc
	Tractors

Empirical Methodology

- Let j denote technology, c country, s sector, and t year
- x^j_{ct} : Log of technology adoption (e.g. tons*Km transported by rail)
- y^j_{ct} : Log of income per capita in 1990 dollars
- I^j_c : Technology-Country dummy
- I^j_t : Technology-Time dummy
- d_{ck} : distance between c and k
- We estimate,

$$x^j_{ct} = \beta^j_{1c} I^j_c + \beta^j_{2t} I^j_t + \beta^j_3 y_{ct} + \beta^j_4 x^j_{-ct} + \beta^j_5 y_{-ct} + \epsilon^j_{ct}$$

where

$$x^j_{-ct} = \sum_{\forall k \neq c} d_{ck} x^j_{kt} \text{ and } y_{-ct} = \sum_{\forall k \neq c} d_{ck} y_{kt}$$

- We refer to x^j_{-ct} and y_{-ct} as the spatial distance from technology (SDT) and income (SDI), respectively.

Other geographic interactions

$$x^j_{ct} = \beta^j_{1c} I^j_c + \beta^j_{2t} I^j_t + \beta^j_3 y_{ct} + \beta^j_4 x^j_{-ct} + \beta^j_5 y_{-ct} + \epsilon^j_{ct}$$

where

$$y_{-ct} = \sum_{\forall k \neq c} d_{ck} y_{kt}$$

- y^j_{-ct} is the spatial distance from income (SDI)
- Captures geographic interactions through channels other than technology
- Note gravity equation terms captured by FEs and income controls (because of logs).

Identification

$$x^j_{ct} = \beta^j_{1c} I^j_c + \beta^j_{2t} I^j_t + \beta^j_3 y_{ct} + \beta^j_4 x^j_{-ct} + \beta^j_5 y_{-ct} + \epsilon^j_{ct}$$

$$x^j_{-ct} = \sum_{\forall k \neq c} d_{ck} x^j_{kt}$$

- At a given year, x^j_{-ct} reflects geographic distance from adoption leaders.
- Over time, d_{ck} is fixed, and x^j_{-ct} reflects how technology diffuses in far vs. nearby countries.
- β^j_4 identified by how diffusion of technology in far vs. nearby countries affects diffusion of technology in c.

Endogeneity

- SDT introduces an endogeneity bias
  - Exogenous increases in x^j_{ct} increase SDT_{-c}
  - if $\beta^j_4 < 0$, SDT_{-c} reduces x_{-c}
  - Lower x_{-c} reduces SDT_c
- This bias is insignificant because:
  - There is a full distribution of adoption levels that enter in computing SDT_c
  - Under the null, ($\beta^j_4 = 0$) there is no bias and, in practice, β^j_4 is small.

Specifications

- Estimate

$$x_{ct}^{j} = \beta_{1c}^{j} I_{c}^{j} + \beta_{2t}^{j} I_{t}^{j} + \beta_{3}^{j} y_{ct} + \beta_{4}^{j} x_{-ct}^{j} + \beta_{5}^{j} y_{-ct} + \epsilon_{ct}^{j}$$

- Use various specifications that differ in whether coefficients (β_4^j and β_5^j) can vary across technologies/sectors.
  1. $\beta_4^j = \beta_4, \forall j$
  2. $\beta_5^j = \beta_5, \forall j$
  3. $\beta_4^j = \beta_4^s, \forall j \in s$
  4. $\beta_4^j = \beta_4^s, \forall j \in s$, and $\beta_5^j = \beta_5^s, \forall j \in s$

Pooled Regressions

	Specification			
	1	2	3	4
SDT	-.000171***	-.000126***	-.000109***	-.000080***
	($8.00e^{-6}$)	($6.82e^{-6}$)	($1.68e^{-5}$)	($1.30e^{-5}$)
SDI		.000659***		-.000300***
		($4.56e^{-5}$)		($7.30e^{-5}$)
SDT Com.			-.000089***	.000070***
			($1.90e^{-5}$)	($1.60e^{-5}$)
SDT Ind.			-.000053***	.000043**
			($2.80e^{-5}$)	($4.30e^{-5}$)
SDI Com.				.000770***
				($1.00e^{-4}$)
SDI Ind.				.000450***
				($1.40e^{-4}$)
# Obs.	53579	53579	53579	53579

Rich Countries

- Countries with income per capita greater than 8000 dollars in 1990

	Specification			
	1	2	3	4
SDT	-.000397*** ($3.80e^{-5}$)	-.000317*** ($3.37e^{-5}$)	-.000560*** ($6.74e^{-5}$)	-.000423*** ($5.46e^{-5}$)
SDI		.000459*** ($6.44e^{-5}$)		.000390*** ($1.02e^{-4}$)
SDT Com.			-.000175*** ($8.38e^{-5}$)	.000099*** ($7.20e^{-5}$)
SDT Ind.			-.000597*** ($1.36e^{-4}$)	.000469*** ($1.19e^{-4}$)
SDI Com.				.000145 ($1.44e^{-4}$)
SDI Ind.				-.000181 ($2.02e^{-4}$)
# Obs.	20151	20151	20151	20151

Poor Countries

- Countries with income per capita smaller than 8000 dollars in 1990

	Specification			
	1	2	3	4
SDT	-.00026***	-.000250***	-.000130***	-.000140***
	($1.40e^{-5}$)	($1.00e^{-5}$)	($2.90e^{-5}$)	($2.40e^{-5}$)
SDI		-.000510***		-.000810***
		($5.80e^{-5}$)		($9.00e^{-5}$)
SDT Com.			-.000190***	-.000150***
			($3.50e^{-5}$)	($2.80e^{-5}$)
SDT Ind.			-.000096**	-.000100***
			($4.80e^{-5}$)	($4.00e^{-5}$)
SDI Com.				.000600
				($1.40e^{-4}$)
SDI Ind.				.000250*
				($1.60e^{-4}$)
# Obs.	33428	33428	33428	33428

Decomposing source of Geographic Interaction

- Countries with income per capita smaller than 8000 dollars in 1990

	Specification	
	1	2
SDT Rich	-.000530***	-.000670***
	($1.50e^{-5}$)	($1.80e^{-5}$)
SDT Poor	-.000110***	-.000140***
	($5.12e^{-6}$)	($5.07e^{-6}$)
SDI Rich		.001530***
		($1.17e^{-4}$)
SDI Poor		-.000136*
		($4.00e^{-5}$)
# Obs.	53579	53579

Balanced Panel: Early Adopters

- Use the balanced panel of 15 countries (the early adopters) that are in the dataset throughout

	Specification			
	1	2	3	4
SDT	-.000700***	-.001100***	-.000850***	-.00160***
	($1.10e^{-4}$)	($1.00e^{-4}$)	($1.60e^{-4}$)	($1.30e^{-4}$)
SDI		.000480***		.000720***
		($5.90e^{-5}$)		($8.60e^{-5}$)
SDT Com.			.000270	.000870****
			($2.40e^{-4}$)	($2.10e^{-4}$)
SDT Ind.			.000600	.000140***
			($4.00e^{-4}$)	($4.00e^{-4}$)
SDI Com.				-.000510***
				($1.40e^{-4}$)
SDI Ind.				-.000510***
				($1.60e^{-4}$)
# Obs.	12540	12540	12540	12540

Latitude or longitude

- According to Diamond (1997):
  - Technology diffuses along latitudes (i.e., East-West dimension)
- Intuitive argument for why this happens for agricultural technologies
  - Why for non-agricultural technologies?
  - Maybe existing networks facilitate the diffusion of technologies
- Is this effect in the data?
  - Yes, much larger effect of distance across latitudes than across longitudes

Longitude and Latitude Distances

- Use longitude (East-West) or latitude (North-South) distance to calculate SDT and SDI

	Specification Longitude		Specification Latitude	
	1	2	1	2
SDT	-.000046***	-.000069***	-.000310***	-.000230***
	$(6.10e^{-6})$	$(4.91e^{-6})$	$(1.30e^{-5})$	$(1.20e^{-5})$
SDI		-.000027		-.000480***
		$(3.50e^{-5})$		$(7.00e^{-5})$
# Obs.	53579	53579	53579	53579

Longitude and Latitude Distances

- Use longitude (East-West) and latitude (North-South) distance to calculate SDT and SDI in the same regression

	Specification	
	1	2
SDT NS	-.005310***	-.006700***
	$(1.50e^{-5})$	$(1.80e^{-5})$
SDT EW	-.000110***	-.000140***
	$(5.10e^{-6})$	$(5.07e^{-6})$
SDI NS		.001500***
		$(1.20e^{-4})$
SDI EW		.000140***
		$(4.00e^{-5})$
# Obs.	53579	53579

The Simplest Model

- Borrow from literature on contagion as well as recent papers in growth (Lucas 2010 and Lucas and Moll (2011))
  - Simplified since the only goal is to the diffusion process
- Consider an economy where a mass N of agents are located uniformly in space. Space is given by the close interval $[0, 1]$.
- Let $G(0, \ell, t)$ denote the fraction of agents at location ℓ and time t that have not adopted the technology
  - The fraction of agents that have adopted is $G(1, \ell, t) = 1 - G(0, \ell, t)$
- Agents meet randomly with α agents per period
  - If an agent that has not yet adopted and meets with an agent that has adopted he adopts immediately
- The probability that an agent at location ℓ meets an agent at location r is $e^{-\delta|\ell - r|}$ times lower than the probability of meeting an agent that lives at ℓ

The Simplest Model

- Thus the probability of not adopting in period $t+h$ conditional on not having adopted in period t is given by

$$G(0,r,t+h) = G(0,\ell,t)\left[\frac{\int_0^1 G(0,\ell,t)\, e^{-\delta|\ell-r|}d\ell}{\int_0^1 e^{-\delta|\ell-r|}d\ell}\right]^{\alpha h}$$

which implies, taking limits as $h \to 0$ that

$$\frac{\partial \ln G(0,r,t)}{\partial t} = \alpha \ln\left(\int_0^1 G(0,\ell,t)\, e^{-\delta|\ell-r|}d\ell\right) - \alpha \ln\left(\int_0^1 e^{-\delta|\ell-r|}d\ell\right)$$

- The above equation implies that if $G(0,\ell,0) < 1$ for some $\ell \in [0,1]$, $G(0,\ell,t) < 1$ for all ℓ and $t > 0$, and $G(0,\ell,t)$ is increasing over time for all ℓ

The Location of Inventions

- Initial conditions

$$G(0,\ell,0) = \begin{cases} g < 1 \text{ for } \ell \in [0,a] \\ 1 \text{ otherwise} \end{cases}$$

$$\begin{aligned} \frac{\partial \ln G(0,r,0)}{\partial t} &= \alpha \ln \left(g \int_0^a e^{-\delta|\ell - r|} d\ell + \int_a^1 e^{-\delta|\ell - r|} d\ell \right) \\ -\alpha \ln \left(\int_0^1 e^{-\delta|\ell - r|} d\ell \right) &< 0 \text{ for all } r \end{aligned}$$

so for $r < r'$

$$\frac{\partial \ln G(0,r,0)}{\partial t} < \frac{\partial \ln G(0,r',0)}{\partial t}, \text{ and } \frac{\partial \ln G(0,r,t)}{\partial t} < \frac{\partial \ln G(0,r',t)}{\partial t}$$

- Since $G(0,r,0)$ is decreasing in r

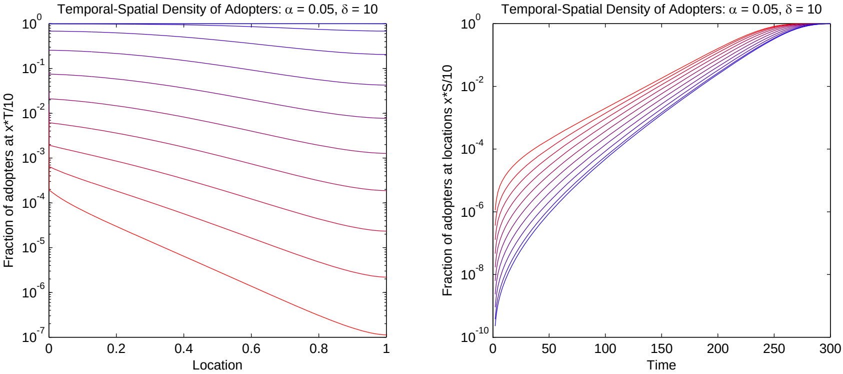
$$\frac{\partial \ln G(0,r,t)}{\partial r} > 0, \text{ for all } t.$$

The Location of Inventions

- Implication1: *The fraction of non-adopters is lower in locations closer to the source of innovation.*
- Implication 2: *The fraction of non-adopters declines at a higher rate in locations closer to the source of innovation.*
- Implication 3: The effect of distance on the level of adoption vanishes over time
  - Since in the limit all locations adopt fully so $G(1, \ell, t) = 1$ for all ℓ

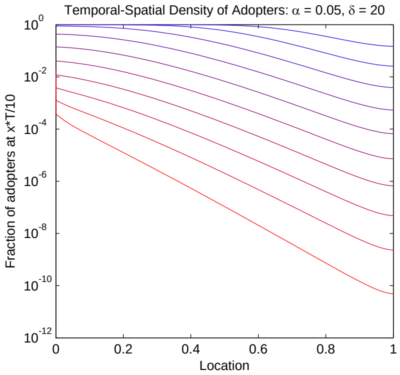
Some Examples

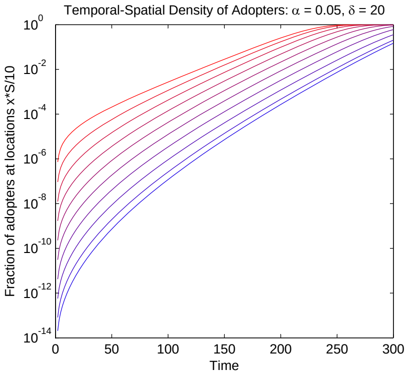
- Simulate the model for the case when $a = 1/1000$, $g = .99$, $\alpha = 0.05$, and $\delta = 10$



Some Examples

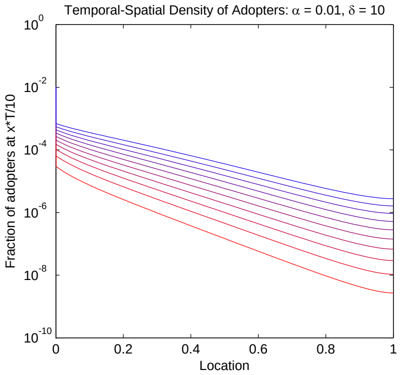
- Increase δ to $\delta = 20$. So agents far from each other meet less often

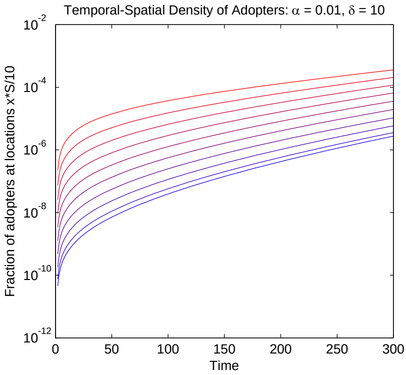




Some Examples

- Reuce α to $\alpha = 0.01$. So agents meet less frequently





Does the Effect of SDT Decrease over Time?

- Estimate

$$x_{ct}^j = \beta_{1c}^j I_c^j + \beta_{2t}^j I_t^j + \beta_3^j y_{ct} + \beta_{4t}^j x_{-ct}^j + \beta_{5t}^j y_{-ct} + \epsilon_{ct}^j$$

- Use the time series of α_{4t}^j for each technology j to estimate

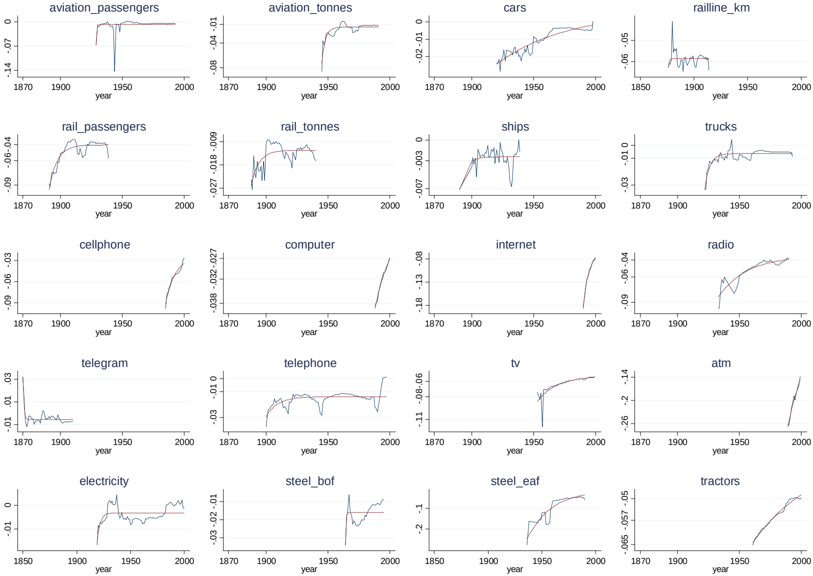
$$\beta_{4t}^j = c^j + e^{-b^j(t-t_0)}(a^j - c^j) + \tilde{\epsilon}_t^j$$

- Note the interpretation of the coefficients:
  - a^j : Determines the initial level of β_{4t}^j
    - Should be negative according to theory
  - b^j : Determines the rate of decline of β_{4t}^j
    - Should be positive according to theory
  - c_j : Determines the long run level of β_{4t}^j if b^j positive
- Estimate using balanced panel of 15 countries

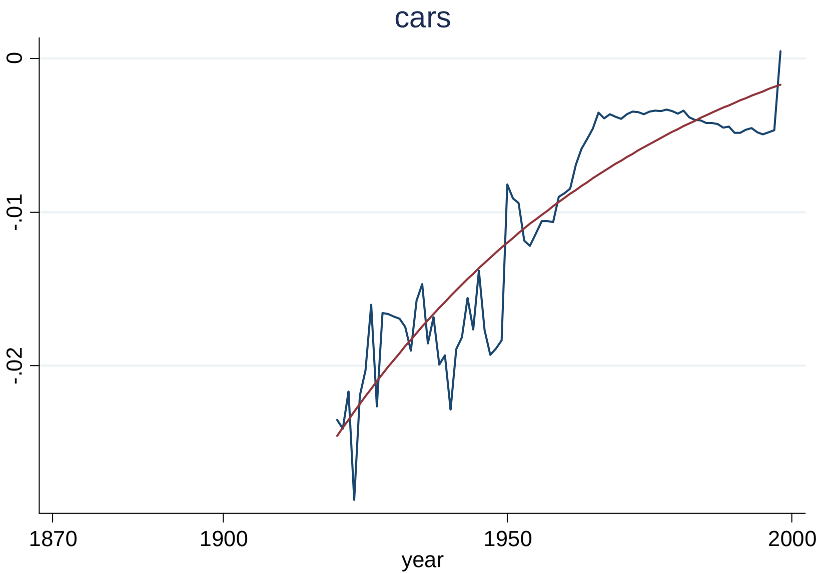
Results: Balanced Sample

Sector	Name	Invented	a^j	s.e. a^j	b^j	s.e. b^j	c^j	s.e. c^j	R^2	# Obs.
Transport	Aviation Passengers	1903	-0.0666	0.0181	3.1571	6.9464	-0.0070	0.0023	0.2748	65
	Aviation Tonnes	1903	-0.0723	0.0056	0.3367	0.0586	-0.0146	0.0011	0.9254	47
	Cars	1885	-0.0246	0.0011	0.0177	0.0049	0.0060	0.0046	0.9535	79
	Rail Line Km	1825	-0.0628	0.0039	1.1609	2.6294	-0.0586	0.0007	0.9959	39
	Rail Passengers	1825	-0.0962	0.0049	0.1556	0.0250	-0.0404	0.0014	0.9821	49
	Rail Tonnes	1825	-0.0258	0.0025	0.1554	0.0525	-0.0123	0.0007	0.9445	53
	Ships	1776	-0.0072	0.0014	0.1948	0.1232	-0.0024	0.0003	0.7981	41
	Trucks	1903	-0.0338	0.0023	0.2414	0.0339	-0.0063	0.0004	0.9177	72
Communication	Cellphone	1973	-0.0914	0.0035	0.1287	0.0354	-0.0244	0.0080	0.9949	16
	Computer	1973	-0.0391	0.0004	0.0667	0.0233	-0.0173	0.0053	0.9998	13
	Internet	1983	-0.1841	0.0019	0.2240	0.0165	-0.0659	0.0035	0.9998	11
	Radio	1920	-0.0837	0.0031	0.0439	0.0105	-0.0359	0.0041	0.9872	52
	Telegram	1835	0.0324	0.0030	1.3812	0.2852	-0.0054	0.0005	0.8683	41
	Telephone	1876	-0.0294	0.0030	0.1070	0.0346	-0.0138	0.0006	0.9131	95
	TV	1927	-0.0860	0.0041	0.0629	0.0232	-0.0528	0.0038	0.9898	46
Industry	ATM	1971	-0.2677	0.0069	0.0464	0.0497	0.0575	0.2780	0.9987	11
	Electricity	1882	-0.0155	0.0028	0.3644	0.1450	-0.0032	0.0004	0.6276	82
	Steel Bof	1950	-0.0340	0.0051	2.0971	2.2733	-0.0159	0.0010	0.9227	31
	Steel Eaf	1907	-0.2466	0.0181	0.0433	0.0100	-0.0121	0.0194	0.9438	47
	Tractors	1903	-0.0644	0.0003	0.0209	0.0043	-0.0364	0.0040	0.9999	40

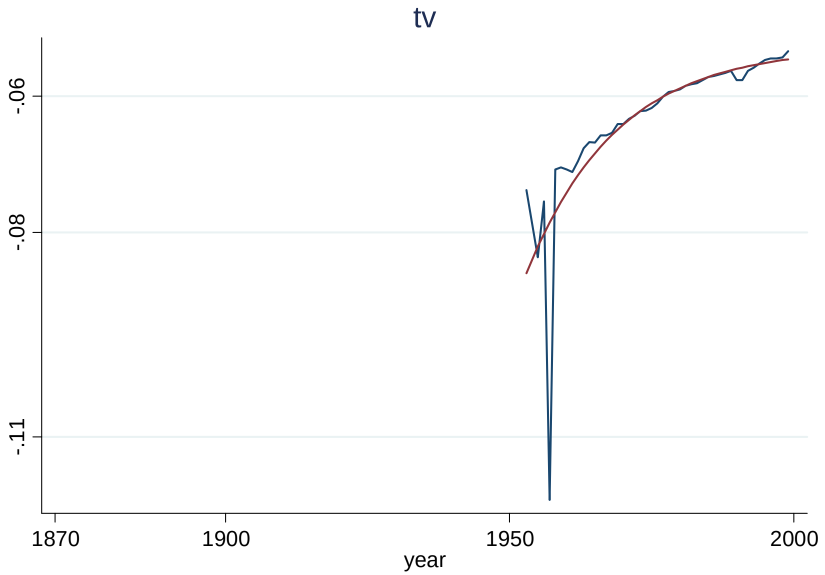
Results: Balanced Sample



Results: Cars



Results: TV



Matching model and data

- We estimate parameters of the model α and δ so that the model generates a pattern of interaction terms that comes as close as possible to observed β_{4t}^{j} in the data.
- For a given, α, δ :
  - Simulate Model
  - Compute SDT for model we compute

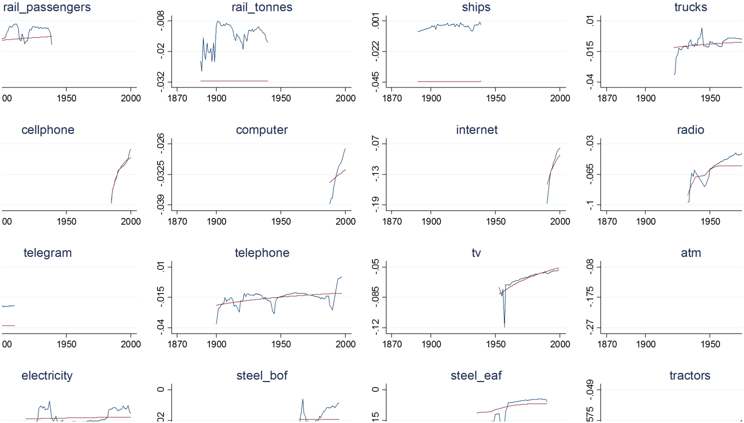
$$SDTM_{it}^{j} = \sum_{k \neq i} x_{it}^{j} d_{ik}$$

- Obtain Estimate the regression

$$x_{it}^{j} = I_{i}^{j} + I_{t}^{j} + \sum \tilde{\beta}_{t}^{j} SDIM_{it}^{j} + \epsilon_{it}$$

$$\underset{\alpha,\ \delta}{Min} \ \sum_{t} (\beta_{4t}^{j} - \tilde{\beta}_{t}^{j}(\alpha, \delta))^2$$

Results: Model



Results: Parameters

- Remember that α denotes the frequency of meetings and δ the spatial discount in the probability of meeting agents

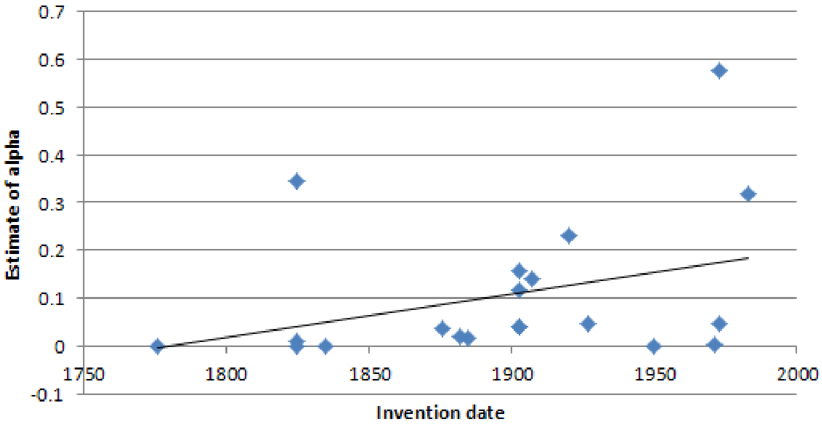
Sector	Name	Invented	α	δ	R^2
Transport	Aviation Passengers	1903	0.117	3.900	0.042
	Aviation Tonnes	1903	0.156	0.563	0.488
	Cars	1885	0.015	3.900	0.202
	Rail Line Km	1825	0.346	0.119	-26.2
	Rail Passengers	1825	0.008	3.900	0.113
	Rail Tonnes	1825	0.000	10.000	-13.9
	Ships	1776	0.000	3.500	-703.8
	Trucks	1903	0.040	0.626	0.270
Communication	Cellphone	1973	0.577	0.137	0.943
	Computer	1973	0.046	1.082	0.384
	Internet	1983	0.318	0.227	0.788
	Radio	1920	0.232	0.105	0.405
	Telegram	1835	0.000	2.700	-17.0
	Telephone	1876	0.036	0.662	0.269
	TV	1927	0.046	0.483	0.599
Industry	ATM	1971	0.001	0.100	-8.676
	Electricity	1882	0.018	0.666	0.050
	Steel Bof	1950	0.000	3.900	-0.226
	Steel Eaf	1907	0.139	3.498	0.362
	Tractors	1903	0.038	1.500	0.960

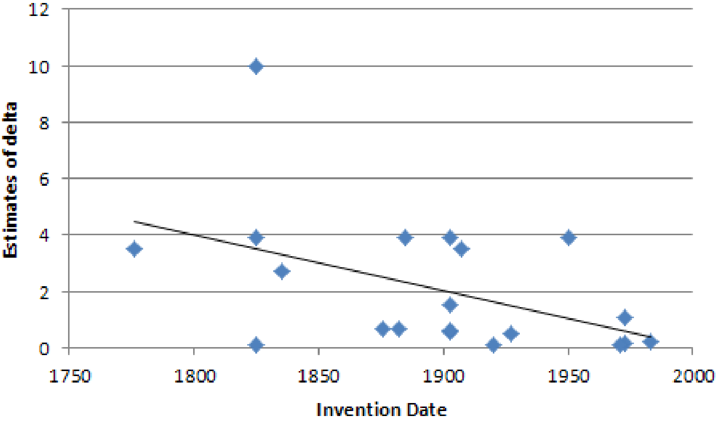
Intuition

- For 5 techs, $\tilde{\beta}_t^j$ are flat and below the series of β_{4t}^j observed in the data. Why?
- The geographic distribution of the followers is such that the SDT variable has a very small dispersion (relative to the dispersion in adoption).

$$x_{it}^j = I_i^j + I_t^j + \sum \tilde{\beta}_t^j SDIM_{it}^j + \epsilon_{it}$$

- As a result, the coefficients $\tilde{\beta}_t^j$ are higher (in absolute value) than the observed β_4^j.
- The α and δ that minimize the distance between $\tilde{\beta}_t^j$ and β_4^j are those that generate a series for $\tilde{\beta}_t^j$ that is less negative.
- This typically is the case when α is very low and δ is very high.


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
Estimate of alpha
1750
1800
1850
1900
1950
2000
Invention date



12
10
8
6
4
2
0
Estimates of delta
1750
1800
1850
1900
1950
2000
Invention Date

Discussion of Results

- α and δ vary substantially across technologies indicating very different speeds of diffusion and spatial discounting
    - On average, α is higher and δ is lower for newer technologies
    - α seems to be larger for newer and network-based technologies: Cellphone and Internet
    - δ is also low for these technologies.

Heterogeneity

- For those with location index $j \in [1, 7]$, we set the (log) of the share of initial adopters to

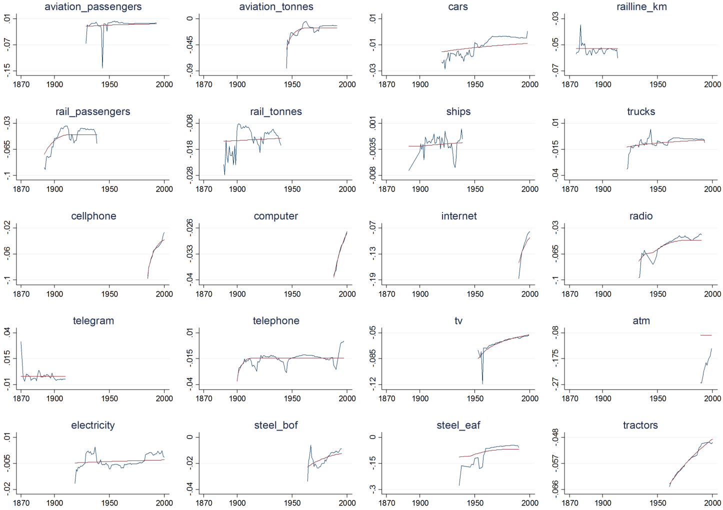
$$\log G_{k0}^{j} = \frac{\sum_i x_{i0}^{j}}{15} - \sigma_x + 2\sigma_x \frac{j-1}{6} - \bar{x}^{j},$$

  and for countries with index $j \in [9, 15]$, we set it to

$$\log G_{k0}^{j} = \frac{\sum_i x_{i0}^{j}}{15} - \sigma_x + 2\sigma_x \frac{15-j}{6} - \bar{x}^{j}.$$

σ_x is the standard deviation of initial adoption across followers.

Heterogeneity: results



Heterogeneity: results

Table 13: Structural Estimates of α, δ and σ with Initial Heterogeneity

Sector	Technology	Year	α	δ	σ
Transport	Aviation Passengers	1903	0.144	0.497	1.170
	Aviation Tonnes	1903	0.777	0.476	1.270
	Cars	1885	0.040	3.900	2.490
	Rail Line Km	1825	0.001	0.210	1.010
	Rail Passengers	1825	0.160	3.900	0.235
	Rail Tonnes	1825	0.008	0.396	1.610
	Ships	1776	0.003	0.097	2.540
	Trucks	1903	0.058	0.416	1.380
Communication	Cellphone	1973	0.577	0.137	0.000
	Computer	1973	0.309	0.418	2.520
	Internet	1983	0.318	0.227	0.000
	Radio	1920	0.121	1.160	0.280
	Telegram	1835	0.000	0.096	2.110
	Telephone	1876	0.679	0.010	1.640
	TV	1927	0.028	0.016	0.020
Industry	ATM	1971	0.001	0.100	0.000
	Electricity	1882	0.034	0.449	1.080
	Steel Bof	1950	0.070	0.145	2.400
	Steel Eaf	1907	0.139	3.500	0.000
	Tractors	1903	0.046	0.770	0.205

Conclusions

- This paper has provided evidence on the importance of geography for the diffusion of technologies
  - Using actual data for specific technologies, not only looking at the consequences of this diffusion
- We found very robust and pervasive patterns hard to account for OVB
- Patterns uncovered seem naturally explained by a model of spatial diffusion of technology
- Basic model fits data quite well
- Structural estimates imply that Spatial frictions have declined but are still quite important to explain cross-country differences in adoption.

Estimating the model's parameters

- The model considers percentage of adopters, while the data consider level of production in units per capita
- We can assign $\log G_{i0}^j = \log x_{i0}^j - \max_i \log x_{iT}^j$ where T is the last period observed in the data, since

$$\lim_{T\to\infty} \log G_{iT}^j = 0 \text{ and } \lim_{T\to\infty} x_{it}^j = \bar{x}^j$$

- We place the leader in the middle of the unit interval and assign to it

$$\max_i \log G_{i0}^j = \max_i x_{i0}^j - \bar{x}^j$$

all other countries are assigned the value

$$\log G_{k0}^j = \frac{\sum_i \log x_{i0}^j}{15} - \bar{x}^j$$

- Then we use our model of diffusion and generate a time-series of G_{it}^j and obtain predicted output in the model using

$$x_{it}^j = \log G_{it}^j + \bar{x}^j$$