The Spatial Diffusion of Technology*

Diego Comin
Harvard University

Mikhail Dmitriev
Boston College

Esteban Rossi-Hansberg
Princeton University

March 23, 2013

Abstract

We study technology diffusion across countries and over time empirically. We find significant evidence that technology diffuses slower to locations that are farther away from adoption leaders. This effect is stronger across rich countries and also when measuring distance along the south-north dimension. A simple theory of human interactions can account for these empirical findings. The theory suggests that the effect of distance should vanish over time, a hypothesis that we confirm in the data, and that distinguishes technology from other flows like goods or investments. We then structurally estimate the model. The parameter governing the frequency of interactions is larger for newer and network-based technologies and for the median technology the frequency of interactions decays by 73% every 1000 Kms. Overall, we document the significant role that geography plays in determining technology diffusion across countries.

1 Introduction

Technology disparities are critical to explain cross-country differences in per capita income.[1] Despite being non-rival in nature,[2] and involving no direct transport costs, technology diffuses slowly both across and within countries. These slow flows can result in significant lags between the time of invention and the time when a technology is initially used in a country. Even when a technology has arrived in a country, it takes years and even decades before it has diffused to the point of having a significant impact on productivity. These observations have led economists to study why does technology diffuse slowly, and what explains cross-country differences in its speed of diffusion.

*We thank Pol Antràs, Antonio Ciccone, Giancarlo Corsetti, Walker Hanlon, Stefania Garetto, Philippe Martin, Alex Monge, Francesc Ortega, Julio Rotemberg, Catherine Thomas, and seminar participants at various institutions for useful comments.

[1] See Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), Comin and Hobijn (2010) and Comin and Mestieri (2010) among many others.

[2] See, for example, Romer (1990). This property steams from the fact that once invented, the use of a technology by one producer does not preclude others from using it.

Existing empirical studies on technology adoption have treated adoption units (e.g. countries, cities, or firms) as independent.[3] Consequently, they have tried to link a country's technology adoption patterns to the country's characteristics (e.g. human capital, institutions, policies, adoption history, etc.).[4] This empirical approach to the drivers of technology adoption ignores the possibility of cross-country interactions in the adoption process. This assumption might be restrictive. Adopting a technology requires acquiring knowledge[5] which often comes from interactions with other agents. The frequency and success of these interactions is likely to be shaped by geography. Technological knowledge is likely to be more easily transmitted between agents in countries that are close than between agents located far apart. Similarly, the payoff to adopting a given technology (e.g. railways) may be affected by the adoption experience of neighboring countries. These mechanisms may generate correlated adoption patterns across nearby countries. To explore the empirical importance of these mechanisms, in this paper we explore the presence of cross-country interactions in technology adoption that are mediated by geographical distance. In particular, we study empirically the diffusion of technology across time and space.

A clear impediment to collecting evidence on the presence of geographic interactions in technology adoption is the lack of comprehensive datasets that directly document the diffusion of specific technologies across countries. In this paper we study the diffusion over time and space of 20 major technologies in 161 countries over the last 140 years using data from the CHAT dataset (Comin and Hobijn 2004, 2010). Our interest lies in exploring the presence of cross-country correlations in technology adoption that are mediated by geographical distance. To this end, we measure how far a country is from the high-density points in the distribution of technology adoption in the other countries. We denote this measure the spatial distance from other country's technology or, to abbreviate, the spatial distance from technology (SDT).[6] After controlling for country and time fixed effects, a negative correlation between SDT and adoption implies that countries that are further away from those where the technology diffuses faster tend to experience a slower adoption of the technology.[7]

In Section 2 we present a battery of empirical findings. We estimate a strong and significant negative partial correlation between SDT and a country's adoption, after controlling for per capita income and technology-specific country and time fixed effects. We also explore whether the spatial correlations in adoption are purely driven by the spatial correlation in income or in other variables strongly correlated with income. To this end, we control for a measure of the spatial distance from

[3] See, for example, Griliches (1957) and Mansfield (1961).

[4] See Comin and Hobijn (2004) and Comin, Easterly and Gong (2010).

[5] Potential users of the technology first learn about its existence and properties, then, they need to learn how to use it, and finally they need to figure out how to apply it to an existing production process or to new ones.

[6] Formally, the SDT of a country is defined as the scalar product of the (log) adoption levels in the rest of the countries and the distance to each of these countries.

[7] The construction of SDT might raise concerns of endogeneity, since adoption is a function of all other countries' adoption rates. In the next section we argue that these concerns are minor if we have many countries and so the contribution of any one country to the distribution of adoption is small. In Appendix B we also argue that, even with a small sample of countries, the upper bound of the bias generated by endogeneity is rather small and certainly irrelevant for our substantive results.

other countries per-capita income (SDI), constructed in a way parallel to SDT. We find that the sign and magnitude of SDT in our diffusion equations is unaffected by the inclusion of SDI. We also explore the robustness of the association between SDT and technology adoption to various specifications and samples. We find robust and significant effects of SDT on technology diffusion across sectors, technologies, and income levels, as well as when we use distinct country samples to compute SDT.

The correlated adoption patterns we document could result from the spatial correlation of other drivers of technology adoption which are independent of income.[8] We further explore this possibility in two robustness checks. First, we follow the diffusion literature and control for measures of the three most significant predictors of diffusion: human capital, trade openness, and institutions.[9] Introducing these controls does not affect the significance or magnitude of the association between SDT and adoption. Second, we show that although the correlation between SDT and adoption is robust across all the technologies in our sample, the correlation between technologies of the rank of countries in their speed of adoption is fairly low (around 0.2). Hence, a technology-specific pattern of omitted variables would be necessary to account for the geographic patterns of technology dynamics that we uncover. Even though it is hard to imagine what this set of omitted variables could be, we acknowledge that our methodology cannot rule out this possibility. Our goal is to describe, for what we believe is the first time, the spatial patterns of technology diffusion across countries and to make the case that these patterns can be parsimoniously rationalized by models of spatial technology diffusion.

In order to provide a richer characterization of the spatial patterns of technology diffusion, we also take on Jared Diamond's hypothesis that technologies diffuse along latitudes.[10] We construct measures of SDT based on latitude and longitude distances, and run a horse race between the two. We find that, consistent with Diamond (1997), SDT across latitude distances has a stronger association with technology adoption than SDT across longitude distances. This finding is remarkable since our sample does not include any technology where climatic reasons might suggest that distance across latitudes is a larger impediment for diffusion than distance across longitudes.

We explore further the mechanisms that drive these spatial diffusion patterns by developing a simple model which borrows from the literature on external effects and contagion[11], as well as recent papers that have emphasized the importance of individual knowledge exchanges for growth (e.g. Eaton and Kortum, 1999, Lucas, 2009, and Lucas and Moll, 2011). In particular, our stylized model is based on two key assumptions. First, technology is diffused through interactions between adopters and non-adopters. Second, interactions are random and more likely between agents located nearby. The model's implications are consistent with our empirical findings. The theory also guides us in exploring new dimensions of the data. In particular, it predicts that the geographic interactions in adoption, as measured by the effect of SDT, should diminish over time as technology diffuses.

[8] This possibility is related to the reflection problem emphasized by Maski (1993).

[9] Caselli and Coleman (2001), Comin and Hobijn (2004), Comin and Hobijn (2008).

[10] See "Guns, Germs and Steel," Diamond (1997).

[11] See, for example, Fujita and Thisse (2002) and the survey in Duranton and Puga (2004).

Going back to the data, we document that this implication holds in virtualy all the technologies and samples studied. Furthermore, we provide a parsimonious statistical characterization that fits well the time variation in the effect of SDT on technology adoption.

We then go a step further and use the time patterns of geographic interactions in adoption to structurally estimate the two parameters that characterize the model. Using a simulated method of moments (SMM) estimator where the initial geography of adoption matches the variance of the initial distribution of adoption in the data, we show that our simple model can generate time-varying interaction effects that closely resemble the data for nineteen out of the twenty technologies in the sample. Our estimates of the structural parameters of the model help us understand better the spatial diffusion process. In particular, they show that the frequency of interactions has been higher for newer than for older technologies and that spatial interactions decline by 73% every 1000 Kms.

The fact that the impact of distance on technology diffusion dies out over time in this very systematic way, distinguishes technology from other economic flows, like migration, trade, or FDI. These other flows have also been shown to decline with distance due to transport cost and other migration restrictions. However, in clear contrast with technology, for people, goods, and investment flows, the effect of distance does not dissipate over time. Once technology is diffused, distance does not matter because ideas and innovations only need to be conveyed to each individual once and can then be used repeatedly afterwards. This particular characteristic of technology, which distinguishes it from other flows, is very much present in the data, both in our purely empirical specification, and in the estimated structural model.

Despite the intuitive appeal of cross-country interactions in technology adoption, the literature has not been able to document its presence and to assess their contribution to the large cross-country differences in technology adoption. Some strands of the literature have explored the presence of knowledge spillovers associated with research and development activities. Broadly speaking, this approach has been pursued in two different ways. One has used patent citations data mostly within rich countries (Jaffe, Trajtenberg, and Henderson, 1993, Almeida, 1996, Thompson and Fox-Kean, 2005). Another has used cross-country data to study the effects of a country's R&D expenditures on other nearby countries' TFP (see Keller, 2004, for a comprehensive survey). However, innovation and adoption are distinct phenomena and it is unclear whether the knowledge and factors relevant to adopt a technology are related to those that matter for innovating. Furthermore, to explain cross-country differences in adoption it seems more appropriate to rely on cross-country spillovers than within-country spillovers. In addition, given the typical length of gestation lags, a positive correlation between Solow residuals and R&D expenditures may just reflect international cyclical co-movement rather than international technology diffusion. A final strand of the literature has studied adoption directly using micro-level data for simple agricultural technologies such as new crops or high yield seeds (e.g. Foster and Rosenzweig,1995, and Bandeira and Rasul, 2006). These studies have also found evidence of spatial correlations in adoption patterns across individuals.

We view our approach to identify the presence and strength of geographic interactions as complementary to approaches that use more detailed and specific data, for particular circumstances, that allows for somewhat cleaner identification. Those approaches are necessarily limited to study one (or at most a few) technologies which typically are very simple (e.g. fertilizers, seeds) and not representative of the continuum of technologies in the economy. Furthermore, studies based on one quasi-natural experiment are, by design, anecdotal and so inference to the wider economy, other periods, or other technologies remains a challenge. Finally, and perhaps most important for our purpose, micro data cannot identify the presence of cross-country interactions in technology adoption. Geographic interactions in adoption may take place along different channels depending on the distance between the adoption units. As a result, the forces identified at short distances might be distinct to the ones we uncover here.[12]

The rest of the paper is organized as follows. In Section 2, we present our non-structural empirical approach, the data, and the findings from our non-structural exploration. In Section 3 we develop our simple stylized model and derive some testable implications. We contrast them with the data in Section 4. In Section 5, we estimate the model structurally. Section 6 concludes.

2 Empirical exploration

In this section, we start our investigation of the role of geographic interactions on technology diffusion. We begin by imposing minimal structure in order to try to uncover general and robust patterns in the data.

Our empirical methodology is based on the following econometric specification:

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

The dependent variable, x^j_{ct}, is the level of adoption of a technology j in country c in year t. Technology adoption measures come from the cross-country historical adoption technology (CHAT) dataset (Comin and Hobijn, 2004, 2009, and 2010). To maximize the country representation of the sample, we focus on 20 major technologies, listed in Table 1.[13] Broadly speaking, the technologies studied belong to three sectors, transportation, communication, and industry. They cover, in an unbalanced way, technology diffusion in 161 countries going back until 1825. For each technology measure, (e.g. tons*kilometer transported by rail per capita), we take logarithms and add a technology-specific constant that ensures that x^j_{ct} is always positive.[14] Adding a constant is inconsequential for the dependent variable, but it is relevant for the interpretation of the SDT term that we discuss below.

Given the time series length for some of our technologies, they may eventually become dominated by newer technologies. Because our interest is on the phase in which technologies diffuse, we censor

[12] Micro estimates of the strength of geographic interactions in adoption (Rode and Weber, 2012) or in innovation (Kerr and Duke, 2012) find that the effect of interactions becomes negligible beyond very short distances (e.g. a couple of kilometers).

[13] We select the 20 technologies with observations for the largest number of countries and that are relevant for a variety of sectors.

[14] In particular, we add the minimum of x^j_{ct} along c and t, for the years used in the regressions.

the time series to eliminate the obsolescence phase. We achieve this by censoring the data once the adoption per capita of the leader (i.e. the country with higher adoption per capita) starts to decline.

Income affects the demand for the goods and services that embody new technologies. This mechanism is orthogonal to the forces we explore in this paper and we control for it by including log of domestic income per capita (y_{ct}) as an independent variable. Controlling for income also takes care of the potential effects of foreign business cycles on domestic technology adoption. This is the case, to the extent that international business cycles affect the domestic economy (only) through the domestic income level. Of course, we can control for domestic income because the dependent variable in our analysis is a direct measure of technology adoption (as opposed to something much closer to income such as TFP). We use Madison (2005) data to construct the (log) of per capita GDP (in 1990 dollars).

Table 1: List of Technologies

Sector	Name	Measure (Per 1000 People)	Range	# of Countries
Transportation	Aviation Passengers	Passenger*Kilometers	1929-1993	102
	Aviation Tons	Tons*Kilometers	1945-1991	101
	Cars	# of Cars	1920-1998	145
	Rail Line	# of Kilometers	1876-1914	67
	Rail Passengers	Passenger*Kilometers	1891-1939	47
	Rail Tons	Tons*Kilometers	1888-1940	52
	Ships	Tonnage (Motor & Steam)	1890-1939	29
	Trucks	# of Trucks	1922-1993	118
Communication	Cellphone	# of Users	1985-2000	142
	Computer	# of Devices	1988-2000	126
	Internet	# of People With Access	1990-2000	144
	Radio	# of Devices	1933-1990	126
	Telegram	# of Telegrams Sent	1870-1910	38
	Telephone	# Of Devices	1900-1997	148
	TV	# Of Devices	1953-1999	148
Industry	ATM	# of Machines	1989-1999	33
	Electricity	KWHr	1919-2000	144
	Steel From Blast Oxygen	Tons	1964-1995	52
	Steel From Electric-arc	Tons	1936-1990	57
	Tractors	# of Tractors	1961-2000	146

In all our regressions we include technology-country dummies, I_c^j, and technology-year dummies, I_t^j. Country dummies capture country-specific factors that affect technology diffusion and that are relatively constant over the time-span in which the technology diffuses. These might include geographical variables (e.g., remoteness, size of the country, density, ruggedness, climate,...), institutional variables (e.g. political regime, expropriation risk,...) or historical endowment (e.g., familiarity with related technologies, education system, ...). Note that we allow country dummies to differ across technologies to capture the possibly different effect that persistent factors have on different technologies.

The inclusion of time and country dummies affects the identification of the estimated coefficients. The country fixed effects imply that the estimates reflect correlations of the change in the dependent variable with the change in the adoption level, x_{ct}^j. That is, with the diffusion of the technology. Technology-year dummies remove the average evolution of the diffusion process for each technology which may vary across technologies for a variety of factors largely orthogonal to our analysis.[15] As a result, the estimated coefficients capture the differential effect on technology diffusion of the dependent variables in a country relative to the rest.

The centerpiece of our exploration of the presence of geographical interactions in adoption is the spatial distance from other countries' technology (SDT). Intuitively, SDT is just an interaction between the (log) of adoption in other countries and how distant they are. In principle, there are many different ways to construct these interactions. In Appendix A, we present several alternatives and show the robustness of the basic empirical findings to these various specifications of SDT. Our baseline way to compute the interaction between technology and distance is as the scalar product of a vector of (log) adoption levels in the other countries and the vector of distances (thousands of kilometers) to these other countries. Formally,

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

where d_{ck} is the distance between countries c and k.

Note that, when the number of countries is large, the vector of adoption measures in the rest of the world, x_{kt}^j, is almost the same across countries, and the cross-country variation in SDT comes from differences in the matrix of distances. Therefore, in the cross-section, SDT is highly correlated with the remoteness of the country. Because the matrix of distances is constant over time (other than due to changes in the sample composition) this direct effect of remoteness on adoption is captured by the technology-specific country fixed effects. Therefore it does not affect the identification of β_4^j.

Since the matrix of distances is constant over time, time variation in SDT is generated by the diffusion of technology (i.e., from x_{kt}^j). As technology diffuses, x_{-ct}^j increases slowly in countries located close to places where technology diffuses faster. Conversely, x_{-ct}^j increases faster in places that are farther from countries where technology diffuses faster. Therefore, if being close to adoption leaders is beneficial for the diffusion of technology, we should observe that x_{-ct}^j is negatively correlated with adoption, x_{ct}^j. This is the logic behind the identification of β_4^j. Note that, since we

[15] These may include the nature of the technology, its capital intensity, when the technology was invented (Comin and Hobijn, 2010), etc.

include time dummies, the identification of β_4^j comes from the relative change of SDT in countries that are close to adoption leaders vs. those that are far (not from absolute changes in SDT).[16]

Of course, geographic interactions may take place along variables other than technology. Trade is an obvious example. However, the terms that arise in standard gravity equations used in international trade[17] are all captured in the regressors included in (1) independently of our variable of interest x_{-ct}^j.[18] A literature in political science (e.g. Simmons et al., 2007) has also emphasized the international diffusion of institutions and markets.[19] These other forms of geographic interactions may, in principle, affect the adoption dynamics in a country. To increase our confidence that the geographic interactions we are identifying with SDT occur through technology and not through these alternatives mechanisms, we introduce another control that we call spatial distance from (other countries') income (SDI). SDI is defined in an analogous way to SDT but rather than computing it with other countries' adoption, we use other countries' (log) per capita income. Formally, SDI is defined as

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

The controls we add in equation (1), and below in Section 2.8, are a way of addressing, in an imperfect way, the reflection problem (see Maski, 1993) that arises in our specification. Of course, there might be other geographic interactions that affect the adoption dynamics in a country. Given the scope of our study in terms of number of countries, technologies, and time, it is extremely hard (and impossible given our data) to identify exogenous changes in SDT and their effect on adoption dynamics. Hence, we have to rely on the argument that it is hard for us to think of variables that affect diffusion, that are geographically correlated, that change over time according to the patterns we uncover below, and that are not correlated with income (and therefore captured by SDI). In Sections 2.8 and 4, we further explore the data to uncover new features of the relationship between SDT and diffusion. These features further constrain the set of potential omitted variables that can account for our findings by forcing them to be technology-specific and to present stringent time-varying patterns. Still, if such variables existed, they could be influencing our results and we could

[16] That is the reason why our findings are robust to various specifications of SDT as shown in Appendix A.

[17] See Anderson, 2004, and Anderson and van Wincoop, 2003, among many others.

[18] The gravity equation that has proven to be an accurate way to predict bilateral trade flows, $TR_{c,c'}$, between countries, is given by

$$TR_{c,c't} = \beta * \frac{Y_{ct} * Y_{c't}}{d_{c,c'}} * \epsilon_{c,c't}. \tag{2}$$

Taking logs and adding across all other countries, we obtain

$$tr_{ct} = \alpha + \beta_1 * y_{ct} + \beta_2 * \sum_{c'} y_{c't} + \beta_3 * \sum_{c'} d_{c,c'} + \epsilon_{ct}. \tag{3}$$

Note that the regressors in (3) are captured by the controls in (1). In particular, the log income term controls for the effect of own income, the country fixed effect controls for the distance term and, when there are many countries, the time dummies basically capture the average income of the other countries. Hence, SDT identifies effects distinct from standard gravity effects.

[19] We explicitly explore the relevance of this hypothesis for technology diffusion below.

be confounding the effect of diffusion with the effect of these other variables. In Sections 4 and 5 we show that a simple parsimonious model of technology diffusion can capture well the pattern we find in the data. This, we believe, lends further credibility to our interpretation of the technology diffusion patterns we document in the data.

It is important to be aware that there is a potential endogeneity bias concern in the estimation of the regression in (1). Specifically, adoption in country c enters in the construction of the SDT of the other countries. If SDT affects adoption (i.e. $\beta_4^j \neq 0$), then the adoption levels of the other countries will also be affected by adoption in c. But because SDT in c is computed using adoption in the other countries, it will indirectly be affected also by adoption in c. If in reality β_4^j is negative, the endogeneity of SDT is likely to introduce a negative bias in β_4^j. This is the case because a higher adoption in country c, x_{ct}^j, increases SDT in the other countries which, because β_4^j is negative, should result in smaller adoption, x_{kt}^t, which in turn leads to a smaller SDT for country c.

There are two reasons to believe that this bias is not a significant concern in practice. First, when the number of countries in the sample is large, the effect of a country's adoption on the other countries SDT is negligible. Second, under the null (i.e. $\beta_4^j = 0$) there is no endogeneity bias and so the standard test to reject the null is still valid. Still, in Appendix B we conduct some back-of-the-envelope calculations and show that even in regressions where we use smaller samples, the endogeneity bias generates less than 0.3% of the standard deviation in SDT and can account for less than 3% of the magnitude of the estimated coefficients. We conclude that the estimates of β_4^j reported below are not significantly affected by an endogeneity bias.

Table 2 presents descriptive statistics of the variables used in the empirical exploration. We report the standard deviations for the raw variables and also for the residuals after regressing each variable against a full set of technology-specific country and time dummies.

Table 2: Descriptive Statistics

	Standard Deviation	Residual after Removing Country and Time*Technology FE Standard Deviation
X (technology)	2.85	2.74
Y (income)	1.00	0.99
Distance Interaction	3246.45	2731.10
Income Interaction	416.29	263.62
Distance Interaction (latitude)	1461.37	1240.60
Distance Interaction (longitude)	3533.42	3130.42
Distance Interaction (abs. latitude)	943.21	767.37
# Obs.	53579	53579

We consider five possible specifications of (1) which differ on the restrictions imposed on parameters β_4^j and β_5^j. In Specification 0 we just set $\beta_5^j = 0$ for all j. In our baseline specification (i.e. Specification 1) β_4^j is the same across technologies while we allow β_5^j to differ across j's. In Specification 2 both β_4^j and β_5^j are the same across technologies. In Specification 3, we allow β_4^j to differ across sectors, though not within sectors, and β_5^j varies across technologies. Finally, in the fourth specification, both β_4^j and β_5^j differ across sectors but not across technologies in the same sector.

2.1 Empirical findings

Our empirical approach is flexible and we shall take advantage of this flexibility in several ways. We first investigate the presence and strength of geographic interactions in technology adoption in various sectors and country samples. We are also able to disentangle the nature of geographic interactions in adoption by decomposing the SDT variable along several dimensions.

Table 3: Pooled Regressions[20]

	Specification				
	0	1	2	3	4
SDT	*-.000147****	*-.000171****	*-.000126****	*-.000109****	*-.000080****
	($4.50e^{-6}$)	($8.00e^{-6}$)	($6.82e^{-6}$)	($1.68e^{-5}$)	($1.30e^{-5}$)
SDI		T.S.	.000659***	T.S.	-.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	53579

2.2 Pooled regressions

We start by running regression (1) in our full sample of countries. Table 3 reports the estimates of the effects of SDT on technology adoption in our four specifications. The column labeled Specification 1 reports the estimate of β_4^j in the first specification, where only the effect of SDT is constant across technologies. This is the most flexible specification. We find a negative, significant (at the 1% level) effect of SDT on a country's adoption. As discussed above, this suggests that countries

[20] Each column corresponds to one specification of regression (1). Specifications 1 and 3 allow for technology-specific SDI coefficients. Specification 0 is the only one that does not include the SDI controls.

that are far from adoption leaders tend to adopt new technologies more slowly than countries that are close. From the statistics in Table 2, it follows that the magnitude of this effect is economically relevant. In particular, a reduction of one standard deviation in SDT leads to an increase in adoption by 17% of one standard deviation.[21]

In the first column, labeled Specification 0, we report the estimate of β_4^j in a regression that does not include the SDI control (in Specification 1 SDI is included but it is allowed to vary by technology, we denote this by T.S., for 'technology specific', in all tables). Comparing the first two columns it seems clear that controlling for SDI does not reduce the estimate of geographic interactions in technology. Columns 3 through 5 show that the effect of SDT on technology adoption is robust across the four specifications we explore. Columns 3 and 4 explore the sectoral variation in geographic interactions in adoption. Transportation is the default option. Therefore, the coefficient of SDT for transportation technologies corresponds to the first row. The rows labeled "SDT Com." and "SDT Ind." report the differential coefficient of SDT for communication and industry technologies, respectively, relative to the coefficient for those in transportation. The ranking of the coefficients of SDT on diffusion across technologies is not robust. In Specification 3, the strongest effect is in communication technologies, while in Specification 4, the strongest effect of SDT on technology diffusion is found in transportation technologies.

2.3 The importance of geography for rich and poor countries

After showing the significance of geographic interactions in adoption dynamics, one may wonder whether their relevance is uniform across countries. To explore this question, we split the countries in our sample in two groups depending on whether in 1990 their income per capita was above or below 12000 dollars according to the estimates in Madison (2005).[22] Then we run regression (1) separately in both subsamples.

Tables 4 and 5 report the estimates for the sample of rich and poor countries, respectively. The effect of SDT on a country's adoption is significant both for rich and poor countries. Still, it is significantly larger for rich than for poor countries implying that distance from adoption leaders slows down adoption more in rich countries than in poor ones. There are also differences in the sectors where geographic interactions in adoption are most relevant. For poor countries, we observe stronger effects of SDT on a country's adoption of communication technologies. For rich countries, it depends on the specification. In Specification 3, the estimates of β_4^j are highest in industry while in the fourth specification they are highest in transportation.

[21] One potential concern with the results presented in Table 3 is that they depend on the units and specification of geography we use in constructing the SDT variable. The results above are robust to two alternative specifications. First, we can eliminate the time-technology dummies and add as a regressor $\sum_{\forall k \neq c} x_{kt}$. That is, on top of the interaction between technology adoption and distance, add the level of technology adoption in all other countries directly. This robustness check yields very similar results both in magnitude and significance. Similarly, we could measure distance as a proportion of the maximum distance and then use one minus this measure interacted with adoption. Again, the alternative specification yields significant results, although, as expected, with an opposite sign.

[22] With this split, there are 36 rich countries and 125 poor countries in our sample.

Table 4: Rich Countries[23]

	Specification			
	1	2	3	4
SDT	*-.000842**** ($6.61e^{-5}$)	*-.000750**** ($5.66e^{-5}$)	*-.001369**** ($1.09e^{-4}$)	*-.000992**** ($8.82e^{-4}$)
SDI	T.S.	.000294*** ($5.17e^{-5}$)	T.S.	.000407*** ($9.24e^{-4}$)
SDT Com.			*.000792**** ($1.39e^{-4}$)	*.000470**** ($1.23e^{-4}$)
SDT Ind.			*.001330**** ($3.30e^{-4}$)	*.001418**** ($2.68e^{-4}$)
SDI Com.				-.000220 ($1.23e^{-4}$)
SDI Ind.				-.000441** ($1.57e^{-4}$)
# Obs.	15098	15098	15098	15098

Table 5: Poor Countries[24]

	Specification			
	1	2	3	4
SDT	*-.000174**** ($1.19e^{-5}$)	*-.000165**** ($9.54e^{-6}$)	*-.000106**** ($2.49e^{-5}$)	*-.000106**** ($2.00e^{-5}$)
SDI	T.S.	-.000397*** ($5.76e^{-5}$)	T.S.	-.000627*** ($9.11e^{-5}$)
SDT Com.			*-.000098**** ($2.91e^{-5}$)	*-.000087**** ($2.36e^{-5}$)
SDT Ind.			*-.000045* ($3.95e^{-5}$)	*-.000065**** ($3.39e^{-5}$)
SDI Com.				.000557 ($1.38e^{-4}$)
SDI Ind.				.000160* ($1.68e^{-4}$)
# Obs.	38481	38481	38481	38481

[23] Estimates of β_4^j for countries with income in 1990 higher than 12000 USD. The first and third specifications allow for technology-specific SDI coefficients.

[24] Estimates of β_4^j for countries with income in 1990 lower than 12000 USD. The first and third specifications allow

2.4 Decomposing SDT

A natural next step consists in exploring whether the characteristics of the countries with whom a country interacts also matter. To investigate this possible dependence, we decompose the SDT variable in two parts. Namely,

$$x_{-ct}^{jRICH} = \sum_{\forall k \neq c \ \& \ k \in RICH} d_{ck} x_{kt}^{j},$$

and

$$x_{-ct}^{jPOOR} = \sum_{\forall k \neq c \ \& \ k \in POOR} d_{ck} x_{kt}^{j},$$

where x_{-ct}^{jRICH} captures the geographic interactions in adoption with rich countries and x_{-ct}^{jPOOR} the interactions with poor countries. Note that, for all countries, $x_{-ct}^{j} = x_{-ct}^{jRICH} + x_{-ct}^{jPOOR}$. The SDI variable can be decomposed in an analogous way.

Table 6 estimates regression (1) allowing for different coefficients in the rich and poor components of SDT. The data again speaks clearly. Adoption interactions with rich countries affect technology adoption between four and five times more than interactions with poor countries. Including a similar decomposition for SDI does not affect significantly the estimates of the effects of the two SDT terms. Since adoption leaders are rich countries, we interpret these findings as evidence that it is particularly detrimental to be far from adoption leaders.

Table 6: Rich and Poor Countries[25]

	Specification	
	1	2
SDT Rich	*-.000339**** ($5.72e^{-5}$)	*-.000303**** ($4.81e^{-5}$)
SDT Poor	*-.000180**** ($9.88e^{-6}$)	*-.000104**** ($7.92e^{-6}$)
SDI Rich	T.S.	-.000426* ($1.67e^{-4}$)
SDI Poor	T.S.	.000144*** ($3.18e^{-5}$)
# Obs.	53579	53579

for technology-specific SDI coefficients.

[25] In the first column SDI Rich and SDI Poor vary by technology. In the second column they are constant across technologies.

2.5 Longitude vs. latitude

Jared Diamond conjectured in his 1997 best-seller book "Guns, Germs and Steel" that, due to the specificity of crops to particular climates, technologies have diffused along a given latitude rather than across latitudes. Our simple econometric framework can be adapted to test Diamond's hypothesis. In particular, we can compute separate SDT variables using distances in the east-west dimension (SDT EW) and in the north-south dimension (SDT NS). Diamond's hypothesis is that distance along the north-south axis slows down technology diffusion more than distance along the east-west axis. Therefore, if Diamond's hypothesis is correct, we should observe a higher effect of SDT NS on adoption than of SDT EW.

We start by introducing separately the two SDT terms in Table 7. We find that the coefficient on SDT across latitudes (SDT NS) is higher than SDT across longitudes (SDT EW). In all regressions we include SDI terms that use the same measures of distance as the corresponding SDT term. The absolute and relative size of the effects of SDT NS and SDT EW on adoption is robust to whether the coefficient of the SDI terms varies or not across technologies.

Table 7: Longitude and Latitude Individually[26]

	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	T.S.	-.000027	T.S.	-.000480***
		($3.50e^{-5}$)		($7.00e^{-5}$)
# Obs.	53579	53579	53579	53579

In Table 8 we compare the spatial distance interactions across latitudes and longitudes. As hypothesized by Diamond, being in a distant latitude is a higher barrier to the diffusion of technologies than being in a distant longitude. The estimates imply that distance across latitudes slows down adoption approximately forty seven times more than distance across longitudes!

Confirming the Diamond hypothesis in a sample of technologies without any crops is somewhat surprising. Actually, other than tractors, our sample does not contain any agricultural technology. Clearly, Diamond's rationale for the greater importance of latitude distances for technology diffusion is not relevant for technologies such as cars or telephones which can work equally well at different latitudes or longitudes. Providing and testing alternatives explanations for this finding is beyond the scope of this paper. However, we can advance one hypothesis that may be worthwhile investigating in future work. Namely, the diffusion of early agricultural technologies could have created a series of networks and trade routes along latitudes that have been used since then for the diffusion of

[26] Each column corresponds to either Specification 1 or 2 of the regression in (1). SDT and SDI are computed using distance either along longitude (first two columns) or along latitude (third and fourth columns).

more modern technologies. Clearly, the empirical relevance of this hypothesis remains a topic for future research.

Table 8: Longitude and Latitude Simultaneously[27]

	Specification	
	1	2
SDT NS	*-.000414**** $(1.40e^{-5})$	*-.000298**** $(1.25e^{-5})$
SDT EW	*-.000068**** $(5.67e^{-6})$	*-.000077**** $(4.52e^{-6})$
SDI NS	T.S.	.000067*** $(3.79e^{-5})$
SDI EW	T.S.	.000057*** $(1.54e^{-5})$
# Obs.	53579	53579

2.6 Early adopters

We focus next on the countries that adopt each technology relatively early. For each technology, we define early adopters as the 15 countries with earliest data on adoption and track them until the end of the sample. Specifically, we limit the left hand side variable to observations from early adopters and we compute the SDT and SDI variables using only information from countries that are early adopter.[28] This exercise is relevant because, by design, the panel used in the estimation and in constructing the interactions variables is roughly balanced (there is still the possibility that a country drops from the sample, but this is not a significant concern in CHAT). Therefore, this exercise may provide reassurance that the geographic interactions in adoption we have uncovered are robust to controlling for the sample of countries considered. In addition, early adoption dynamics may be interesting in themselves.

Table 9 reports the estimates of (1) for early adopters. Qualitatively the results are the same as when studying the full sample. The coefficient of SDT is negative and significant, and it is largest for transportation technologies. However, there are significant quantitative differences between the estimates reported in Tables 3 and 9. The estimates of the geographic interactions in adoption for early adopters (with other early adopters) are four times larger than the equivalent effects for the full sample. This should not be surprising since early adopters are rich countries and we have

[27] SDT NS and SDI NS are computed using distance along latitudes. SDT EW and SDI EW are computed using distance along longitudes. In Column 1, SDI NS and SDI EW vary by technology.

[28] The smaller sample of countries makes the potential endogeneity problem raised above a more relevant concern. In Appendix B we calculate a bound on the effect of this endogeneity bias. We find that the true coefficient can be larger than the coefficient reported by only 0.000022. Given the magnitude of the estimated coefficients reported in Table 7, this proves the endogeneity bias essentially irrelevant in practice.

already established that (i) rich countries are more sensitive to geographic interactions and that (ii) geographic interactions with rich countries have a larger impact on a country's adoption.

Table 9: Early Adopters[29]

	Specification 1	2	3	4
SDT	*-.000700****	*-.001100****	*-.000850****	*-.001600****
	($1.10e^{-4}$)	($1.00e^{-4}$)	($1.60e^{-4}$)	($1.30e^{-4}$)
SDI	T.S.	.000480***	T.S.	.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

2.7 Robustness

So far, our exploration has identified geographic patterns of technology diffusion and has shown, with more to come in the rest of the paper, that these patterns are general to the extend that they hold for a large number of countries and technologies. However, as argued above, interpreting the source of these patterns is not obvious. Even though our findings are consistent with the presence of significant geographic interactions in adoption, the correlation between SDT and adoption might just reflect the omission of some relevant driver of adoption that presents the appropriate cross-country correlation. The shadow of this possibility is impossible to rule out completely. However, we can perform further exercises that inform us about the nature of the geographic diffusion patterns we have uncovered.

The first exercise consists of studying the robustness of the our findings to controlling for some of the variables that have been previously documented as drivers of adoption. Though the number of potential controls is unlimited, relatively few variables have been identified empirically as drivers of technology diffusion. The cross-country literature, by-and-large, has been centered on three

[29] Each column corresponds to a specification of the regression in (1) for the balanced sample of early adopters. SDT and SDI are computed only with early adopters. In Specifications 1 and 3, the coefficient of SDI can vary by technology.

variables: human capital, institutions, and trade openness.[30]

We introduce these drivers of technology, z_{ct}, in our regressions in two ways.[31] First, we include them as controls in regression (1) to allow for a direct effect of a country's level of z in its technology level. Note that the bias from omitting these controls may generate the estimates of β_4^j we estimate. In addition, the return to adopting a technology may be affected by its neighbors level of z (e.g. the more open they are the higher the returns to adoption). To control for this possibility, we construct spacial distance from other countries z's for the three controls (human capital, openness, and democracy) in the same way we have constructed SDT and SDI.

Table 10: Early Adopters With Additional Controls, Education[32]

	Specification			
	1	2	3	4
SDT	-.001366***	-.002378***	-.001408***	-.002405***
	$(2.00e^{-4})$	$(1.85e^{-4})$	$(2.11e^{-4})$	$(1.91e^{-4})$
SDI	T.S.	.001084***	T.S.	.00107***
		$(1.10e^{-4})$		$(1.28e^{-4})$
Educ			T.S.	.043091
				$(3.63e^{-2})$
SD_Educ			T.S.	.000161
				$(3.49e^{-4})$
# Obs.	8203	8203	8203	8203

Table 10 through 12 present the results from introducing these controls one at a time and Table 13 introduces all three at once. The main finding is that the importance of SDT in regression (1) does not diminish by controlling for other drivers of technology or for the spatial distance from other drivers. The only control that reduces somewhat the estimate of β_4^j is democracy, in the specifications where we allow the coefficient of democracy and SD-democracy to vary by technology (see Table 11). On average, the coefficient of both democracy and SD-democracy are positive suggesting that competition in the political system favors technology diffusion but that a country does not benefit from being close to other democracies. Human capital and trade are

[30] See, for example, Caselli and Coleman (2001) for human capital, Comin and Hobijn (2008) for institutions, Coe and Helpman (1995) for trade and Comin and Hobijn (2004) for all three.

[31] The data on democracy comes from Polity IV. In particular we use the "Polity2" variable which measures the degree of competition in the political system. Data on trade openness comes from the Penn World Tables 7. in particular we use the variable "openc" which measures the share of exports plus imports over GDP. The data on human capital comes from CHAT which contains data on enrollment rates in secondary education. The data on human capital and democracy goes back to the beginning of the 20^{th} century, while the data on openness starts in 1960.

[32] Each column corresponds to a specification of the regression in (1) for the balanced sample of early adopters. Educ denotes secondary enrollment rate. SDT, SDI and SD_Educ are computed only with early adopters. In Specifications 1 and 3, the coefficient of SDI can vary by technology. In Specification 3, Sd_Educ can vary by technology.

insignificant because their effects on adoption are captured by the (time and country) fixed effects and income. In any event, the robustness of the estimate of β_4^j to controlling for the most-studied drivers of technology adoption in the literature suggests that SDT is introducing a new source of variation in diffusion dynamics.

Table 11: Early Adopters With Additional Controls, Polity[33]

	Specification			
	1	2	3	4
SDT	-.000626***	-.001115***	-.000306***	-.000953***
	$(1.15e^{-4})$	$(1.00e^{-4})$	$(1.19e^{-4})$	$(1.01e^{-4})$
SDI	T.S.	.000460***	T.S.	.000380***
		$(6.07e^{-5})$		$(6.07e^{-5})$
Polity			T.S.	.015291***
				$(1.23e^{-3})$
SD_Polity			T.S.	.000037***
				$(1.09e^{-5})$
# Obs.	11880	11880	11880	11880

Table 12: Early Adopters With Additional Controls, Openness[34]

	Specification			
	1	2	3	4
SDT	-.001427***	-.001626***	-.001557***	-.001604***
	$(1.63e^{-4})$	$(1.43e^{-4})$	$(1.91e^{-4})$	$(1.44e^{-4})$
SDI	T.S.	.000534***	T.S.	.000745***
		$(6.35e^{-5})$		$(1.29e^{-4})$
Open			T.S.	-.001322
				$(5.76e^{-4})$
SD_Open			T.S.	$(-3.81e^{-6}$
				$(1.85e^{-6})$
# Obs.	7085	7085	7085	7085

[33] Each column corresponds to a specification of the regression in (1) for the balanced sample of early adopters. Polity denotes Polity 2 from Polity IV, a measure of the degree of competition in teh political system. SDT, SDI and SD_Polity are computed only with early adopters. In Specifications 1 and 3, the coefficient of SDI can vary by technology. In Specification 3, Sd_Polity can vary by technology.

[34] Each column corresponds to a specification of the regression in (1) for the balanced sample of early adopters. Open denotes the share of nominal exports plus imports in GDP. SDT, SDI and SD_Open are computed only with early adopters. In Specifications 1 and 3, the coefficient of SDI can vary by technology. In Specification 3, Sd_Open

The relationship between SDT and adoption is pervasive. Throughout the paper we show that the patterns we identify hold for many diverse technologies and for a variety of country subsamples. How likely are these patterns to arise from the omission of a relevant driver of adoption in regression (1)? The answer depends on whether the identity of the countries where technology diffuses faster and more slowly is the same across technologies or not. If the rankings of countries based on their speed of diffusion are the same across technologies, the estimates of β_4^j we have observed could be generated by a country-level (omitted) variable that presents the appropriate correlation with diffusion, the appropriate cross-country correlation, and which is not correlated with our controls. Instead, if the country rankings in terms of the speed of diffusion differs significantly across technologies, then the required pattern of omitted variables needs to also be technology-specific. Given the number of technologies we study, the odds of this being the case are significantly lower.

Table 13: Early Adopters With Additional Controls

	Specification 1	2	3	4
SDT	-.002829*** ($2.89e^{-4}$)	-.003215*** ($2.47e^{-4}$)	-.002256*** ($3.48e^{-4}$)	-.002554*** ($2.71e^{-4}$)
SDI	T.S.	.001141*** ($1.23e^{-4}$)	T.S.	.001833*** ($2.34e^{-4}$)
Educ			T.S.	T.S.
Polity			T.S.	T.S.
Open			T.S.	T.S.
SD_Educ			T.S.	.001888*** ($3.45e^{-4}$)
SD_Polity			T.S.	.0001566*** ($2.33e^{-5}$)
SD_Open			T.S.	-.000016*** ($3.83e^{-6}$)
# Obs.	4219	4219	4219	4219

A few examples may help illustrate the nature of our findings. Take the case of aviation cargo. The countries where our measure of adoption has increased the most over the period studied are Netherlands, Switzerland, Belgium and France. Clearly, these countries are geographically close to others where the technology has diffused fast. Two of the countries where it has diffused most slowly are Brazil and Australia which, in the balanced sample, are among the furthest from the fast diffusion area. Tractors, whose sample covers mostly developing economies over the period 1961-2000, also diffused slowly in Australia, as well as in Argentina and Bolivia, while, over this period, they

can vary by technology.

diffused faster in sub-saharan Africa (Burkina Faso, Botswana, Angola). These examples illustrate the correlation between a neighbor's diffusion experience and a country's own experience. But they beg the following question. Does technology systematically diffuse more slowly in Australia? The answer is negative. Raillines is a technology that diffuses relatively fast over the period studied (1876-1914) in Australia. Interestingly, Australia is close to some of the countries in our sample where raillines diffused fast over this period (e.g. Japan and Sri Lanka) and far from others where it diffused more slowly (e.g. Egypt and UK).

To investigate more systematically adoption experiences across technologies, for each technology we rank countries according to the increment in adoption between the first and last year in the balanced sample. Then, we compute the pairwise correlation between country rankings for each pair of technologies. The average number of countries that coincide in a pair of technology rankings is 7. That is, on average, there are 8 non-coincident countries when comparing the samples for a pair of technologies. The average correlation among the country rankings between technology pairs is 0.19. When we weight the pairwise correlations by the number of observations the average is only 0.22. These low correlations suggest that countries with faster diffusion for a particular technology might experience slower diffusion for others.[35] Hence, a technology-specific pattern of omitted variables is needed to account for the negative correlation between SDT and adoption that we document. This finding, we believe, significantly raises the bar for the omitted variable explanation of our results, and supports our claim that there are important geographic interactions in adoption.

3 The simplest model

3.1 Description

We now present a simple mechanical model to represent and analyze the forces we have uncovered so far. A model can also help us parametrize the effects we find in the data and can point to some new hypothesis to test. For these purposes we want to propose the simplest theory of human interactions that can accommodate the temporal and geographic effects that are present in the data. The theory we propose is a theory of social interactions in which agents meet randomly with other agents and adopt new technologies if the agents they meet have already adopted (similar to the mechanism in Eaton and Kortum, 1999, and Lucas, 2009). We also assume that agents meet more frequently agents that are close by. The model specifies stochastically who do agents meet and when do they adopt new technologies. Agents make no decisions. The result is a mechanical, mathematical description of adoption rates over time and space. Adoption dynamics are governed by the rate of meetings among agents (α) and the decay in the meeting probability over space (δ). These are the two key parameters we estimate for each technology in Section 5.

[35] Note that this is only the case for the *change* in adoption. For the level of adoption Comin, Hobijn and Rovito (2006) finds a much larger correlation.

3.2 Formalization

Consider an economy where a mass N of agents are located uniformly in space. Space is given by the closed interval $[0,1]$. Time, $t = 0, 1, ...$ is discrete. We consider the diffusion of a technology that is first adopted at time $t = 0$. 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, therefore, given by $G(1,\ell,t) = 1 - G(0,\ell,t)$.

Agents meet randomly with α agents per period. We assume that the new technology strictly dominates the old one, so if an agent meets someone that has adopted the new technology already, he adopts immediately. A meeting between two agents that have not adopted does not lead to any technology upgrading. The parameter α governs the frequency of meetings and therefore determines the speed of technology adoption.

Agents meet more frequently with agents that locate close to them. In particular, 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 parameter δ governs the importance of space for technology adoption. A high δ implies that agents meet with agents far away from them very infrequently and therefore that diffusion is very localized.

The probability of not adopting in period $t+h$ conditional on not having adopted in period t at location r is then given by

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

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

$$\frac{\partial \ln G(0,r,t)}{\partial t} = \alpha \ln\left(\int_0^1 G(0,\ell,t)\, e^{-\delta|\ell-r|}d\ell\right) - \alpha \ln\left(\int_0^1 e^{-\delta|\ell-r|}d\ell\right). \tag{4}$$

The above equation implies that if $G(0,\ell,0) < 1$ for some interval of positive Lebesgue measure $L \in [0,1]$, $G(0,\ell,t) < 1$ for all ℓ and t and $G(0,\ell,t)$ is increasing over time for all ℓ. That is, if a non-trivial number of agents adopted in period $t = 0$, then the technology diffuses to all locations and adoption increases over time at all locations.

The effect of geography enters the model only through the distribution of the first adopters, $G(0,\cdot,0)$. To illustrate this, consider an example without geography where $G(0,\ell,0) = g < 1$ for all ℓ. So the same fraction of agents in all locations have already adopted at time zero. Then $\partial \ln G(0,\ell,t)/\partial t = \alpha \ln G(0,\ell,t)$ so $\partial \ln G(0,\ell,0)/\partial t = \alpha \ln g$. One can then guess and verify that the solution of the differential equation is given by $G(0,\ell,t) = e^{e^{\alpha t}\ln g} = g^{e^{\alpha t}}$.[36] In this example

[36] Guess that $G(0,\ell,t) = e^{\lambda(t)\ln g}$, so $\ln G(0,\ell,t) = \lambda(t)\ln g$ and so

$$\frac{\partial \ln G(0,\ell,t)}{\partial t} = \alpha \frac{\partial \lambda(t)}{\partial t}\ln g.$$

Hence, $\lambda(t) = \alpha \partial \lambda(t)/\partial t$ and so $\lambda(t) = e^{\alpha t}$. This implies that $G(0,\ell,t) = e^{e^{\alpha t}\ln g} = g^{e^{\alpha t}}$.

space plays no role. Technology diffuses slowly and uniformly and eventually all agents adopt, since $\lim_{t\to\infty} G(0,\ell,t) = \lim_{t\to\infty} g^{e^{\alpha t}} = 0$.

The example above eliminates the importance of space using two assumptions. First, assuming that the number of meetings is independent of the location (α is constant). An assumption we will maintain throughout. Second, it assumes that the density of adoption at $t = 0$ is uniform. This second assumption is unrealistic and should be modified. Initial adoption is in general concentrated geographically. For example, it is probably concentrated close to the inventor of the new technology. Therefore, a natural way to add geography is to add heterogeneity in initial conditions. In Section 5 we do this by estimating the variance of initial adopters from diffusion dynamics and then comparing this variance with the one we observe in the data. However, to illustrate the implications of the model, the simplest way is to start with an interval of locations that adopts initially, while all other areas start with no adoption whatsoever. Formally, the initial conditions now are

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

The resulting dynamics are more complicated than before and cannot be fully solved analytically. However, since $g < 1$,

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

for all ℓ and so for $a < \ell < \ell'$, $\partial \ln G(0,\ell,0)/\partial t < \partial \ln G(0,\ell',0)/\partial t$. Since $G(0,\ell,0)$ is decreasing in ℓ, this implies that $\partial \ln G(0,\ell,t)/\partial t < \partial \ln G(0,\ell',t)/\partial t$, and thus[37]

$$\frac{\partial^2 \ln G(0,\ell,t)}{\partial t \partial \ell} > 0, \text{ for all } t \text{ and all } \ell > a.$$

The previous arguments result in the following two implications:

Implication 1: *The fraction of non-adopters is lower in locations closer to the source of innovation.*

Implication 2: *The fraction of non-adopters declines proportionally faster in locations closer to the source of innovation.*

Since this process implies that in the limit all locations adopt fully so $G(1,\ell,t) = 1$ for all ℓ, we can also conclude that:

Implication 3: *The effect of distance on the level of adoption vanishes over time.*

The parameters α and δ affect the growth in the fraction of adopters as well as their level. It is easy to conclude from equation (4) that $\frac{\partial^2 \ln G(0,\ell,t)}{\partial t \partial \alpha} < 0$. Therefore, the larger α the faster adoption grows over time. It is harder to draw analytically other conclusions on the effects of α and δ on

[37] The restriction that $\ell > a$ is imposed to avoid boundary effects within the interval $[0,a]$. Of course, if the interval of initial adopters was in the middle of the interval $[0,1]$, then the same implications would apply within the interval of initial adopters as we move away from the center.

the evolution of adoption. However, we can illustrate them with the help of numerical examples. Figures 1 to 3 show three examples with $\alpha = 0.05$ and $\delta = 10$, $\alpha = 0.05$ and $\delta = 20$, and $\alpha = 0.01$ and $\delta = 10$.[38] The left panel represents the fraction of adopters over space in 10 different time periods, with equal intervals between them. The right panel represents the density of adopters over time for 10 points in space (again, equally spaced).

The results are clear, intuitively, and expected: First, the density of adopters decreases as we move away from ℓ, and the slope (in logs) decreases with time. The slope increases in absolute value with δ. Second, in all locations adoption increases monotonically over time, with the fraction of non-adopters falling proportionally slower in locations farther away from the initial innovation (in the examples $\ell = 0$). Finally, the growth rate of adoption increases with the number of meetings per period, α.

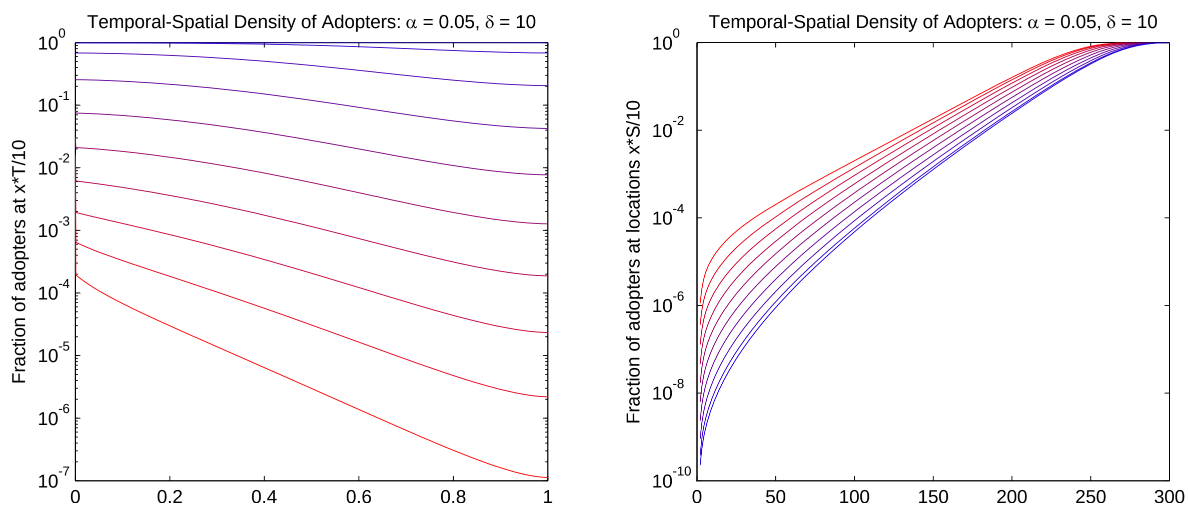

Figure 1: $\alpha = 0.05$ and $\delta = 10$

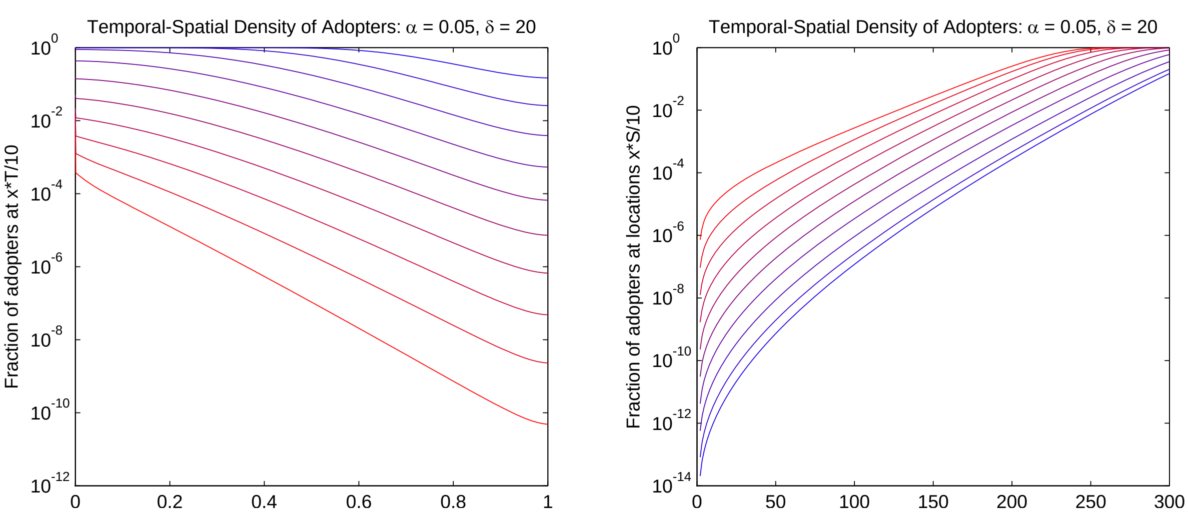

Figure 2: $\alpha = 0.05$ and $\delta = 20$

[38] We simulate the model for the case where $a = 1/1000$, $g = .99$ and two levels of α and δ. The fraction of adopters is plotted in log scale. We use a grid of 1000 points for space and run the model over 300 periods.

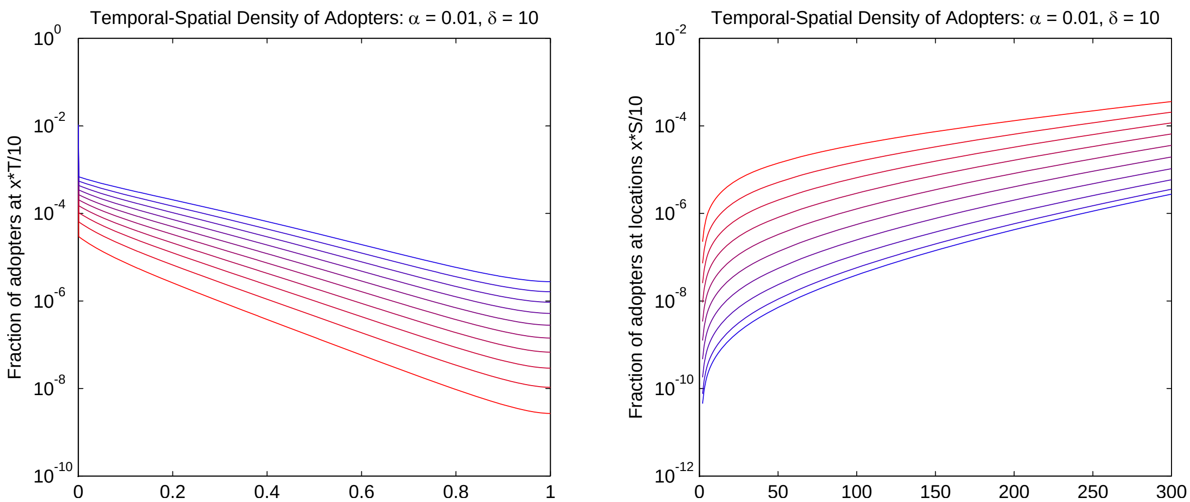

Figure 3: $\alpha = 0.01$ and $\delta = 10$

The regressions presented in Section 2 show that Implications 1 and 2 are consistent with the data. In particular we found that the coefficient on SDT in (1) is negative and significant. We now proceed to contrast the other prediction. In particular, we are interested in Implication 3, which tells us that the effect of SDT on adoption should vanish over time. This prediction is specific to the fact that agents that have adopted a technology can use it repeatedly in the future, and do not require any future interactions with adopters. In our view, this is a fundamental feature that distinguishes technological flows from migration, trade, or FDI flows.

4 Exploring the model's predictions

To contrast Implication 3 with the data we proceed in two steps. First, we modify the specification in (1) to allow for time varying coefficients of β_4^j and β_5^j. The new specification is given by

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

In a second step, for each technology, we take the series of estimates of β_{4t}^j, and fit them to the following three-parameter non-linear specification

$$\beta_{4t}^j = c^j + e^{-b^j(t-t_0)}(a^j - c^j) + \tilde{\epsilon}_t^j, \tag{6}$$

where $\tilde{\epsilon}_t^j$ is a residual, and t_0 is the initial adoption year. The parameter a^j determines the initial level of β_{4t}^j, and, according to our model, it should be negative. The parameter b^j determines the rate of increase of β_{4t}^j, and should be positive according to our theory. When b^j is positive, c_j is the long run level of β_{4t}^j.

We apply this two-stage procedure both for the balanced (15 countries) and unbalanced (161 countries) samples. Figures 4 and 5 plot, for each technology, the estimates of β_4^j in the unbalanced (Figure 4) and balanced (Figure 5) samples together with the fitted curves from (6). The first observation is that, for a large majority of the technologies, the model predictions are borne by the data. In particular, in 13 out of 20 technologies in the unbalanced sample, and in 19 out of 20 in the balanced sample, we observe estimates of β_{4t}^{j} that are initially negative and increase over time.

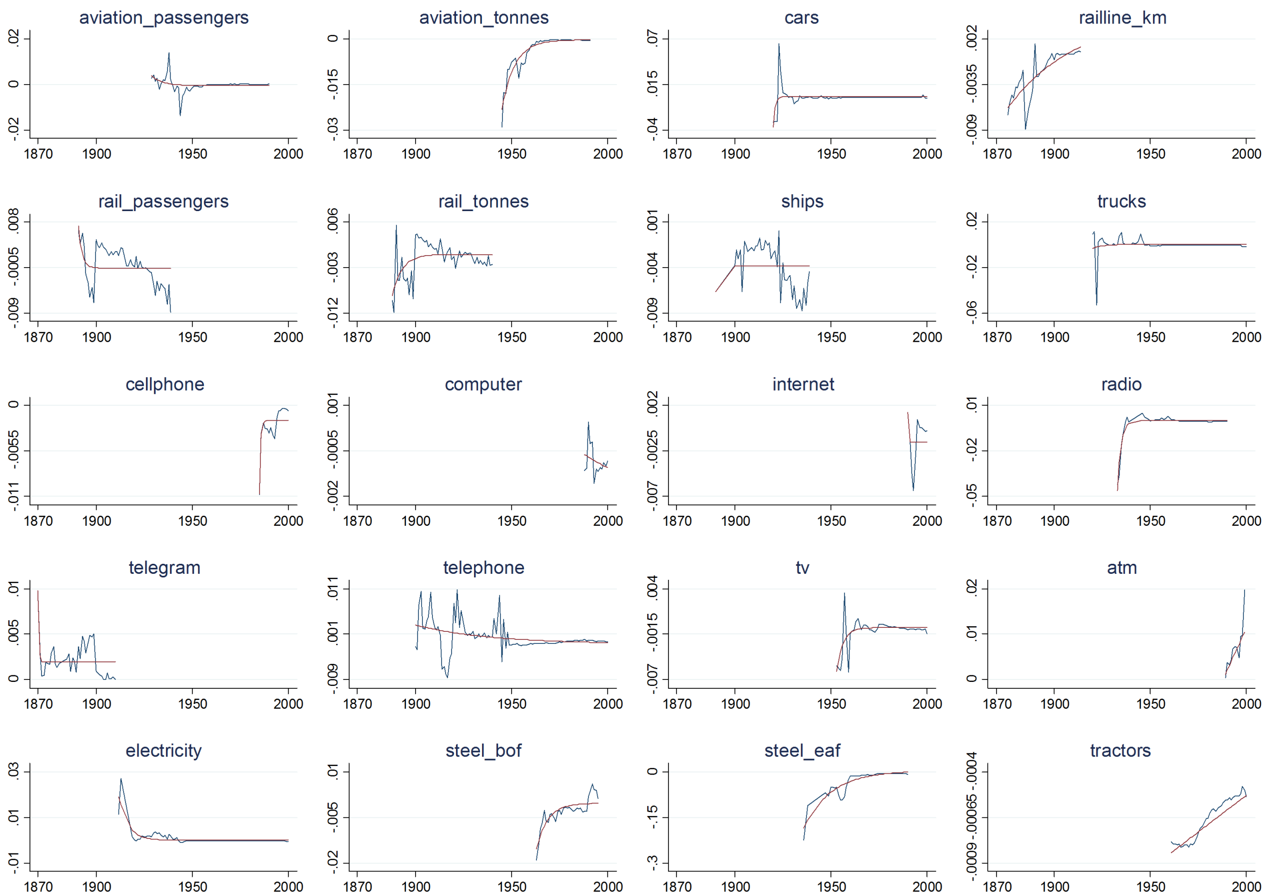

Figure 4: Estimates of β_{4t}^{j} in the unbalanced sample; fitted lines from the regression in (6)

Tables 14 and 15 report, for each technology, the estimates and standard errors of a^j, b^j and c^j, together with the R^2 of regression (6) for the unbalanced (Table 14) and balanced (Table 15) samples. The tables also report the year of invention of the technology. The point estimates confirm that the data conforms to the model's predictions. The point estimates of a large majority of technologies in both the unbalanced and balanced samples have negative estimates of a^j, positive estimates of b^j and estimates of c^j that often are close to zero and are almost always smaller (in absolute value) than the point estimates of a^j.[39]

[39] The time-varying pattern of the estimates of β_4^j that we uncover contrasts with the evidence from gravity equations that the elasticity of trade with distance has increased over time (Head and Mayer, 2011). This further suggests that the geographic interactions in adoption we are identifying are distinct from traditional geographic trade interactions.

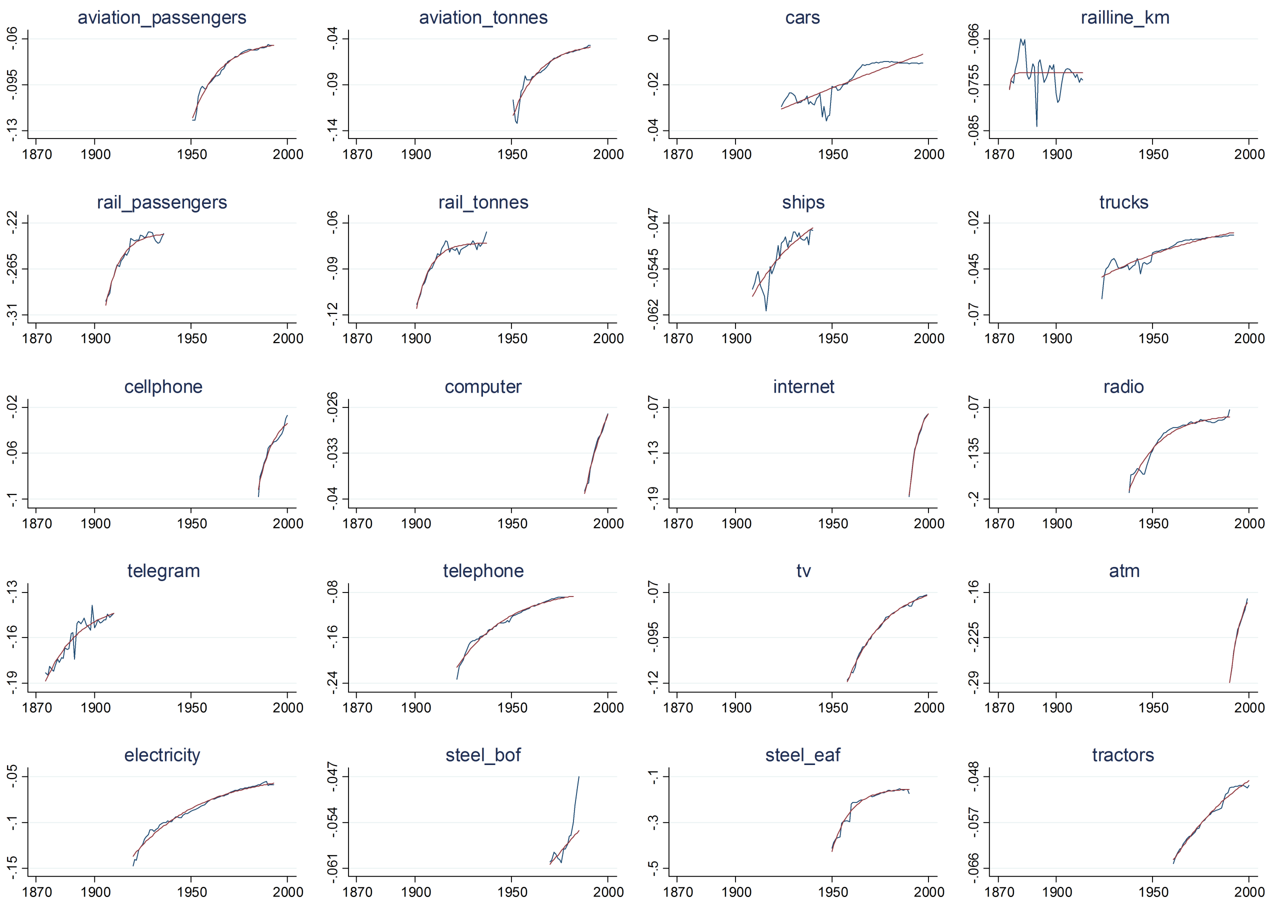

Figure 5: Estimates of β_{4t}^{j} in the balanced sample; fitted lines from the regression in (6)

Note also the goodness of fit of the three-parameter specification (6) to the time-varying estimates β_4^j. Both in the unbalanced and (especially) in the balanced samples the R^2's are very high. The median R^2 for the unbalanced sample is 0.71 and for the balanced sample it is 0.96.

5 Structural estimation

After exploring the presence of geographic interactions in adoption in the data with a reduced form specification, it is informative to study the spatial diffusion of technology in a more structured way. In particular, a structural estimation may serve two purposes. First, it allows us to understand whether a model as stylized as ours can generate key features of the diffusion patterns we observe in the data. Second, it helps identify deep parameters that in our model govern the frequency of interactions between agents and how geographic distance affects the probability of a successful interaction. These parameter values can in turn be used to quantify spatial growth models, as in Desmet and Rossi-Hansberg (2011).

Table 14: Unbalanced Sample[40]										
Sector	Technology	Year	a^j	$s.e.\ a^j$	b^j	$s.e.\ b^j$	c^j	$s.e.\ c^j$	R^2	# Obs.
Trans.	Aviation Pass.	1903	0.0037	0.0021	*0.1772*	0.1500	-0.0005	0.0005	0.09	62
	Aviation Ton.	1903	-0.0234	0.0014	*0.1507*	0.0174	-0.0004	0.0004	0.92	47
	Cars	1885	-0.0365	0.0091	*0.9820*	0.5466	0.0001	0.0010	0.19	81
	Rail Line Km	1825	-0.0063	0.0008	*0.0270*	0.0243	0.0052	0.0065	0.79	39
	Rail Pass.	1825	0.0073	0.0033	*0.5296*	0.4084	-0.0006	0.0005	0.14	49
	Rail Ton.	1825	-0.0085	0.0023	*0.1950*	0.0950	-0.0004	0.0005	0.35	53
	Ships	1776	-0.0067	0.0025	*2.6075*	0.0000	-0.0039	0.0004	0.03	41
	Trucks	1903	-0.0031	0.0050	*0.2088*	0.5213	0.0003	0.0008	0.01	81
Comm.	Cellphone	1973	-0.0107	0.0013	*1.6287*	0.7547	-0.0018	0.0004	0.88	16
	Computer	1973	-0.0006	0.0004	*0.0240*	0.6814	-0.0024	0.0447	0.75	13
	Internet	1983	0.0013	0.0023	*25.779*	0.0000	-0.0016	0.0007	0.15	11
	Radio	1920	-0.0464	0.0020	*0.5203*	0.0413	0.0004	0.0003	0.94	52
	Telegram	1835	0.0098	0.0015	*2.5420*	2.5098	0.0019	0.0002	0.74	41
	Telephone	1876	0.0030	0.0012	*0.0257*	0.0258	-0.0012	0.0015	0.12	101
	TV	1927	-0.0060	0.0012	*0.2978*	0.1059	-0.0007	0.0002	0.58	47
Industry	ATM	1971	0.0011	0.0021	*0.0095*	0.2052	0.1031	2.0728	0.86	11
	Electricity	1882	0.0191	0.0016	*0.2144*	0.0328	0.0000	0.0002	0.69	84
	Steel Bof	1950	-0.0154	0.0019	*0.1711*	0.0459	-0.0003	0.0008	0.81	32
	Steel Eaf	1907	-0.1843	0.0137	*0.0699*	0.0098	0.0033	0.0065	0.89	47
	Tractors	1903	-0.0008	0.0000	*0.0000*	0.0099	0.3239	126.38	0.87	40

We consider a sample of 15 countries and locate them evenly spaced in the unit interval so that their locations can be indexed by $j = \{1, 2, ..., 14, 15\}$.[41] As we have seen in the simulations, geography matters for the diffusion of technology. In particular, diffusion dynamics are affected by the location of adoption leader. We place the leader in the middle of the interval (i.e. $j = 8$).

When bringing the model to the data we need to specify three initial characteristics of the modeled economy: the length of the geographic area, the initial adoption level for the leader, and the initial adoption level for the followers. We set the length of the interval of locations in the model to the maximum distance across countries in the balanced sample for each technology. This calibration is important to generate SDT variables with similar variance as in the data. Furthermore, with this calibration, we can interpret our estimates of δ as implying that the probability of a meeting declines by a factor of $e^{-\delta}$ for every additional 1000 Kms of distance between agents.

[40] Estimates of a^j, b^j and c^j from regression (6), and goodness of fit.

[41] This sample size corresponds to the balanced panel we have used above. In our structural execises we focus on this sample because simulating the unbalanced sample has the additional complexity of countries entering the sample at different times.

Sector	Technology	Year	a^j	s.e. a^j	b^j	s.e. b^j	c^j	s.e. c^j	R^2	# Obs.
Trans.	Aviation Pass.	1903	-0.1202	0.0014	*0.0708*	0.0052	-0.0620	0.0013	0.98	43
	Aviation Ton.	1903	-0.1234	0.0034	*0.0716*	0.0100	-0.0448	0.0036	0.93	41
	Cars	1885	-0.0307	0.0014	*0.0029*	0.0070	0.0938	0.2712	0.76	74
	Rail Line Km	1825	-0.0765	0.0034	*1.0135*	2.2587	-0.0731	0.0006	0.03	39
	Rail Pass.	1825	-0.3009	0.0027	*0.1311*	0.0126	-0.2297	0.0018	0.96	31
	Rail Ton.	1825	-0.1157	0.0019	*0.1386*	0.0136	-0.0728	0.0009	0.95	37
	Ships	1776	-0.0589	0.0011	*0.0334*	0.0258	-0.0417	0.0084	0.76	32
	Trucks	1903	-0.0492	0.0011	*0.0143*	0.0055	-0.0104	0.0096	0.87	69
Comm.	Cellphone	1973	-0.0914	0.0035	*0.1287*	0.0354	-0.0244	0.0080	0.95	16
	Computer	1973	-0.0391	0.0004	*0.0667*	0.0233	-0.0173	0.0053	0.99	13
	Internet	1983	-0.1841	0.0034	*0.2240*	0.0165	-0.0659	0.0035	1.00	11
	Radio	1920	-0.1867	0.0031	*0.0649*	0.0056	-0.0794	0.0026	0.96	53
	Telegram	1835	-0.1880	0.0028	*0.0542*	0.0152	-0.1362	0.0064	0.88	36
	Telephone	1876	-0.2109	0.0022	*0.0370*	0.0025	-0.0719	0.0037	0.98	61
	TV	1927	-0.1189	0.0005	*0.0494*	0.0021	-0.0649	0.0010	1.00	46
Industry	ATM	1971	-0.2878	0.0030	*0.1606*	0.0262	-0.1408	0.0127	0.99	11
	Electricity	1882	-0.1370	0.0013	*0.0277*	0.0019	-0.0452	0.0026	0.98	74
	Steel Bof	1950	-0.0604	0.0006	*0.0026*	0.0661	0.0725	3.2345	0.75	16
	Steel Eaf	1907	-0.4256	0.0080	*0.1034*	0.0073	-0.1518	0.0046	0.97	41
	Tractors	1903	-0.0644	0.0003	*0.0209*	0.0043	-0.0364	0.0040	0.98	40

Table 15: Balanced Sample[42]

We recognize that while diffusion in the model is measured by the percentage of adopters, CHAT variables measure the amount of output produced with the technology (per capita) or the number of units of the technology (per capita). The difference between adoption measures in the model and data is that the data includes an intensive margin (i.e. number of units of technology per adopter) that in the model is absent. We make the model and data comparable by introducing an intensive margin specified as a log-linear function of income. Since the baseline regression in (1) already controls for log income, adding an intensive margin amounts to just adding a technology-specific constant. We compute this constant from the leader's adoption (in CHAT) in the terminal period, T. In particular, in the model, as time goes to infinity, the fraction of adopters goes to 1. Therefore, the (log) intensive margin, $\bar{x}^j$, is equal to

$$\bar{x}^j = \max_i x^j_{iT}.$$

[42] Estimates of a^j, b^j and c^j from regression (6), and goodness of fit.

Given our calibration of the intensive margin, $\bar{x}^j$, the initial (log) fraction of adopters in the leading country is given by

$$\max_i \log G^j_{i0} = \max_i x^j_{i0} - \bar{x}^j, \tag{7}$$

where $\max_i x^j_{i0}$ is the maximum (log) adoption level in CHAT for technology j in the initial year.

In the data, followers differ in their initial adoption levels in the initial year. In order to capture this initial heterogeneity we set the (log) of the share of initial adopters to

$$\log G^j_{k0} = \frac{\sum_i x^j_{i0}}{15} - \sigma + 2\sigma\frac{j-1}{6} - \bar{x}^j, \text{ for locations with index } j \in [1, 7],$$

and to

$$\log G^j_{k0} = \frac{\sum_i x^j_{i0}}{15} - \sigma + 2\sigma\frac{15-j}{6} - \bar{x}^j, \text{ for locations with index } j \in [9, 15],$$

where σ represents the standard deviation of initial adoption across followers.

We then use our model of diffusion to generate time-series of the share of adopters, G^j_{it}, associated with a given α, δ and σ. For each technology, the time series we generate have the same length as the CHAT time series. We then construct the model adoption levels as

$$\hat{x}^j_{it} = \log G^j_{it} + \bar{x}^j.$$

Estimation of α, δ, and $\sigma-$ Once we can generate synthetic adoption data for a given triplet (α, δ, σ), we are ready to apply the estimation procedure. For each technology the estimates of α, δ and σ are those values that minimize the distance in the time-varying coefficients, β^j_{4t}, between the model and the data.

To compute the model's counterpart to β^j_{4t} we proceed as follows. First, we compute the model time-series for SDT ($SDTM$) as,

$$SDTM^j_{it} = \sum_{k \neq i} \hat{x}^j_{it} d_{ik}.$$

Then, the model's counterpart to β^j_{4t} is given by the estimated $\tilde{\beta}^j_t$ from the following regression:

$$\hat{x}^j_{it} = I^j_i + I^j_t + \sum_t \tilde{\beta}^j_t SDTM^j_{it} + \epsilon_{it}. \tag{8}$$

We choose α, δ, and σ to minimize the sum of squared distances between the series of data estimates β^j_{4t} and the series of model estimates $\tilde{\beta}^j_t$.

Estimation results $-$ Figure 6 plots, for each technology, the estimates of β^j_{4t} from the data and the estimates of $\tilde{\beta}^j_t$ associated with the optimal α, δ, and σ. Broadly speaking, the model fits the evolution of the effect of SDT on adoption well for nearly all technologies in our sample. The average R^2 is 89%, and the median 93%. For all technologies except one, $\tilde{\beta}^j_t$ is initially negative, then it starts increasing, and it ends at a less negative level. These patterns reflect the presence of geographic interactions in adoption and the decline in their intensity as technology diffuses and adoption levels become more uniform across countries.[43]

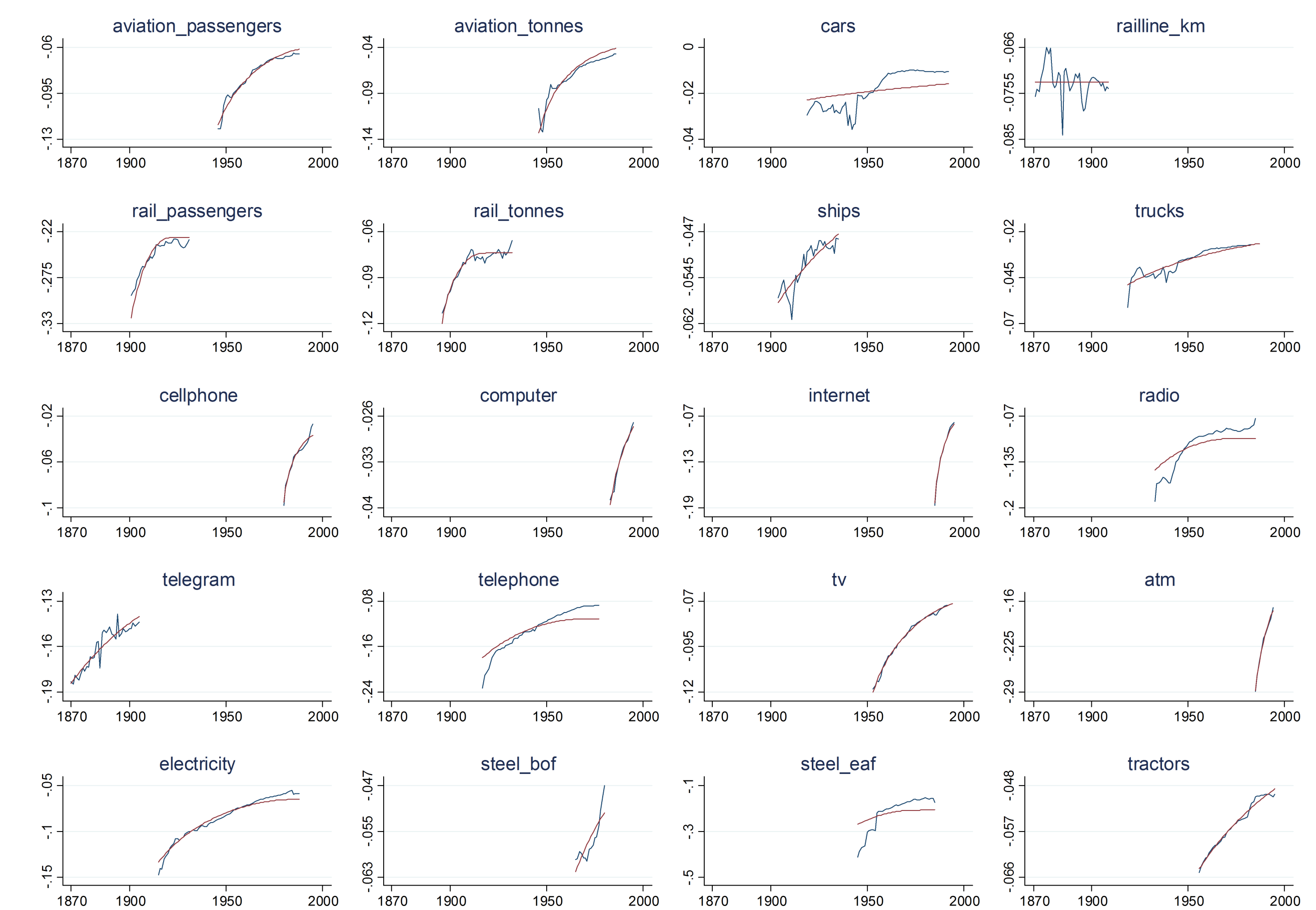

Figure 6: Estimates of geographic interactions in adoption in the balanced sample (data in blue and model in red)

Table 16 reports the point estimates of α, δ, and σ for each technology. Their identification is quite intuitive. α has two effects on $\tilde{\beta}_t^j$. First, a higher probability of contacts, α, increases the unconditional probability of adopting a technology especially for those close to the adoption leader. This force leads to a higher (in absolute value) initial estimate of $\tilde{\beta}_t^j$. Second, a higher α leads to faster diffusion and to a more even cross-country distribution of adoption which reduces the benefits of the proximity to adoption leaders. As a result, a higher α tends to reduce (in absolute terms) estimates of $\tilde{\beta}_t^j$ in subsequent periods. The steepness of the profile of β_{4t}^j is particularly important to identify α because, as we shall see below, the other two parameters also affect the level of $\tilde{\beta}_t^j$,

[43]The fact that typically $\tilde{\beta}_t^j$ does not converge to zero is due either to the fact that technologies have not fully diffused or that, with country fixed effects, $\tilde{\beta}_t^j$ does not need to asymptotically converge to zero. Intuitively, the country fixed effects introduce positive (asymptotic) dispersion on the RHS. Therefore, $\tilde{\beta}_t^j$ cannot converge to zero as the dispersion of the LHS goes to zero. $\tilde{\beta}_t^j$ needs to converge to some negative value to undo the dispersion introduced on the RHS by the country fixed effects.

but they do not have a clear impact on the steepness of the profile. These effects are evident in Table 16 and Figure 6. Technologies with very high estimates of α (e.g., internet, cellphones) tend to have point estimates of $\tilde{\beta}_t^j$ that are initially large and, importantly, tend to exhibit very steep profiles for $\tilde{\beta}_t^j$. Conversely, technologies with low estimates of α such as ships, tractors, or trucks tend to have low (in absolute terms) initial values of $\tilde{\beta}_t^j$ and flatter profiles.

Table 16: Structural Estimates of α, δ, and σ with Initial Heterogeneity

Sector	Technology	Year	α	δ	σ	R^2
Transport	Aviation Passengers	1903	0.136	0.113	0.671	0.966
	Aviation Tons	1903	0.202	5.683	0.000	0.883
	Cars	1885	0.026	2.117	1.625	0.365
	Railline Km	1825	0.0001	0.741	2.750	-0.002
	Rail Passengers	1825	0.155	1.623	0.000	0.778
	Rail Tons	1825	0.202	0.000	1.226	0.936
	Ships	1776	0.031	0.788	2.359	0.758
	Trucks	1903	0.058	1.936	1.788	0.858
Communication	Cellphone	1973	0.509	0.510	1.723	0.953
	Computer	1973	0.115	0.819	4.198	0.979
	Internet	1983	0.800	0.324	0.166	0.997
	Radio	1920	0.098	7.276	0.565	0.681
	Telegram	1835	0.060	2.160	0.005	0.855
	Telephone	1876	0.105	4.933	0.211	0.794
	TV	1927	0.091	0.556	1.387	0.995
Industry	ATM	1971	0.182	9.417	0.053	0.993
	Electricity	1882	0.104	1.785	0.592	0.961
	Steel Bof	1950	0.059	0.643	1.315	0.696
	Steel Eaf	1907	0.126	7.606	0.231	0.432
	Tractors	1903	0.035	0.998	2.528	0.982

The parameter δ also has two effects on $\tilde{\beta}_t^j$. For δ very high, contacts take place predominantly with agents from the same location.[44] Therefore, there is little room for geographic interactions in adoption and the estimates $\tilde{\beta}_t^j$ are low (in absolute value). A reduction in delta allows for the possibility of contacts with other locations which opens the possibility of benefiting from being close to the leader. As a result, a decline in delta leads to higher $\tilde{\beta}_t^j$ (in absolute terms) when δ is high. This force can be seen at work, for example, when comparing the profiles of $\tilde{\beta}_t^j$ for radios vs. telephones.[45] For lower values of δ, a reduction in δ reduces the relevance of location for the frequency of contacts leading to lower absolute values of $\tilde{\beta}_t^j$.

[44] Recall that we simulate a version of the model with a discrete number of locations.

[45] Note that α and σ are similar in both technologies and the lower estimate of δ for telephones leads to larger (absolute) values of $\tilde{\beta}_t^j$ for telephones than for radios.

Finally, a higher σ tends to decrease (in absolute terms) the profile of $\tilde{\beta}_t^j$. Intuitively, heterogeneity in initial adoption increases the dispersion in SDT. This tends to reduce the coefficients $\tilde{\beta}_t^j$ (in absolute value). This effect is evident, for example, when comparing steel produced with electric arc furnaces (eaf) vs. blast oxygen furnaces (bof), or trucks with telegrams.

Analysis of estimates – To gain further confidence in our structural estimation, we can compare the estimates of σ with the actual cross-country dispersion in initial adoption among followers. After all, our motivation to introduce σ in the estimation was to account for initial heterogeneity in adoption among followers. Reassuringly, the estimates of σ are quite similar to the standard deviations in initial adoption levels. In particular, the correlation between the estimated and actual cross-country dispersions in initial adoption is 0.67. The average dispersion in the data is 1.54, while the mean of σ is 1.17.

The point estimates of α and δ provide valuable information about the spatial and temporal diffusion processes. The average estimate of α is 0.15 with a median of 0.10. The average estimate of δ is 2.50 with a median of 1.31. These estimates suggest that the probability of a meeting declines by 73% ($= 1 - e^{-1.31}$) for the median technology for every additional 1000 Kms of distance between agents.

The estimates vary significantly by technology. The standard deviation of the estimates of α across technologies is 0.19, while for δ it is 2.84. α is highest for cell phones and the internet, and lowest for telegrams and tractors. δ is highest for ATMs, radios, and cars and lowest for aviation passengers, internet and cellphones.

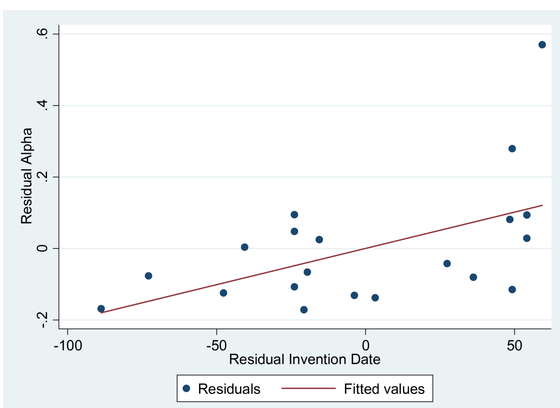

Figure 7A: α vs. invention date.[46]

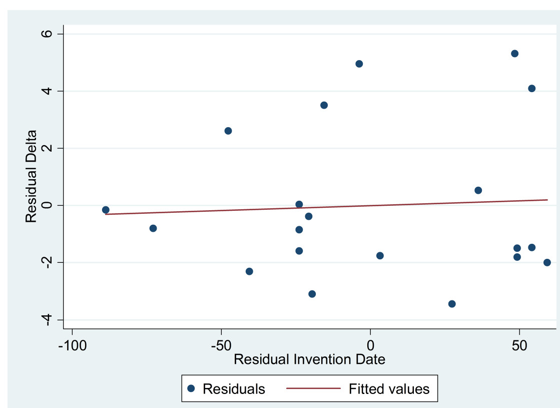

Figure 7B: δ vs. invention date.[47]

The number of technologies (20) is probably too small to identify all the cross-technology patterns in α and δ present in the data. However, careful examination of Table 16 suggests a positive association between the estimates of α and the invention dates of technologies. The correlation between these two variables is 0.48 and it is statistically significant at the 5% level. That is, the diffusion of newer technologies has been fostered by more frequent contacts among agents and therefore has been faster. Figure 7A plots the relationship after controlling for the sector. The

[46] The coefficient in this regression is 0.0019, with a robust standard error of 0.0008.

[47] The coefficient in this regression is 0.004, with a robust standard error of 0.012.

regression line implies that technologies invented ten years later have a probability of contact two percentage points higher. This finding is consistent with the acceleration in the speed of diffusion of technologies documented in Comin and Hobijn (2010) using a very different model and econometric techniques.[48] In Figure 7B we conduct a similar exercise for δ. However, we do not find an association between δ and the invention date.

6 Conclusions

In this paper we have used a new data set of direct measures of technology to study technology diffusion across time and space. Our findings indicate that understanding technology diffusion over space is crucial to understand the speed of technology diffusion. Countries that are far away from the adoption leaders benefit less rapidly from these technologies. In contrast to flows of goods, people, or investments, the spatial effects we identify for technologies vanish over time. For most technologies this implies that the effect of geography is initially strong, decays over time, and eventually disappears. As far as we know, this is the first paper to document these patterns in adoption rates for a large number of technologies and countries.

The empirical pattern of technology adoption over time and space is well accounted for by a simple model of random interactions. The model determines a pattern of adoption for each technology given two key parameters. The frequency of interactions (governed by α) and the spatial decay in the probability of interactions (given by δ). Our structural estimation of the model provides estimates of these parameters for each technology. These estimates show that interactions are more frequent for more recent technologies. Perhaps more important is that our paper provides estimates of structural parameters that can be used to inform spatial theories of growth (as in Eaton and Kortum, 1999, and Desmet and Rossi-Hansberg, 2011). The speed and spatial scope of technology diffusion is a key component to the quantitative implications of these theories. Thus, we hope that the evidence on the significance of the spatial and temporal links in technology adoption we document will prove helpful to stimulate future research in these areas.

[48] Note also that α is statistically lower for technologies in the industry sector.

7 Appendix A: Robustness to other specifications of SDT

This appendix illustrates the robustness of our basic results to different specifications of the SDT variable. To this end, we compute the technology distance interactions, SDT, using the following two alternative specifications. Define SDT2 and SDT3 as

$$SDT2 = \sum_{\forall k \neq c} \frac{x_{kt}}{d_{ck}},$$

and

$$SDT3 = \log(\sum_{\forall k \neq c} e^{(x_{kt})^{-\gamma d_{ck}}}),$$

where, as above, x_{kt} denotes the log of adoption per capita in country k at time t, and d_{ck} denotes the distance (in thousands of km) between countries c and k, and γ is a parameter that we calibrate. As our baseline specification, these alternatives are also sensible ways to capture the interaction between adoption in other countries and how far they are from c. One important difference is that the presence of geographic interactions in adoption *would lead to positive coefficients* (rather than negative) of $SDT2$ or $SDT3$ on country c technology level.

We use the alternative specification to define SDI for each of these formulations, and run Specification 1 of the pooled regression (1) for each of the three different measures of SDT. Table 18 presents the coefficients of SDT. (We denote our baseline specification $SDT1$.) It is clear that we obtain significant geographic interactions from adoption regardless of the specification. In the context of $SDT3$, we have tried various values of γ and the results are not sensitive to its value.

Table 18: Pooled Regressions

	Specification		
	SDT1	SDT2	SDT3 ($\gamma = 1$)
SDT	*-.0001468**** ($4.50e^{-6}$)	*.0013544**** ($1.04e^{-4}$)	*.2873901**** ($5.01e^{-3}$)
# Obs.	53579	53579	53579

Table 19 reproduces the estimates of the coefficients of the interaction terms when decomposing them between the latitude vs. longitude parts for the two specifications of SDT presented above. Again, it is clear from the table that, as in our baseline specification of SDT, interactions along latitudes have stronger effects on adoption than along longitudes.

Table 19: Longitude and Latitude with Alternative Measures of SDT				
	Specification: SDT2		Specification: SDT3	
	1	2	1	2
SDT NS	*$1.87e^{-6}$****	*$2.23e^{-6}$****	*.1603851****	*.1765056****
	($1.65e^{-7}$)	($1.78e^{-7}$)	($1.10e^{-2}$)	($1.07e^{-2}$)
SDT EW	*$-1.50e^{-6}$****	*$-3.24e^{-7}$*	*.1468938****	*.1505348****
	($1.85e^{-7}$)	($1.95e^{-7}$)	($5.24e^{-3}$)	($5.24e^{-6}$)
SDI NS	T.S.	$7.97e^{-7}$	T.S.	-.0810472***
		($1.15e^{-6}$)		($1.83e^{-2}$)
SDI EW	T.S.	$-2.99e^{-6}$	T.S.	.1269351*
		($1.61e^{-6}$)		($1.36e^{-2}$)
# Obs.	52731	52731	53579	53579

8 Appendix B: Upper bound of endogenity bias

In the appendix we detail the back-of-the envelope calculations about the impact of the endogeneity of the SDT variable on the estimates of β_4^j. To this end, let's suppose that country $c's$ adoption level increases by one standard deviation (2.46 in the balanced sample). Since the average distance in the balanced sample is 7.5 (thousands km), the SDT of the other countries will increase on average by 18.45. Since the coefficient β_4^j is -0.0007 (from Table 7, Column1), this should lead to an average reduction in the adoption for the 14 countries other than c of 0.0129 ($= -0.0007 * 18.45$). If the average country is 7.5 thousand Kms from country c, then these declines in adoption will reduce the SDT of country c by 1.35 ($= 0.0129 * 7.5 * 14$). Since the standard deviation of SDT in the balanced sample is 385, the endogenous increase of SDT represents just 0.35% (= 1.35/385) of the observed dispersion of the independent variable (i.e. SDT).

The small share of the dispersion of SDT generated by its endogeneity limits the magnitude of the bias this has on the estimate of β_4^j. To get an approximate bound on the size of the bias, suppose that SDT can be decomposed between the exogenous (SDT^x) and the endogenous (SDT^n) components as follows:

$$SDT = SDT^x + SDT^n \tag{9}$$

To get a back-of-the-envelope bound on the effect of SDT^n on β_4^j, let's consider a univariate version of regression (1) where adoption (x) is the dependent variable and SDT the independent one. In that case,

$$\hat{\beta}_4^j = \frac{Cov(SDT, x)}{\sigma^2(SDT)}$$

where Cov stands for covariance and $\sigma^2(.)$ is the variance. Using (9) and some straightforward manipulations, β_4^j can be decomposed between the exogenous and the endogenous components as follows:

$$\begin{aligned}
\hat{\beta}_4^j &= \overbrace{\frac{Cov(SDT^x, x)}{\sigma^2(SDT)}}^{\text{Exogenous}} + \overbrace{\frac{Cov(SDT^n, x)}{\sigma^2(SDT)}}^{\text{Endogenous}} \\
&= \frac{Cov(SDT^x, x)}{\sigma^2(SDT)} + Corr(SDT^n, x)\frac{\sigma(SDT^n)}{\sigma(SDT)}\frac{\sigma(SDT)}{\sigma(x)}
\end{aligned}$$

With the information we have, it is possible to bound the endogenous component (i.e. the second term). $Corr(SDT^n, x)$ must be higher than -1. From our previous calculations, $\frac{\sigma(SDT^n)}{\sigma(SDT)} = 0.0035$. And from the descriptive statistics in the balanced sample, $\frac{\sigma(SDT)}{\sigma(x)} = \frac{2.46}{385} = 0.0064$. Therefore, the endogenous component of $\hat{\beta}_4^j$ is higher than -0.000022. This represents 3% of the estimate we obtain for β_4^j which is -0.0007.

9 References

Almeida, Paul. 1996. "Knowledge Sourcing by Foreign Multinationals: Patent Citation Analysis in the U.S. Semiconductor Industry." *Strategic Management Journal* 17(Winter Special Issue):421-31

Jaffe, Adam, and Manuel Trajtenberg, and Rebecca Henderson. 1993. "Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108(3):577–98.

Anderson, James and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93(1):170–192.

Anderson, James and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature* 42(3):691–751.

Bandeira, Oriana and Imran Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *Economic Journal* 116:869–902.

Basu, Susanto. 1996. "Procyclical Productivity: Increasing Returns or Cyclical Utilization?" *Quarterly Journal of Economics* 111(3):719–751.

Burnside, Craig, Martin Eichenbaum and Sergio Rebelo. 1995. "Capital Utilization and Returns to Scale." in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER Macroeconomics Annual.* Cambridge, MA: MIT Press. pp. 67-110.

Caselli, Francesco and Jon W. Coleman. 2001. "Cross-country Technology Diffusion: The Case of Computers."American Economic Papers and Proceedings 91:328–335.

Comin, Diego and Bart Hobijn. 2004. "Cross-Country Technology Adoption: Making the Theories Face the Facts." *Journal of Monetary Economics* 51(1):39–83.

Comin, Diego, Bart Hobijn and Emilie Rovito. 2006. "Five Facts You Need to Know about Technology Diffusion." NBER WP 11928.

Comin, Diego and Bart Hobijn. 2009. "The CHAT Dataset." NBER WP 15319.

Comin, Diego and Bart Hobijn. 2009. "Lobbies and Technology Diffusion." *Review of Economics and Statistics* 91, no. 2 (May 2009): 229-244.

Comin, Diego and Bart Hobijn. 2010. "An Exploration of Technology Diffusion." *American Economic Review* 100(5):2031–59.

Comin, Diego, William Easterly and Erick Gong. 2010. "Was the Wealth of Nations Determined in 1000 BC?" *American Economic Journal: Macroeconomics* 2(3):65–97.

Desmet, Klaus and Esteban Rossi-Hansberg. 2011. "Spatial Development."Working Paper, Princeton University.

Diamond, Jared. 1997. *Guns, Germs, and Steel: The Fates of Human Societies.* Norton, W. W. & Company, Inc.

Duranton, Gilles and Puga, Diego. 2004. "Micro-foundations of Urban Agglomeration Economies," in: J. V. Henderson & J. F. Thisse (ed.), *Handbook of Regional and Urban Economics*, Ed. 1, 4, 48, 2063-2117, Elsevier.

Eaton, Jonathan and Samuel Kortum. 1999. "International Technology Diffusion: Theory and Measurement." *International Economic Review* 40(3):537–70.

Foster, Andrew and Mark Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103(6):1176–1209.

Fujita, Masahisa and Jacques-Francois Thisse. 2002. *Economics of Agglomeration.* Cambridge University Press.

Griliches, Zvi. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25(4):501–522.

Hall, Robert E. and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114(1): 83-116.

Head, Keith and Thierry Mayer. 2011. "Gravity, Market Potential, and Development." Journal of Economic Geography 11(2): 281-94.

Keller, Wolfgang. 2004. "International Technology Diffusion." *Journal of Economic Literature* 42(3):752–782.

Kerr, William, and Scott Duke Kominers 2012. "Agglomerative Forces and Cluster Shapes." *mimeo* HBS.

Klenow, Peter J. and Andres Rodriguez-Clare. 1997. "Economic growth: A review essay," *Journal of Monetary Economics*, 40(3): 597-617.

Lucas, Robert and Benjamin Moll. 2011. "Knowledge Growth and the Allocation of Time." NBER WP 17495.

Lucas, Robert. 2009. "Ideas and Growth." *Economica* 76: 1-19.

Maddison, Angus. 2005. "Measuring And Interpreting World Economic Performance 1500-2001." *Review of Income and Wealth* 51(1):1–35.

Mansfield, Edwin. 1961. "Technical Change and the Rate of Imitation." *Econometrica* 29(4):741–766.

Manski, Charles. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60(3):531–542.

Rode, Johannes, and Alexander Weber. 2012. "Does Localized Imitation Drive Technology Adoption? A Case Study on Solar Cells in Germany." *mimeo* TU Darmstadt.

Simmons, Beth, Frank Dobbin & Geoffrey Garrett. 2007. *The Global Diffusion of Markets and Democracy.* Cambridge University Press, First edition, Cambridge, UK.

Thompson, Peter and Melanie Fox-Kean. 2005. "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment." *American Economic Review* 95(1):450–60.