Germs, Social Networks and Growth

Alessandra Fogli and Laura Veldkamp*

August 1, 2013

Abstract

Does the pattern of social connections between individuals matter for macroeconomic outcomes? If so, how does this effect operate and how big is it? Using network analysis tools, we explore how different social network structures affect technology diffusion and thereby a country's rate of technological progress. The network model also explains why societies with a high prevalence of contagious disease might evolve toward growth-inhibiting social institutions and how small initial differences can produce large divergence in incomes. Empirical work uses differences in the prevalence of diseases spread by human contact and the prevalence of other diseases as an instrument to identify an effect of social structure on technology diffusion.

How does the pattern of social connections between individuals affect a country's income? This paper uses tools from network analysis to explore how and how much different social network structures might affect a country's rate of technological progress. Our network model explains why societies might adopt growth-inhibiting structures and allows us to quantify the potential size of these effects. Motivated by the model, we use differences in the prevalence of diseases spread by human contact and the prevalence of other diseases as an instrument to measure an effect of social network structure on technology diffusion.

There is a long history of measuring the speed of information or technology diffusion within various kinds of networks (Jackson (2008), Granovetter (2005)). Given these findings, a simple way to explain the effect of social structure on GDP is to show that some types of social networks disseminate new technologies more efficiently than others and append a production economy where the average technology level is related to output and income. There are two problems with this

*Corresponding author: afogli@umn.edu, Department of Economics, University of Minnesota, 90 Hennepin Ave. Minneapolis, MN 55405. lveldkam@stern.nyu.edu, 44 West Fourth St., rm 7-77, New York, NY 10012. We thank participants at the Minnesota Workshop in Macroeconomic Theory, NBER EF&G meetings, SED meetings, the Conference on the Economics of Interactions and Culture and Einaudi Institute, the Munich conference on Cultural Change and Economic Growth, SITE, NBER Macro across Time and Space, and NBER growth meetings and seminar participants at Bocconi, Brown, USC, Stanford, Chicago, Western Ontario, Minnesota, Penn State, George Washington, and NYU for their comments and suggestions. We thank Corey Fincher and Damian Murray for help with the pathogen data, Diego Comin, Pascaline Dupas, Chad Jones, and Marti Mestieri, for useful comments, and Isaac Baley, David Low, and Amanda Michaud for invaluable research assistance. Laura Veldkamp thanks the Hoover Institution for their hospitality and financial support through the national fellows program. Keywords: growth, development, technology diffusion, economic networks, social networks, pathogens, disease. JEL codes: E02, O1, O33, I1.

explanation. First, social contacts are presumably endogenous. If so, why would a social network structure that inhibits growth evolve and persist? Second, this explanation is difficult to quantify or test. How might we determine if its effects are trivial or not? While researchers have mapped social networks in schools or on-line communities (Jackson, 2008), mapping the exact social network structure for an entire economy is not feasible.

Our theory for why some societies have growth-inhibiting social structures revolves around the idea that communicable diseases and technologies spread in similar ways - through human contact. We explore an evolutionary model, where some people favor stable, local social networks and others do not. Stable, local and fractionalized networks are more insular. They have fewer links with the rest of the community. This limited connectivity reduces the risk of an infection entering the collective, allowing the participants to live longer. But it also restricts the group's exposure to new technologies. In countries where communicable diseases are inherently more prevalent, the high risk of infection makes nodes with distant linkages more likely to die out. A stable, local and fractionalized social network that inhibits the spread of disease and technology will emerge. In countries where communicable diseases are less prevalent, nodes with only local connections will be less economically and reproductively successful. Greater reproductive success of networks that diffuse ideas and germs quickly lead them to dominate social networks in the long run.

The idea that disease prevalence and social networks are related can help to isolate and quantify the effect of social networks on technology diffusion. Isolating this effect is a challenging task because technology diffusion and social networks both affect each other: Technology diffusion is a key determinant of income, which may well affect a country's social network structure. To circumvent this problem, we instrument for social network structure using disease prevalence data. By itself, disease prevalence would be a poor instrument because it is not likely to be exogenous: higher income levels would likely translate into better health and lower disease levels. Therefore, our instrument uses differences in the prevalence of two types of disease. The first type is diseases that are spread directly from person-to-person. These diseases might plausibly affect social structure because changing one's relationships with others can prevent transmission. The second type of diseases are those transmitted only through animals. Since direct human contact does not affect one's probability of infection, the prevalence of such diseases should not affect social networks. Thus, a main contribution of the paper is to use the difference in prevalence of communicable disease and animal-transmitted disease as an instrument to measure the effect of social network structure on income.

Our model explains why communicable disease might be correlated with social network structure, how networks can influence a country's technology diffusion and average productivity, and why less productive social networks might persist. We isolate four aspects of social networks, be-

cause they are important determinants of diffusion speed and we have cross-country data measuring them. Of course, this means that we hold fixed many other aspects of networks that may also differ across countries. Measuring these other aspects of social networks and understanding their effects on economic growth would be useful topics for further research.

Section 1 begins by considering a series of exogenous networks and examines the effect of each network feature on technology and disease diffusion. Then, it considers networks that evolve and explores the reverse effects: how technology and disease affect the types of networks that emerge. Specifically, disease prevalence creates the conditions for growth-inhibiting networks to emerge. Section 2 proposes a framework for identifying the effect of networks on growth and uses model simulations to investigate the magnitude of the predicted effect as well as the rationale for the proposed instrument. Section 3 describes our measures of pathogen prevalence, social networks, and technology diffusion. Section 4 uses this data to test the model's predictions for the relationship between disease prevalence and social network structure. This establishes that disease prevalence is a powerful instrument for social networks. The section then goes on to estimate the effect of social networks on technology diffusion, using the difference in communicable and non-communicable diseases as an instrument. A main finding is that a 1-standard-deviation change in each network feature changes output per worker by 75-135%.

Related literature The paper contributes to four growing literatures. A closely related literature is one that considers the effects of networks on economic outcomes. Most of this literature considers particular firms, industries or innovations and how they were affected by the social networks in place (e.g., see Granovetter (2005) or Rauch and Casella (2001)). In contrast, this paper takes a more macro approach and studies the types of social networks that are adopted throughout a country's economy and how those affect technology diffusion economy-wide. Ashraf and Galor (2012) and Spolaore and Wacziarg (2009) also take a macro perspective but measure social distance with genetic distance. Our network theory and findings complement this work by offering an endogenous mechanism to explain the origins of social distance and why it might be related to the diffusion of new ideas.

Thus in its scope, the paper is more related to a second literature, that on technology diffusion. Recent work by Lucas and Moll (2011) and Perla and Tonetti (2011) uses a search model framework where every agent who searches is equally likely to encounter any other agent and acquire their technology. Greenwood, Seshadri, and Yorukoglu (2005) models innovations that are known to all but are adopted when the user's income becomes sufficiently high. What sets this paper apart is its assumption that agents only encounter those in their network. Our insights about why societies adopt networks that do not facilitate the exchange of ideas and our links to empirical measures

of social networks arise because of this focus on the network topology. This focus is similar to Oberfield (2013). But Oberfield models firms who optimally choose a single firm to connect to, which precludes thinking about the network features we examine.

The third literature is on culture and its macroeconomic effects. Gorodnichenko and Roland (2011) focus on the psychological or preference aspects of collectivism. They use collectivism to proxy for individuals' innovation preferences and consider the effects of these preferences on income. In contrast, we use collectivism as one of many measures of human relationships and assess the effect of those relationships on the speed of technology *diffusion*. Similarly, most work on culture and macroeconomics regards culture as an aspect of preferences.[1] Greif (1994) argues that preferences and social networks are intertwined because culture is an important determinant of a society's network structure. While this may be true, we examine a different determinant of networks that is easily measurable for an entire country, pathogen prevalence. Our evolutionary-sociological approach lends itself to quantifying the aggregate effects of social networks on economic outcomes.

Finally, our empirical methodology draws much of its inspiration from work on the role of political institutions by Acemoglu, Johnson, and Robinson (2002) and Acemoglu and Johnson (2005) and the role of social infrastructure by Hall and Jones (1999). But instead of examining institutions or infrastructure, which are not about the pattern of social connections between individuals, we study an equally important but distinct type of social organization, the social network structure.

1 A Network Diffusion Model

Our model serves three purposes. First, it is meant to fix ideas. The concept of social network structure is a fungible one. We want to pick particular aspects of networks to anchor our analysis on. In doing this, we do not exclude the possibility that other aspects of social or cultural institutions are important for technology diffusion and income. But we do want to be explicit about what we intend to measure.

Second, the model guides the choice of variables that we explore empirically. The model teaches us that four different aspects of social networks facilitate technology diffusion. Informed by these results, we use measures of these aspects of social networks as our independent variables to determine the effect of social networks on technology diffusion.

Third, the model motivates our choice of disease as an instrument for social network struc-

[1] See e.g., Tabellini (2010) and Algan and Cahuc (2007) who examine the relationship between cultural characteristics and economic outcomes, and Bisin and Verdier (2001) and Fernández, Fogli, and Olivetti (2004) who examine the transmission of culture. Durlauf and Brock (2006) review work on social influence in macroeconomics, but bemoan the lack of work that incorporates social network interactions.

ture. Specifically, it explains why disease that is spread from human-to-human might influence a society's social network in a persistent way. The disease-based instrumental variable we use is a valid instrument, regardless of the veracity of this theory. The model simply offers one possible explanation for why disease and social networks might have the robust relationship we see in the data.

The final role of the model is that it helps us answer the following question: The richest countries have income and productivity levels that are 100 times higher than the poorest countries. Can differences in social network structure plausibly explain such large income disparities? To answer this kind of question requires a model. Section 2 takes up this quantitative exercise.

A key feature of our model linking social networks to technological progress is that technologies spread by human contact. This is not obvious since one might think new ideas could be just as easily spread by print or electronic media. However, at least since Foster and Rosenzweig (1995), a significant subbranch of the growth literature has focused on the role of personal contact in technology diffusion; see Conley and Udry (2010) or Young (2009) for a review. In his 1969 AEA presidential address, Kenneth Arrow remarked,

> "While mass media play a major role in alerting individuals to the possibility of an innovation, it seems to be personal contact that is most relevant in leading to its adoption. Thus, the diffusion of an innovation becomes a process formally akin to the spread of an infectious disease."

With this description of the process of technological diffusion in mind, we propose the following model.

1.1 Economic Environment

Time, denoted by $t = \{1, \ldots, T\}$, is discrete and finite. At any given time t, there are n agents, indexed by their location $j \epsilon \{1, 2, \ldots, n\}$ on a circle. Each agent produces output with a technology $A_j(t)$:

$$y_j(t) = A_j(t).$$

Social networks Each person i is socially connected to γ other people. If two people have a social network connection, we call them "friends." Let $\eta_{jk} = 1$ if person j and person k are friends and $= 0$ otherwise. To capture the idea that a person cannot infect themselves in the following period, we set all diagonal elements (η_{jj}) to zero. Let the network of all connections be denoted N.

Spread of technology Technological progress occurs when someone improves on an existing technology. To make this improvement, they need to know about the existing technology. Thus, if

a person is producing with technology $A_j(t)$, they will invent the next technology with a Poisson probability λ each period. If they invent the new technology, $\ln(A_j(t+1)) = \ln(A_j(t)) + \delta$. In other words, a new invention results in a $(\delta \cdot 100)\%$ increase in productivity.

People can also learn from others in their network. If person j is friends with person k and $A_k(t) > A_j(t)$, then with probability ϕ, j can produce with k's technology in the following period: $A_j(t+1) = A_k(t)$.

Spread of disease Each infected person transmits the disease to each of their friends with probability π. The transmission to each friend is an independent event. Thus, if m friends are diseased at time $t-1$, the probability of being healthy at time t is $(1-\pi)^m$. If no friends have a disease at time $t-1$, then the probability of contracting the disease at time t is zero.

An agent who catches a disease at time t loses the ability to produce for that period ($A_j(t) = 0$). Let $d_j(t) = 1$ if the person in location j acquires a transmittable disease (is sick) in period t and $= 0$ otherwise. An agent who is sick in period t dies at the end of period t. At the start of period $t+1$, they are replaced by a new person in the same location j. That new agent inherits the same social network connections as the parent node. When we discuss network evolution, we will relax this assumption. At the start of period t, the new agent begins with zero productivity and learns the technology of each of his friends with probability ϕ, just like older agents do.

1.2 Average path length, infection time and diffusion speed

The speed at which germs and ideas disseminate can be measured by the average path length in a network. To understand the concept of average path length, consider a ring with 10 nodes where each person has 2 friends on either side of them. Node 1 is directly connected to nodes 2,3,9 and 10. The path length from 1 to these four nodes is length 1. They are, in turn, linked to 4,5,7, and 8. The path length from 1 to those 4 nodes is 2. Finally, the shortest path between nodes 1 and 6 is length 3. Since the network is symmetric, there is an identical set of paths from nodes $2-10$ to all the other nodes. Therefore, the average path length is $(4*1+4*2+3)/9 = 1.67$.

Definition 1 *The average path length is the average number of steps along the shortest paths for all possible pairs of network nodes. Let p_{ij} represent the shortest path length between nodes i and j and $\mathcal{N} = \{1, \ldots, n\}$ represent the set of n nodes. Then,*

$$\textit{Average path length} = \frac{1}{n} \sum_{i \in \mathcal{N}} \frac{\sum_{j \in \mathcal{N}/i} p_{ij}}{n-1} \tag{1}$$

If the average path length between individuals is shorter, diseases and ideas disseminate more quickly because they require fewer transmissions to reach most nodes. The next result uses average path length to characterize the mean infection time and the mean discovery time for a new technological innovation. Let $L_j(t)$ represent the last day of person j's life. It is the next period in which the person living in location j gets sick and dies: $L_j(t) = min\{s : s \geq t, d_j(s) = 1\}$. Thus, $L_j(0)$ is number of periods that the person living in location j at time 0 will live. Analogously, let $\alpha_j(0)$ be the number of periods it takes for a new idea, introduced in period 0, to reach person j.

Result 1 *If $\pi = 1$ and $\sum_j d_j(0) = 1$, then the average lifetime $E_j[L_j(0)]$ is monotonically increasing in the average path length of the network.*

If $\phi = 1$, then the average discovery time $E_j[\alpha_j(0)]$ is monotonically decreasing in the average path length of a network.

But faster diffusion is not the same as faster technological innovation. The reason that diffusion accelerates technology growth is that when idea diffusion is faster, redundant innovations are less frequent. Thus, more of the innovations end up advancing the technological frontier. The following result clarifies the mechanism by which the individualist network achieves a higher rate of growth.

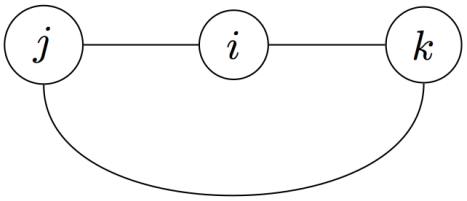
Result 2 *Suppose that at t, two networks have the same $A_j(t)$ $\forall j$. Then the probability that the next new idea arrival will increase the technological frontier is larger in the network with the smaller average path length.*

Taken together, these results explain why ideas and germs spread more quickly in low-path-length networks, why fast diffusion might imply faster technological progress and output growth, and what evolutionary advantages each type of network might offer its adopters. Next, we describe what observable features of a network cause its average path length to be long or short.

1.3 Network feature 1: collectivism vs. individualism

Collectivism is an aspect of a social network structure that has been extensively studied by sociologists. Mutual friendships and interdependence are hallmarks of collectivist societies. To measure this interdependence, we can ask: If i is friends with j and with k, how often are j and k also friends? We refer to a structure where i, j and k are all connected to each other as a *collective*. An example is given below in Figure 1. Therefore, a measure of the extent of shared friendships, and thus the degree of collectivism, is the number of such collectives.

Figure 1: A collective



To count the number of collectives, we look at all the instances in a given network where one node i is connected to two other nodes j, k. Count that as a triple if j and k are connected.[2] In this section, we will fix the number of connections γ to be 4. We vary γ in the following section. We begin by constructing the network that has the largest number of collectives, of any ring network with 4 links to each node and where all the nodes are connected by some path. That maximum-collective network is:

Network 1 *(Ring lattice) In the collectivist social network, each individual j is friends with the 4 closest people. In other words, $\eta_{jk} = 1$ for $k = \{j-2, j-1, j+1, j+2\}$ and $\eta_{jk} = 0$ for all other k.*

This type of ring network, illustrated in figure 2, is a limit of the small work network (Watts and Strogatz, 1998), as the probability of random links goes to zero. Sociologists frequently use the small world network as an approximation to large social networks because of its high degree of collectivism and its small average path length, both pervasive features of real social networks. The appendix shows that there are as many collectives as there are members of the network.

At the other end of the spectrum, we examine a second network that has the lowest possible degree of collectivism. Call it the individualistic network. Because the individualist and collectivist networks should be as similar as possible along all other dimensions, we construct the individualist network by starting from network 1 and changing the smallest number of links, with the smallest distance changes, that eliminates all collectives.

Network 2 *In the individualistic social network, each person is friends with the person next to them and the person m positions away from them, on either side. In other words, for any integer $m \in \{3, \ldots, n/2-3\}$, the network matrix has entries $\eta_{jk} = 1$ for $k = \{j-m, j-1, j+1, j+m\}$ and $\eta_{jk} = 0$ for all other k.*

[2]This collectives measure is related to a common measure of network clustering: Divide the number of collectives by the number of possible collectives in the network to get the *overall clustering* measure (Jackson 2008).

These two network structures are particularly informative because of their starkly different numbers of collectives. This stark difference facilitates matching social institution data with one or the other type of network.

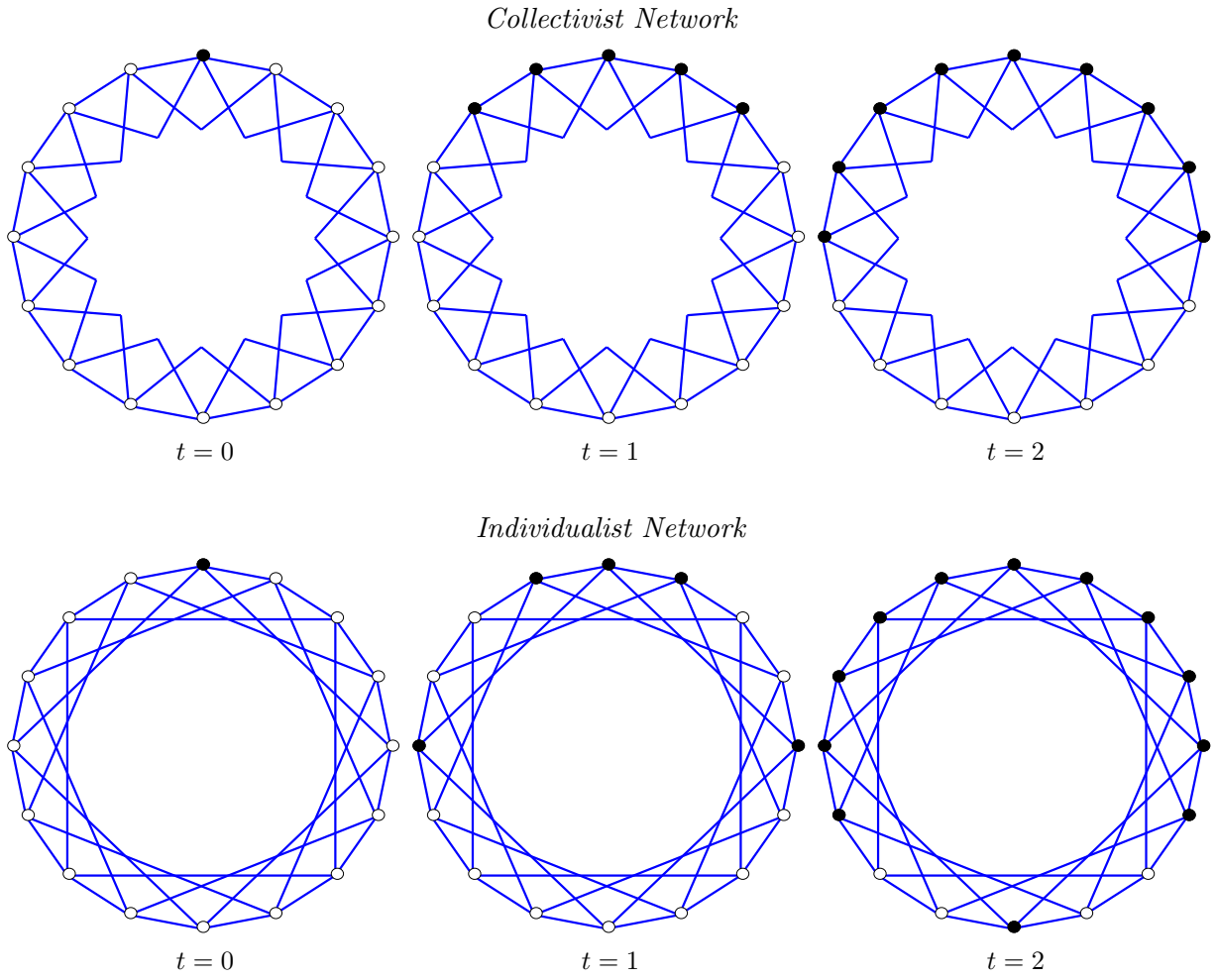
Result 3 *(Collectivism slows diffusion) For any* $m \epsilon (2, n/4)$*, there is a network size* $\bar{n}$ *such that, for any* $n > \bar{n}$*, the average path length in network 1 (collectivist) is longer than the average path length in network 2 (individualist).*

Disease and technology spread more slowly in the collectivist network. The reason is that each contiguous group of friends is connected to at most 4 non-group members. Those are the two people adjacent to the group, on either side. Since there are few links with outsiders, the probability that a disease within the group is passed to someone outside the group is small. Likewise, ideas disseminate slowly. Something invented in one location takes a long time to travel to a far-away location. In the meantime, someone else may have re-invented the same technology level, rather than building on existing knowledge and advancing technology to the next level. Such redundant innovations slow the rate of technological progress and lower average consumption.

Figure 2 illustrates the smaller path length and faster diffusion process in individualist networks. In a simple case where the probability of transmission is 1, each frame shows the transmission of an idea or disease introduced to one node at time 0. The "infected" person transmits that technology to all the individuals she is connected to. In period 1, 4 new people use the new technology, in both networks. But by period 2, there are 9 people using the technology in the collectivist network and 12 using it in the individualist network. In each case, an adopter of the technology transmits the technology to 4 others each period. But in the collectivist network, many of those 4 people already have the technology. The technology transmission is redundant. This example illustrates why, on average, ideas and diseases will diffuse more slowly through a collectivist network than an individualist one.

Could Collectivism Facilitate Technology Diffusion? Perhaps Arrow was not correct and technology diffusion is not a process "formally akin to the spread of infectious disease." Instead, a technology is adopted only when a person comes in contact with multiple other people who have also adopted it. This is called complex contagion. Centola and Macy (2007) show that it is theoretically possible that having many mutual friendships makes it more likely that groups of people adopt a technology together. However, these same authors admit that "We know of no empirical studies that have directly tested the need for wide bridges in the spread of complex contagions." In other words, it is a theoretical possibility, without empirical support. In contrast, the idea that technology is adopted when information about the success of the technology arrives

Figure 2: Slower diffusion in the collectivist network (top) than the individualist network (bottom).



from a single social contact is a well-documented phenomena. (See e.g. Foster and Rosenzweig (1995), Munshi (2004), or Conley and Udry (2010).)

1.4 Network feature 2: Degree

The degree of a node is the number of connections that node has to other nodes. In the context of a social network, degree is the number of friends a person has.

Network 3 *(Network with degree γ)*
Consider a ring lattice with where every node has degree γ, where γ is even. Each individual j is friends with the γ closest nodes. In other words, $\eta_{jk} = 1$ for $k = \{j-\gamma/2, \ldots, j-1, j+1, \ldots, j+\gamma/2\}$ and $\eta_{jk} = 0$ for all other k.

A ring lattice social network with a higher degree has a lower average path length. With more connections, it requires fewer steps to reach other nodes. This will speed diffusion of germs and technologies.

Result 4 *(Higher degree speeds diffusion) The average path length in network 3 is a decreasing function of degree γ.*

1.5 Network feature 3: Link stability

The third network feature we introduce is the possibility that social links change over time. We model this as a small probability that each period, each link might be randomly reassigned to another pair of nodes. This is an extension of a small-world model, proposed by Watts and Strogatz (1998). Links stability then refers to the probability that a link is *not reassigned* in any given period.

Network 4 *(Small world network)*
Begin with network 1. For each link in network 1, randomly break the link with probability p. For each broken link, form one new link (called a shortcut) that connects any pair of non-connected nodes with equal probability. Once formed, the network is unchanged at each date $t \geq 0$.

The small world network is a useful benchmark, but is not dynamic. As such, it is not a useful tool for studying changes in social linkages over time. To study the stability of links, we consider a model where shortcuts are formed and broken every period. This death and rebirth process keeps the network in a non-degenerate state, where link stability is related to expected average path length of the network.

Network 5 *(Unstable network) At time 0, the network is network 4. At each date $t > 0$, break each link with probability $\tilde{p}$. For each broken link, form one new shortcut link which connects each pair of non-connected nodes with equal probability. In every period, each shortcut disappears with probability z.*

The key link stability result is that a lower probability of forming shortcuts (greater link stability) increases the average path length between nodes. As such, it decreases the speed of diffusion.

Result 5 *(Link stability slows diffusion.) In steady state, the expected average path length of network 5 is a decreasing function of the rewiring probability p.*

1.6 Network feature 4: Fractionalization

Another feature of a social network is that there might be fractionalization, meaning social groups that have almost no social ties between them. We start with the static small world network describe above and add factions. These are groups of nodes that have no random links between them. For example, if two groups do not speak the same language, they cannot be socially connected. For simplicity, we consider equal-sized factions.

Definition 2 *(Fractionalized network) In a network with F factions and n nodes, where n/F is an integer, nodes $\{(f-1)n/F+1, (f-1)n/F+2, \ldots, fn/F\}$ comprise faction f.*

Network 6 *Begin with network 1. For each node i in network 1, form a new link with probability p. The new link connects i with a single node j that is not already connected to i $n(i,j)=0$ and that is in the same faction $f(i)=f(j)$. Each of these feasible links if formed with equal probability. The new random link never connects nodes in different factions. Once formed, the network is unchanged at each date $t \geq 0$.*

Result 6 *(More factions slow diffusion.) Let $\alpha > 1$ be an integer. Then, the expected average path length with αF factions is greater than the expected average path length with F factions.*

This network is like a small world network, but where there is zero probability of forming random links outside your faction of size n/F. For example, if there are 2 factions, then start with network 1 (ring). For each node, form a new link with probability p, where that new link randomly connects any two nodes in the *same half of the ring.* If there are 3 factions, the new link randomly connects any two nodes in the same third of the ring, etc.. Adding more factions forces the small world links to connect only nodes that are in smaller and smaller neighborhoods of each other. Because it eliminates long-distance shortcuts, fractionalization makes the geodesic distance between nodes in different faction longer. An increase in F therefore increases the average path length.

1.7 Network Evolution Model

So far, we have simply described diffusion properties of various networks. This leaves open the question of why some societies adopt a network that inhibits growth. One approach would be to work with a network choice model. But equilibria in such models often do not exist and when they do, they are typically not unique. Instead, we consider an evolutionary model where the network changes as agents die and new ones are born in their place. This evolutionary model also helps to explain why growth-inhibiting social networks might persist long after most diseases have died out.

To keep the model tractable and transparent, we focus on one dimension along which networks might evolve: prevalence of collectives. But the logic of these results clearly carries over to the other network features as well.

Production, endowments and the diffusion processes for technology and disease are the same as in the fixed-network model. In addition, at each date t, each person j can be one of two types. In principle, these two types could represent differences in link stability, or membership in a faction. But for concreteness, we will consider an example where nodes types are either collectivist $\tau_j(t) = co$ or an individualist $\tau_j(t) = in$. All agents are linked to the two people adjacent to them. In addition, they are linked to at least one other person. Which other people depends on their type and the type of their neighbors. Individualists form links with those adjacent to them and someone four spaces to their right. For example, if the person is in location j, they are linked to $j-1$, $j+1$ and $j+4$. Collectivists form links with those adjacent to them and someone two spaces to their right. For example, if the person is in location j, they are linked to $j-1$, $j+1$ and $j+2$. In addition, a person of either type might be linked to nodes $j-2$ and/or $j-4$, depending on whether the agents in those locations are individualist or collectivist. In other words, a person's own type governs their links to the right (with indices higher than yours, except near n); others' types govern links to the left.

A person's type is fixed throughout their lifetime. The network structure only changes when someone dies. There are two reasons an individual can die. First, they can acquire the disease. Someone who acquires the disease at time t has zero output in period t. At the end of period t, they die. Second, each period, each person dies with probability z from a disease that is not spread from person-to-person. This probability is independent across time and individuals. When someone at node j dies in period t, then at the start of period $t+1$, a new person inhabits that node. This second cause of death allows the network to evolve, even after the disease has died out.

A newborn person inherits the best technology from the set of people that the parent was socially connected to. He also inherits the type of the person with that best technology. In other words, if the person at node j is socially connected to nodes $\{k : \eta_{jk}(t) = 1\}$ and dies at time t, the new person at node j at time $t+1$ will start with technology $\max_{\{k:\eta_{jk}(t)=1\}} A_{kt}$. Let k^* be the argument that maximizes this expression (i.e. the friend with the highest time-t technology), then the time-$(t+1)$ type of the person is the same as the time-t type of person k^*: $\tau_j(t+1) = \tau_{k^*}(t)$.

The idea behind this process is that evolutionary models often have the feature that more "successful" types are passed on more frequently. At the same time, we want to retain the network-based idea that one's traits are shaped by one's community. Therefore, in the model, the process by which one inherits the collectivist or individualist trait is shaped by one's community, the social network, and by the relative success (relative income) of the people in that network.

1.8 Theoretical Results: Network Evolution

The question we want this model to answer is: Why do some societies end up with a collectivist network even though it inhibits growth? These results describe the long-run properties of networks and disease. They explain how disease prevalence can permanently alter social structure. This is important because it rationalizes differences in social structures that persist even after diseases have been eliminated. The first result shows that eventually, the economy always converges to either the fully collectivist network (1) or the fully individualist one (2).

Result 7 *With probability 1, the network becomes homogeneous:* $\exists T$ *s.t.* $\tau_j(t) = \tau_k(t)\ \forall k$ *and* $\forall t > T$.

In other words, after some date T, everyone will have the same type forever after. They might all be individualist or all be collectivist. But everyone will be the same. The reason for this is that since traits are inherited from neighbors, when a trait dies out, it never returns. The state where all individuals have the same trait is an absorbing state. Since there are a finite number of states, and whenever there exists a j, k such that $\tau_j(t) \neq \tau_k(t)$, every state can be reached with positive probability in a finite number of steps, then with probability one, at some finite time, an absorbing state is reached and the economy stays there forever after.

Similarly, having zero infected people is an absorbing state. Since that state is always reachable from any other state, with positive probability, it is the unique steady state.

Result 8 *With probability 1, the disease dies out:* $\exists T$ *s.t.* $d_j(t) = 0 \quad \forall j$ *and* $\forall t > T$.

What these results teach us is that which network type will prevail is largely dependent on which dies out first, the individualist trait, or the disease. When there is a positive probability of infection, people with individualist networks have shorter lifetimes, on average. If disease is very prevalent, it kills all the individualists and the society is left with a collectivist network forever after. If disease is not very prevalent, its transmission rate is low, or by good luck, it just dies out quickly, individualists will survive. Since they are more economically successful, they are more likely to pass on their individualist trait. So, the economy is more likely to converge to an individualist network. This is not a certain outcome because of exogenous random death. It is always possible that all individualists die, even if the disease itself is no longer present. These results hold, no matter if τ represents collectivist types, link stability or types that stick to their factions. The main take away is that networks can persist long after the conditions to which they were adapted have changed.

2 Connecting Model and Data

We use a calibrated model simulation to motivate our empirical analysis. First, we show that according to our model, differences in networks can potentially explain large differences in incomes across countries. Second, we use the model to show that the difference in prevalence of the two types of disease has no direct effect on technology, making that difference a valid instrument. The model is not rich enough to produce predicted growth rates or disease rates that are accurate. Rather, the objective is simply to understand the nature of the model's predictions and gauge whether the predicted effects are trivial or not.

2.1 Framework for Measurement

Our objective is to better understand how social networks affect technology diffusion and economic development. The difficulty is that economic development also can potentially change the social network structure. The challenge is to isolate each of these two effects. To do this, we consider the following structural model:

$$A = \beta_1 + \beta_2 \tilde{N} + \epsilon \tag{2}$$

where A is the speed of technology diffusion, $\tilde{N}$ is a social network feature (collectivism, link stability or fractionalization), the β's are unknown coefficients and ϵ is a mean-zero residual orthogonal to $\tilde{N}$. Social network structure depends on average productivity $\bar{A}$, as well as the prevalence of socially transmittable diseases $\bar{d} = \sum_i d_i(t)/n$, and the prevalence of other disease z:

$$\tilde{N} = \beta_3 + \beta_4 A + \beta_5 \bar{d} + \beta_6 z + \eta, \tag{3}$$

where η is a mean-zero residual orthogonal to A, $\bar{d}$ and z. The coefficient of interest is β_2, which measures the effect of network structure $\tilde{N}$ on technology diffusion A.

This model recognizes the endogeneity problem inherent in estimating the relationship between A and $\tilde{N}$. It incorporates our main hypothesis, that network structure $\tilde{N}$ matters for technology A, but it also reflects the idea that perhaps technology (and income) can cause social networks to change as well. Because A depends on $\tilde{N}$ and $\tilde{N}$ depends on A, an OLS estimate would be biased.

Our theory suggests that an instrument with power to predict social network structure $\tilde{N}$ is total disease prevalence $\bar{d} + z$. But, this is not likely to be a valid instrument both because technology affects disease (vaccines are a technology, for example) and because poor health reduces productivity and diminishes one's capacity for invention. We capture the correlation between disease prevalence and technology, from both directions of causality, in the following relationship, which says that,

after controlling for networks, there is a residual correlation between technology and disease:

$$\epsilon = \delta_1 + \delta_2(\bar{d} + z) + \xi. \tag{4}$$

If $E[\epsilon(\bar{d} + z)] \neq 0$, in other words, if $\delta_2 \neq 0$, then disease prevalence is an invalid instrument.

To resolve this problem, we use the difference in human disease prevalence and zoonotic disease prevalence $(\bar{d} - z)$ as our instrument. When $var(\bar{d}) = var(z)$, the difference $(\bar{d} - z)$ is orthogonal to the sum $(\bar{d} + z)$. Therefore, in our final exercise, we scale z to give it the same variance as $\bar{d}$ to ensure that the orthogonality holds. Thus, our identifying assumption is

$$E[\epsilon(\bar{d} - z)] = 0. \tag{5}$$

Since in Equation 4 we restrict the coefficients on $\bar{d}$ and z to be the same, we assume that human disease prevalence and zoonotic disease prevalence have the same effect on technology. Hence the total effect on technology is determined by the sum $\bar{d} + z$. This is orthogonal to the composition of the effect between the two types of disease, $\bar{d} - z$, which has no direct effect on A. But as long as $\beta_5 \neq \beta_6$ in (3), then human and zoonotic diseases have different effects on social networks $\tilde{N}$. Therefore, since the diseases have different effects on networks $\tilde{N}$ and similar effects on the speed of technology diffusion A, the instrument $(\bar{d} - z)$ can be a powerful and valid instrument.

Finally, note that we do not need to know all the determinants of social structure. Rather, any subset of the determining variables can serve as valid instruments for $\tilde{N}$. Similarly, we do not need to observe $\tilde{N}$ exactly. A proxy variable with random measurement noise is sufficient for an unbiased instrumental variables estimate of the coefficient β_2.

2.2 Parameter Choice

To evaluate magnitudes, we need to choose some realistic parameter values. The key parameters are the probabilities of disease and technology transmission, the initial pathogen prevalence rate and the rate of arrival of new technologies. These parameters are summarized in Table 1.

For the initial pathogen prevalence rate, we use the annual tuberculosis death rate in China, a country where the disease was endemic. Tuberculosis is the most common cause of death in our sample.[3] One would like to choose the probability of disease transmission to target a steady state rate of infection. But, as we've shown, the only steady state infection rate is zero. Thus, we set the transmission rate so that, on average, the disease disappears in 150 years. This average masks

[3]Note that this is a mortality rate, not an infection rate. Since individuals who get sick in the model die, this is the relevant comparison. Also, it is a conservative calibration because it uses only one disease and it would be easier to get large effects out of a higher disease prevalence rate.

Table 1: Parameters and their empirical counterparts

	Parameter	Value	Target
Initial disease prevalence	$\bar{d}(0)$	0.5%	TB death rate in China
Disease transmission probability	π	32%	Disease disappears in 150 years (indiv country avg)
Innovation productivity increase	δ	30%	2.6% growth rate in individualist country
Technology transfer probability	ϕ	50%	Half-diffusion in 20 years (Comin et. al. '06)
Technology arrival rate	λ	0.25%	1 arrival every 2 years (Comin et. al. '06)
Exogenous death rate	z	1/70	average lifespan
Network degree	γ	4	Modal number of close friends in GSS data
Link instability rate	p	10%	Probability of moving in GSS data (7%)

large heterogeneity. In many simulations, the disease will disappear after 2 periods. In others, it will persist for hundreds of years. Thus, the economy starts with a given fraction of the population being sick and each sick person represents an independent 32% risk (π) of passing the disease on to everyone that person is friends with.

Everyone starts with a technology level of 1. But each period, there is a chance that any given person may discover a new technology that raises their productivity. The rate of arrival of new technologies is calibrated so that a new technology arrives in the economy every 2 years, on average. This corresponds to the average rate of adoption of technologies in the (Comin, Hobijn, and Rovito, 2006) data set. The magnitude of the increase in productivity from adopting a new technology is calibrated so that the individualistic network economy (more likely to be the developed economy in the data) grows at a rate of 2.2% per year. The probability of transmitting a new technology to each friend (λ) is chosen to explain the fact that for the average technology, the time between invention and when half the population has adopted the technology is approximately 20 years (Comin, Hobijn, and Rovito, 2006). Finally, in the evolutionary model, Network 5, there is a probability of exogenous death. We choose this probability to match an average lifespan in a low-disease economy of approximately 70 years.

2.3 How Much Effect Might Networks Have on Output?

To measure the effect of each network feature, we start with a ring network and then vary each of our four network features, one at a time. For each value of individualism, degree, heterophily

and fractionalization, we simulate the models in sections 1.3-1.6, for 5-11 parameter values, with 50 independent runs, for 100 periods each.[4] We record the average growth rate and regress the growth rate on the network feature. Table 2 reports the coefficients of an OLS regression for each network feature. It reveals that while changes in individualism, degree and factions have small effects on growth, a 10% rise in the probability of a long-distance link increases annual growth by 0.09%. Similarly, giving each agent in the economy 10% more social connections would increase growth by 2.3%.

Table 2: Network effects on productivity growth in the model. Column 1 reports $100 \times \beta_2$ coefficient in the regression $\bar{gA} = \beta_1 + \beta_2 \tilde{N} + \epsilon$ where $\bar{gA} = 1/(nT) \sum_{t=1}^{T} \sum_{j=1}^{n} A_{it}$, and $\tilde{N}$ represents one of four network features: the proportion of nodes who are individualistic, the degree m of the network, the probability of forming a shortcut p or the number of factions f. Column 2 reports $100 \times \beta_2$ coefficient in the regression $ln(\bar{gA}) = \beta_1 + \beta_2 ln(\tilde{N}) + \epsilon$.

Characteristic	Levels	Logs
% Individualistic	1.97	3.43
	(0.20)	(0.17)
Degree m	0.06	22.94
	(0.00)	(0.83)
Prob of shortcut p	6.38	0.91
	(0.33)	(0.07)
Number of factions f	-0.01	-7.58
	(0.00)	(0.37)

This simple exercise makes the point that a difference in network structure can create a small, but permanent friction in technology diffusion. When cumulated over a long time horizons, this small friction has the potential to explain large differences in countries' incomes.

2.4 Is the Difference between Diseases a Valid Instrument?

The difference in the prevalence of socially transmittable and other disease is a valid instrument if equation 5 holds. To show that this condition holds in our model, we hold the network fixed (network 4) and vary the initial prevalence of both types of disease.[5] But since socially transmittable disease spreads and typically becomes more prevalent over time, but the other disease does not spread, comparing the rates of initial prevalence is not a valid comparison. Therefore, our statistical analysis considers the relationship between average prevalence of the disease in the first 100 periods and productivity growth. The other parameters used in the simulation are those in table 1.

[4]The parameter values for each simulation are on a grid evenly-spaced nodes: $[0 : 0.1 : 1]$ for individualism, $[2 : 2 : 10]$ for degree γ, $[0 : 0.01 : 0.1]$ for the probability $\tilde{p}$, and $[0 : 1 : 10]$ for the number of factions F.

[5]The set of simulation nodes used for initial prevalence of both types of disease are $[.004 : .004 : .02]$. The resulting average prevalence of transmissible disease ranges between $[0, 0.106]$.

Table 3: The effect of socially transmittable and other disease on GDP growth in the model. Table reports 100 times a β_2 coefficient in the regression $\bar{gA} = \beta_1 + \beta_2 x + \epsilon$ where $\bar{gA}$ is the average growth rate defined in table 2, x is the average prevalence (fraction of the population infected at a given time) of either transmittable disease $\bar{d}_0$ or other (zoonotic) disease z.

Dependent variable: Productivity growth		
Transmissible disease	$\bar{d}_0$	-1.95
		(0.51)
Zoonotic disease	z	-6.33
		(0.70)
Difference	$\bar{d}_0 - z$	-0.94
		(0.55)

Both the transmissible and zoonotic diseases reduce productivity in a significant way. A 10% increase in prevalence reduces average GDP growth by 0.2-0.6%. This is not surprising since by assumption, a sick agent has zero productivity. What is important here is that the difference between transmissable and zoonotic disease prevalence is not a significant predictor of productivity growth. This coefficient is not significant at the 5 or even 10% significance levels, despite the fact that we generated 10,000 independent simulations of the model, under different starting conditions, to run these regressions on. What we learn from this finding is that, if the network connections are held fixed, there should be no significant direct effect of the difference in diseases on productivity growth. The reason is that both diseases affect productivity in the same way, by making people sick and thus unproductive. Since the two diseases have similar effects, the difference in prevalence has no effect. Thus, the model motivates our choice of disease difference as an instrument to capture network effects, without affecting technology directly. Of course, there may be forces outside the model that invalidate our instrument. We address those in the next section.

3 Data

Our theory is about the relationship between pathogen prevalence, social networks, and technology diffusion. We have assembled a data set that contains these variables for at least 62 countries. This section describes how each one is measured. Additional details, maps and summary statistics are in the appendix.

3.1 Measuring Pathogen Prevalence

We measure the presence of deadly pathogens in two ways. The first approach recognizes that disease conditions may take a long time to affect social networks and therefore it is desirable to

use historical data. At the same time, because our identification strategy relies on differences in disease prevalence, our data must be available for many different diseases, across many countries. One can go back to the colonial period (as in Acemoglu, Johnson, and Robinson (2001)), but the different kinds of diseases that we need to implement our identification strategy are not present in that data. Our approach recognizes that, if we want to uses differences in diseases as an instrument, it is useful to have a large number of each type of disease. Therefore, we use contemporary (2005) data with the prevalence of 34 diseases in 78 geopolitical regions. This data does appear to capture some long-run features of the epidemiological environment because they are remarkably consistent with the colonial data.[6] Furthermore, we compiled historical prevalence of 9 different pathogens in the 1930's and used it to re-estimate our key results. Our estimates of the effect of social network structure on technology diffusion (appendix B) are nearly identical to those with the contemporary disease data.

Our contemporaneous disease data come from GIDEON (Global Infectious Diseases and Epidemiology Network) and use a 3-point coding scheme to report the 2011 prevalence of 34 of the most common infectious diseases. For many of these diseases, the scheme is coded directly by GIDEON; in these cases, a value of "1" means "not endemic" (cases do not originate in this country), a value of "2" means "sporadic" (<1 case per million people, per year), and a value of "3" means "endemic" (an ongoing presence). The countries with the highest pathogen prevalence are Brazil, India, China, Nigeria and Ghana. Countries with the lowest prevalence include Canada, Switzerland, Luxembourg, Hungary and Sweden. The complete list of diseases we use, along with characteristics of each disease, is reported in table 13.

To identify the effect of disease on social network structure, we will use the difference in the prevalence of various types of diseases. Epidemiologists often classify infectious diseases by reservoir.[7] The reservoir is any person, animal, plant, soil or substance in which an infectious agent normally lives and multiplies. From the reservoir, the disease is transmitted to humans. Animals often serve as reservoirs for diseases. There are also nonliving reservoirs, such as soil, which is a reservoir for fungi and tetanus. Figure 13 summarizes the properties and classification of all the pathogens that we collected data on.

Human-specific d_{hs} Many diseases have only human reservoirs, even though they historically may have arisen in other species, such as measles which originated in cattle. Such diseases may be spread with the help of an animal (called a vector), such as a mosquito that injects one person's blood in another person. But it is in the human, not in the mosquito, where the

[6]The data on pathogen prevalence from 2005 and the 1930's line up quite well the data on mortality rates from the colonial period (see Figure 5 in the appendix) for the subset of countries we have in common, showing that our data captures the same long run differences in the epidemiological environment.

[7]See e.g., Smith, Sax, Gaines, Guernier, and Gugan (2007) or Thornhill, Fincher, Murray, and Schaller (2010).

disease flourishes. Human-specific diseases in our data set include Diptheria, Filaria, Measles and Smallpox. The variable d_{hs} is defined as $d_{hs} \equiv \sum_{l \epsilon \mathcal{HS}} \text{prevalance}_l$, where l is a disease and $\mathcal{HS}$ is the set of all human-specific diseases.

Zoonotic z Other diseases, although they infect and kill humans, develop, mature, and reproduce entirely in non-human hosts. These are zoonotic diseases. Humans are a dead-end host for infectious agents in this group. Our zoonotic diseases include anthrax, rabies, schistosomiasis (SCH), tetanus, and typhus (TYP). The variable G_s is defined as $G_s \equiv \sum_{l \epsilon \mathcal{Z}} \text{prevalance}_l$, where l is a disease and $\mathcal{Z}$ is the set of all zoonotic diseases.

Multi-host d_m Some infectious agents can use both human and non-human hosts to complete their lifecycle. We call these "multi-host" pathogens. Our multi-host diseases include leishmaniasis (LEI), leprosy (LEP), trypanosomes (TRY), malaria (MAL), dengue (DEN) and tuberculosis (TB). The variable d_m is defined as $d_m \equiv \sum_{l \epsilon \mathcal{HS}} \text{prevalance}_l$, where l is a disease and $\mathcal{M}$ is the set of all multi-host diseases.

Since multi-host and human-specific pathogens can reside in humans, they have the potential to affect the relative benefits of a social network. Zoonotic pathogens are not carried by people, only by other animals. Their prevalence is less likely to affect the benefits of any particular social network. Therefore, for the purposes of our analysis, we will group human-specific and multi-host diseases together. We define the variable $\bar{d} \equiv d_{hs} + d_m$. It is the sum of 22 human and multi-host diseases, while z is the sum of 12 diseases.

Using the disease prevalence data from each era separately, we construct the following two differences to use as instrumental variables:

$$\Delta\text{germ} \equiv \bar{d} - z \tag{6}$$

$$\Delta\text{germ\_std} \equiv \frac{\bar{d}}{std(\bar{d})} - \frac{z}{std(z)} \tag{7}$$

3.2 Measuring social networks

Measuring collectivism In our model, collectivism is defined as a social pattern of closely linked or interdependent individuals. What distinguishes collectives from sets of people with random ties to each other is that in collectives, it is common that two friends have a third friend in common. This is the sense in which they are interdependent.

In 1970, Hofstede (2001) surveyed IBM employees worldwide to find national differences in cultural values. He performed a factor analysis of the survey responses, and found two factors that

together can explain 46% of the variance in survey responses. He labeled one factor "Collectivism vs Individualism", and uses it to construct an index of individualism that ranges from between 0 (strongly collectivist) to 100 (strongly individualist) for 72 countries. Hofstede describes collectivist and individualist societies as follows: "on the individualist side we find societies in which the ties between individuals are loose... On the collectivist side, we find societies in which people from birth onwards are integrated into strong, cohesive in-groups, often extended families..." While Hofstede's survey asks questions that are not directly about the pattern of social relationships, there is a body of sociological theory and evidence that supports the connection between the behaviors that Hofstede asks about and the pattern of network collectives as described in our model. Appendix B contains more details about Hofstede's survey questions, sociological theories that link these questions to network structure, and other correlated social survey measures that clarify the interpretation of Hofstede's index.

The ideal data to measure collectivism would be each country's prevalence of social collectives. There are a handful of studies that map out partial social networks, but only for small geographic areas, across eight countries. (See Fischer and Shavit (1995) for a review.) But these studies do bolster the connection between Hofstede's survey outcomes and social networks collectives. Table 9 in the Appendix shows that highly collectivist countries, according to Hofstede, have a higher average prevalence of network collectives.

Measuring network degree Since we do not have a large cross-country panel of social network data, we need to use a proxy for the number of social connections an average resident of each country has. Our proxy for network degree in foreign countries is the average number of close friends reported by U.S. immigrants from that country.

Our data comes from the General Social Survey[8] (GSS). The variable *frinum* asks people to report how many close friends they have. Next, we regress this combined response on : (1) age, (2) income, (3) marital status, (4) dummies for education level, and (5) what fraction of the U.S. population shares the respondent's country of origin. Going forward, we use the residuals from this regression as the explanatory variable. Finally, we select respondents that report being first or second-generation immigrants and average their residuals to form a *stability* variable for each reported ethnicity. See appendix B for more details.

Measuring stability Link stability is the probability that two people who share a social link stay linked. While we cannot measure broken relationships directly, we can measure mobility. Presumably many friends are lost and new friends are made when people move frequently from one

[8] *http://www3.norc.org/gss+website/*

community to another. Therefore, we use the length of time a person lives in the same community as a proxy for their social network stability. We do not have a large cross-country panel of mobility data. But we do have extensive data on mobility for U.S. residents, including those born abroad. So our proxy for link stability in foreign countries is the average length of time that first-generation U.S. immigrants from that country have lived in their communities.

Stability data also comes from the GSS. The variable *stability* is constructed by first combining the survey answers to four similar questions asked about the length of time the respondent has lived in his/her present location. As for the degree variable, we regress this combined response on : (1) age, (2) income, (3) marital status, (4) dummies for education level, and (5) what fraction of the U.S. population shares the respondent's country of origin and use the residuals as the explanatory variable. Finally, we select respondents that report being born in a foreign country and average their residuals to form a *stability* variable for each reported ethnicity. See appendix B for more details. The part of stability (or lack of mobility) of first generation U.S. immigrants, that is not explained by demographic differences, is our proxy variable for social network stability is the country of origin. The underlying assumption here is that people who move to the U.S. from countries with stable social networks maintain higher degrees of social network stability than immigrants from less stable countries, after they immigrate to the U.S..

Measuring fractionalization Our theory tells us that we want to measure the number of groups in the country that are socially disconnected. One obvious reason why social connections may not be present is that the groups speak different languages or have different cultures. Therefore, we use data on ethnic fractionalization from Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003). Specifically, ethnic fractionalization is one minus the probability that two randomly selected agents in a country share the same culture. Thus, if all residents share the same culture, there is zero fractionalization. In the model, the probability of belonging to the same faction is decreasing in the number of factions f. Therefore, fractionalization, which is one minus that probability, is increasing in the number of factions f.

Of course, this measure does not tell us who is familiar with both cultures and can serve as a pathway for communication between factions. But it is well-known that even when many people can communicate in two languages, language borders create strong social divides between populations. The same is true for cultural boundaries. Just a slight preference for associating with culturally similar people leads to culturally segmented social groupings. Furthermore, the model has some bi-cultural nodes. There are the two nodes on each side of the factional divide who communicate with each other, despite being in different factions.

3.3 Measuring the Rate of Technology Diffusion

We use a technology diffusion measure that is derived from the cross-country historical adoption of technology data set developed by Comin, Hobijn, and Rovito (2006). The data covers the diffusion of about 115 technologies in over 150 countries during the last 200 years. At a country level, there are two margins of technology adoption: the "extensive" margin (whether or not a technology is adopted at all) and the "intensive" margin (how quickly a technology diffuses, given that it is adopted.) A country can be behind in a technology even though it is adopting it quickly, if the technology was introduced to the country late.

Since our model speaks only to the diffusion rate of a technology, i.e. its intensive margin of adoption, we need to filter the extensive margin from the data. We do this with the results from Comin and Mestieri (2012), where attention is restricted to 15 technologies. Technical details are in that paper, but the idea is the following: For a given country, plotting the normalized level of a given technology (e.g. log telephone usage minus log country income) over time yields an increasing curve. For a given technology, these curves look similar across countries, except for horizontal and vertical shifts. The horizontal shifts correspond to the extensive margin of technology adoption; if country A adopts telephones in exactly the same way as country B, only twenty years later, its curve will be identical to that of B except shifted twenty years to the right. However, if country A adopts telephones less vigorously but at the same time, its curve will be below that of B's. This diffusion rate of technology is what we are interested in, so it is what we focus on. Specifically, Comin and Mestieri (2012) estimate the slope of a non-linear diffusion curve. A higher slope parameter m_{ij} indicates a faster diffusion rate of technology j in country i.

A complication is that the diffusion data set is unbalanced; if data for a country is only available for slowly-spreading technologies, it might artificially appear technologically backward. To control for this problem, we estimate $m_{ij} = \alpha_j + e_{ij}$, where α_j is a technology-specific fixed effect. Our measure of technology diffusion for a given country is the average residual diffusion$_i = \sum_j e_{ij}$.

3.4 Concerns about instrument exogeneity

Our identifying assumption is that while technology diffusion and GDP may affect disease prevalence, even 40 years prior, it affects many diseases similarly. Likewise, the direct effect on GDP of different types of disease is also similar. Thus, the difference in the prevalence of two types of disease is exogenous with respect to GDP. The difference we consider is the difference between diseases that reside in humans (human-specific plus multi-host) and diseases that reside exclusively in non-human animals (zoonotic diseases).

Unequal variance One concern with this instrument might be that the difference between disease prevalence rates might not be orthogonal to the sum. For example, if zoonotic disease had (hypothetically) been eradicated in every country in our sample, then $\Delta germ = \bar{d} - z = \bar{d}$. Since the prevalence of disease is likely to be correlated with income and technology diffusion, this situation would render $\Delta germ$ an invalid instrument. For two variables x and y, $(x+y)$ is uncorrelated with $(x-y)$ when x and y have equal variances. Our human and zoonotic disease variables do not have exactly the same variance. To ameliorate this concern, we also use $\Delta germ\_std$ as an instrument in table 2 and find that it produces estimates of the importance of social networks that are even larger than the initial estimates.

Uneven effects of technology. Our empirical strategy is based on the assumptions that $\bar{d}$ and z have the same relationship with A but different relationships with $\tilde{N}$. One may think this relationship does not necessarily hold. For example, perhaps clean water initiatives are one of the first public health measures adopted when income rises. If this were the case, then there would be a negative correlation between zoonotic illness and technology diffusion, and therefore a positive correlation between (human - zoonotic) diseases ($\Delta germ$) and shocks to technology diffusion ϵ. If $E[\epsilon x] > 0$, how would this bias the results? A positive shock to income (high ϵ) would increase the difference in disease (x), which would decrease individualism $\tilde{N}$ (since we estimate $\beta_5 < 0$). This would induce negative correlation between A and $\tilde{N}$, which would lower the estimated coefficient β_2 in equation 2. So β_2 would be downward biased. Thus, if the instrument is invalid because economic development primarily reduces water-borne illnesses, then the true size of the network's effect on technology diffusion is even larger than what we estimate.

Social networks affect disease. The other hypothetical cause for concern might be that faster technology diffusion and the accompanying higher income cause the social structure to change. In particular, a richer, more modern society is more likely to be market-based and individualist. The change in network structure could affect the difference in disease prevalence by facilitating the transmission of diseases spread from human-to-human. Notice that this logic does not imply that differences in disease x are correlated with the estimation error ϵ in (2). This story suggests that social network structure $\tilde{N}$ depends on A, something already represented in our specification (equation 3), and it suggests that there should be an additional equation representing the idea that the instrument x depends on the network: $x = \psi_1 + \psi_2 S + \nu$. In this structure, as long as $e[\epsilon\nu] = 0$, x is still a valid instrument for $\tilde{N}$. In other words, as long as technology diffusion affects the difference in disease through networks, rather than directly, this form of reverse causality *does not invalidate the use of disease differences as instruments.* It only implies that β_5 is perhaps not an

unbiased estimator of the effect of disease on social institutions. Our estimates suggest that more disease is associated with less individualism. If individualism spreads disease, then this estimate is downwards-biased. In other words, the true effect of disease on social institutions would be larger than the one we estimate.

4 Empirical Results: How Much Do Networks Affect Technology?

4.1 First-Stage Regressions: Disease and Social Networks

We begin by investigating the relationship between our instruments and our measures of social network structure. There are two key findings: First, the instruments are significant predictors of social network networks. Second, the difference in diseases are positively correlated with collectivism, low degree, homophily and fractionalization. Although this effect is not identified, the correlation is consistent with the evolutionary network model.

In each specification, we use multiple instruments to evaluate the validity of each instrument by testing for orthogonality with the residual in equation (2). Following Hall and Jones (1999), we use two language-based variables as additional instruments to test the validity of our own disease-based instruments. The variable pronoun is a dummy variable that is equal to 1 if it is conventional to omit first- and second-person pronouns in a country's dominant spoken language (Kashima and Kashima, 1998). For example, English and German typically do not omit pronouns, while Spanish does. In addition, including the fraction of the population speaking English as a first language contributes additional explanatory power.[9] Because these variables are language-based, they are a product of the country's distant past and are unlikely to be affected by current income or technology. For three of the specifications, English and pronoun are not strong instruments for the social network variable. In order to avoid a weak instruments problem for these specifications, we use latitude and elevation as instruments instead.

We begin by exploring the data on individualism and disease prevalence. Figure 3 illustrates the clear, negative relationship between our network measure, the Hofstede index, and the sum of the prevalence of all nine pathogens in our historical disease data set. The negative relationship is consistent with our theory, in which greater disease prevalence favors the emergence of a collectivist network. Even though collectivism itself inhibits the spread of disease, the net prediction of the evolutionary model is that high pathogen prevalence is correlated with collectivism. This prediction is confirmed in figure 3 where more disease is correlated with lower individualism. Similarly, since networks with high degree, low link stability, and low fractionalization protect against disease transmission, high human disease $\bar{d}$ environments favor the emergence of such networks.

[9]The English variable is available from the Penn World Tables, Mark 5.6.

Figure 3: Hofstede's individualism index plotted against total pathogen prevalence. Total pathogen prevalence is $\bar{d} + z$. This is a sum of the prevalence of all nine diseases described in section 3.

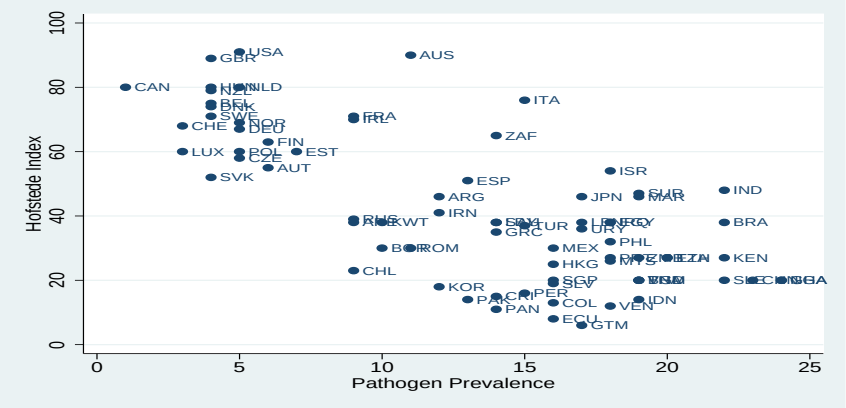


Table 8 quantifies these relationships. The negative correlations between disease and high-diffusion (low average path length) social networks is consistent with the evolutionary network theory. The explanatory power of pathogens can be large; the R^2 of the individualism regressions is over 50%. The economic magnitudes are also large. A one-unit increase in $\Delta germ$ corresponds to one human disease being endemic instead of sporadic. Having one more socially transmittable human disease consistently prevalent corresponds to an individualism index that is 3.77 points lower (16% of a standard deviation). For the stability and fractionalization measures, the instruments are weaker, the R^2 is lower, and sometimes only $\Delta germ$ is significant.

These results are important for the next stage, identifying an effect of institutions on technology diffusion. But they are also interesting on their own because they are consistent with one reason why countries may have adopted different social institutions. Perhaps social networks have evolved, in part, as a defense against the spread of directly-communicable diseases. But further statistical work needs to be done to say conclusively that disease prevalence is part of the reason why some societies have adopted social networks that inhibit technological diffusion and growth.

4.2 Main Results: Social Institutions and Technology Diffusion

Our main result is to quantify the effect of social networks on technology diffusion. Figure 4 illustrates the relationship between social network structure and the speed of technology diffusion in a scatter plot. It reveals that more individualist, higher degree, less socially stable, and less fractionalized societies tend to also be societies where technologies diffuse quickly. In interpreting this correlation, reverse causality is obviously a concern: The economic development that results from technology diffusion could produce a wave of urbanization, which influences social networks.

Table 4: **First-stage regressions of pathogen prevalence variables on individualism index**

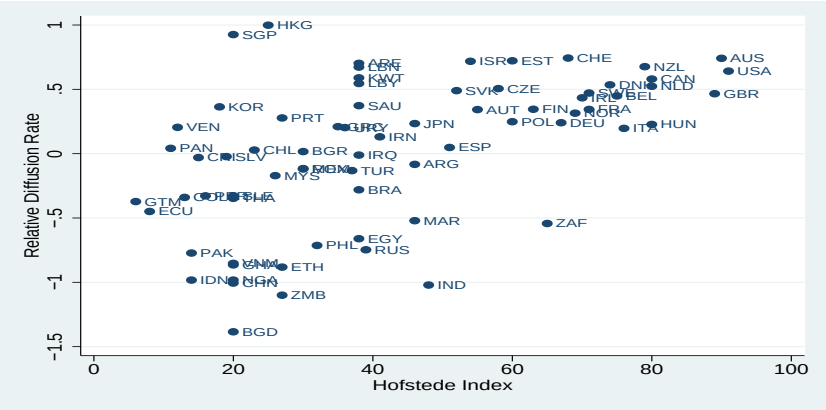
Dependent variable	Individualism		Degree		Stability		Fractionalization	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Human - zoonotic	-2.38		-7.98		1.24		2.76	
pathogens (Δgerm)	(0.48)		(11.36)		(0.94)		(0.79)	
Human - zoonotic		-7.12		-31.95		4.39		-1.38
pathogens (Δgerm_std)		(2.90)		(48.42)		(4.93)		(3.77)
English	23.05	24.90			-7.88	-8.00	7.16	
	(7.42)	(8.43)			(15.31)	(15.42)	(12.98)	
Pronoun	-23.14	-30.02			6.52	9.71	-0.72	
	(4.35)	(4.57)			(8.69)	(8.26)	(7.37)	
Latitude			539.20	615.75				-78.69
			(236.30)	(186.81)				(13.22)
Elevation			-2.06	-1.83				-0.57
			(3.97)	(3.94)				(0.31)
Constant	220.5	69.46	-50.35	-221.64	-33.75	-12.93	-24.80	67.98
R^2	0.72	0.64	0.22	0.22	0.07	0.06	0.17	0.41
Observations	62	62	51	51	71	71	71	67

The table reports OLS estimates of the γ coefficients in $\tilde{N} = \beta_3 + \beta_5 germ + \beta_7 English + \beta_8 Pronoun + \eta$. $\tilde{N}$ is the social network variable, which is Hofstede's individualism index (1)-(4), degree, as measured by the number of close friends reported by immigrants to the US in the GSS (x 100), link stability as measured by length of time US immigrants live in the same community (x 100), or ethnic fractionalization (x 100), from Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003). The *germ* variables are defined in equations (6) and (7). For definitions of the other instruments and more measurement details, see appendix B. Standard errors are in parentheses.

Therefore, we use the differences in pathogen prevalence and either language or geographic variables as instruments for social networks.

The first two columns of table 5 show that the degree of individualism in a country's network has a large effect on a country's rate of technology diffusion. A 1-standard deviation in individualism is 28.5. When we use $\Delta germ\_std$ as an instrument, a 1-standard deviation increase in individualism results in $28.5 \cdot 1.31 = 37.3\%$ increase in the speed of technology diffusion. The mean of the diffusion variable is near zero so this is not easily interpretable relative to its mean. But its standard deviation is 63.4%. Thus, a degree of individualism that is 1 standard deviation above the average is associated with technology diffusion that is 59% of a standard deviation higher than average. Across many specifications, the estimates of the effect of social network structure are remarkably stable. Individualism consistently explains 27-28% of the variation in technology diffusion rates. A 1-standard deviation increase in degree, stability and fractionalization affect technology diffusion by $[0.61, 0.66]$, $[-0.012, -0.0075]$, and $[-0.52, -0.54]$ respectively, depending to the instruments used. This represents roughly 1, 0.02 or $[0.8, 1.2]$ standard deviations of technology diffusion. The finding is that individualism, network degree and fractionalization have potentially large economic effects

Figure 4: Technology and individualism. Comin and Mestieri (2012)'s technology diffusion measure (vertical axis) plotted against Hofstede's individualism index (horizonal axis).



on technology diffusion. The magnitude of the network stability appears to be much smaller.

The Sargan test statistics (in the row labeled over-ID) are chi-square statistics for the test of the null hypothesis that the instruments are uncorrelated with the regression residual ϵ. For every IV specification, we cannot reject this null hypothesis at the 5% or even the 10% level. However, we could reject the null hypothesis at a 15% confidence level in the estimation in columns (1) or (4). But in many cases, the p-value suggests that the disease instrument is likely to be a valid instrument.[10]

Controlling for other possible explanatory variables. A natural question is whether social networks are simply a proxy for some other economic variable. To assess this, we choose a variety of other variables thought to explain technology adoption or income and control for their effects too. In doing so, we recognize that these control variables may themselves be endogenous. Inferring causality from these results would therefore be problematic. However, we continue to use Δ*germ_std*, *pronoun* and *english* as instruments, individualism as an explanatory variable, and add the following variables, one-by-one, to the first- and second-stage estimations:[11] Controlling for life expectancy at birth, which could capture a direct effect of pathogens on technology diffusion, or social infrastructure, which could promote technology diffusion and discourage disease, reduces the size of the coefficient on individualism by a factor of roughly 1/2. The other control variables we try are: (1) ethnic-linguistic fractionalization (the probability that two randomly matched people belong to different ethnic or linguistic groups), which could affect both social networks and the diffusion rate of technology, (2) latitude, which is likely to be correlated with the epidemiological

[10] We also computed Basmann statistics. They were quite close in value to the Sargan statistics in every instance.

[11] Our procedure and our choice of variables here largely follow Hall and Jones (1999). The variable "social infrastructure" is constructed by Hall and Jones to measure the quality of institutions.

Table 5: **Social Networks and Technology Diffusion (main result)**

	Technology Diffusion Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instrument:	Δgerm	Δgerm_std	Δgerm	Δgerm_std	Δgerm	Δgerm_std	Δgerm	Δgerm_std
Individualism	1.62	1.31						
	(0.33)	(0.34)						
Degree			26.41	24.44				
			(8.76)	(8.39)				
Stability					-4.25	-2.63		
					(2.10)	(1.67)		
Fractionalization							-2.89	-2.06
							(0.86)	(0.46)
Over-ID p-val	0.11	0.74	0.29	0.11	0.78	0.82	0.17	0.24
	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
N	62	62	67	67	71	71	67	67

Columns 1-4 report $100 * \beta_2$, columns 5-8 report β_2 coefficient from an IV estimation of $A = \beta_1 + \beta_2 \tilde{N} + \epsilon$. Technology diffusion rate (A) comes from the Comin and Mestieri (2012) measure of the intensive technology adoption in a country. Individualism $\tilde{N}$ is the Hofstede index. The variables $\Delta germ$, $\Delta germ\_std$ are defined in equations (6) and (7). Columns (1)-(2) and (5)-(7) also use *pronoun* and *english* as instruments, as defined in table 1. Rows (3)-(4) and (8) use *latitude* and *elevation* as instruments, in addition to the disease variables listed at the top. The over-ID test is a Sargan test statistic. The null hypothesis is that the instruments are uncorrelated with ϵ. Accept means that null hypothesis cannot be rejected at the 5% or even the 10% confidence level.

environment, (3) disease-adjusted life expectancy, to capture the direct effect of health on technology, (4) a country's degree of capitalism or socialism, which could be highly correlated with social network structure and which probably affects incentives for technology adoption, and (5) population density, which affects disease, networks and technology diffusion. These all leave the estimate of the effect of individualism largely unchanged. Appendix B reports the complete set of results for each of these estimations. In sum, there is a statistical relationship between social network structure and technology diffusion that is above and beyond that which comes from other commonly-used determinants of income.

Effect of social networks on income. To interpret these results economically, it is helpful to re-estimate the effect of social network structure with a dependent variable that is more familiar to macroeconomists: log real output per worker. The coefficients in Table 6 tell us that a 1-standard-deviation increase in the Hofstede index (42 units) increases log output per worker by 0.87, which represents an 87% increase. A 1-standard deviation increase in degree roughly doubles output: $2.54 * 0.43 = 1.09$. For stability and fractionalization, a standard deviation increase decreases output by 135% and 75% respectively.

To get a more concrete idea of what these numbers imply, let's compare the Netherlands and

Table 6: **Social Networks and Income per Worker**

Dependent variable:	Output per worker			
Individualism	2.07			
	(0.45)			
Degree		43.19		
		(13.23)		
Stability			-4.64	
			(2.14)	
Fractionalization				-2.89
				(0.92)
N	63	48	63	63

The entries are $100 * \beta_2$ in columns 1 and 2 and β_2 in columns 3 and 4 from an IV estimation of $Y/L = \beta_1 + \beta_2 S + \epsilon$, where Y/L is log (RGDP) per worker and $\tilde{N}$ are the measures of social network structure. Y/L data come from the Penn World Tables mark 5.6. The instruments are English, Kashima, and $\Delta germ$, as in (6). For the degree estimation, the instruments Latitude and Elevation replace the language instruments.

Ghana. The Netherlands is individualistic ($Indiv = 80$), with large degree reported social networks (-0.35), has unstable social networks (-0.21) and a low degree of ethnic fractionalization (0.11). Ghana is collectivist ($Indiv = 20$), with smaller social networks (-3.06), has moderately stable social networks (-0.04), and a high degree of ethnic fractionalization (0.67). The difference in individualism is $80 - 20 = 60$, while the difference in output per worker between the Netherlands and Ghana is 2.74. Using the coefficient in column 1 of Table 6 the difference in individualism explains $60 * 0.021 = 1.26$ of this difference, which is over one-third of the per worker output gap. The difference in network degree (-2.71) explains a 117% gap in income. The difference in social network stability (0.17), suggests that the Netherlands should be 80% richer. The difference in ethnic fractionalization explains a $0.56 * 2.89 = 160\%$ gap in output per worker. The conclusion is that, while none explain the entire difference in income between these two countries, each of these network features appears to have economically large effects.

4.3 Could Social Networks Really Change in Response to Disease?

The idea that people might choose their social circles based on disease avoidance might sound far-fetched. But researchers in animal behavior have long known that other species choose their mates with health considerations in mind (Hamilton and Zuk, 1982). Furthermore, primate research has shown that the animals most similar to human beings behave similarly to the agents in our model. Their mating strategies, group sizes, social avoidance and barriers between groups are all influenced by the presence of socially transmissible pathogens (Loehle, 1995).

One might also question whether historical societies knew enough about contagion to make in-

formed choices about social networks. Yet, historical documents reveal a reasonable understanding of epidemiology. For example, in the sixteenth century, when smallpox reached the Americas and became a global phenomenon, people understood that the skin lesions and scabs that accompany smallpox could transmit the disease. They knew that survivors of smallpox and other infections were immune to re-infection. The practice of inoculation, whereby people were intentionally exposed to disease was practiced hundreds of years ago in China, Africa and India. Similarly, the plague was recognized to be contagious. Therefore, control measures focused primarily on quarantine and disposal of dead bodies. Even two thousand years ago, in biblical times, leprosy was understood to be contagious. Lepers, or suspected lepers, were forced to carry a bell to warn others that they were coming. Thus, the idea that one should avoid contact with others who carry particular contagious diseases is not just a modern idea.

5 Conclusions

Measuring the effect of social network structure on the economic development of countries is a challenging task. Networks are difficult to measure and susceptible to problems with reverse causality. We use a theory of social network evolution to identify properties of social networks that can be matched with data and to select promising instrumental variables that can predict network structure. The theory predicts that societies with higher disease prevalence are more likely to adopt low-diffusion social networks. Such networks inhibit disease transmission, but they also inhibit idea transmission. This model reveals which social features should speed or slow diffusion. It also suggests that disease prevalence might be a useful instrument for a social network because affects how social networks evolve.

Quantifying the model reveals that small initial differences in the epidemiological environment can give rise to large differences in network structure that persist. Over time, these persistent network differences can generate substantial divergence in technology diffusion and output. We find evidence of this social network effect in the data. Exploiting the differential mode of transmission of germs, we are able to identify the significant effect of social network structure on technology diffusion and income. More broadly, the paper's contribution is to offer a theory of the origins of social institutions, propose one way these institutions might interact with the macroeconomy, and show how to quantify and test this relationship.

References

Acemoglu, D., and S. Johnson (2005): "Unbundling Institutions," *Journal of Political Economy*, 113, 949–995.

Acemoglu, D., S. Johnson, and J. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91(5), 1369–1401.

——— (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distributions," *Quarterly Journal of Economics*, CXVII(4), 1231–1294.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): "Fractionalization," *Journal of Economic Growth*, 8(2), 155–194.

Algan, Y., and P. Cahuc (2007): "Social attitudes and Macroeconomic performance: An epidemiological approach," Paris East and PSE Working Paper.

Ashraf, Q., and O. Galor (2012): "Human Genetic Diversity and Comparative Economic Development," *American Economic Review*, forthcoming.

Bisin, A., and T. Verdier (2001): "The Economics of Cultural Transmission and the Evolution of Preferences," *Journal of Economic Theory*, 97(2), 298–319.

Centola, D., and M. Macy (2007): "Complex Contagion and the Weakness of Long Ties," *Americaon Journal of Sociology*, 113 (3), 702–734.

Coleman, J. (1988): "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, 94, S95–S120.

Comin, D., B. Hobijn, and E. Rovito (2006): "Five Facts You Need to Know About Technology Diffusion," NBER Working Paper 11928.

Comin, D., and M. Mestieri (2012): "An Intensive Exploration of Technology Diffusion," HBS Working Paper.

Conley, T., and C. Udry (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100(1), 35–69.

Durlauf, S., and W. Brock (2006): "Social Interactions and Macroeconomics," in *Post-Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, ed. by D. Colander. New York: Cambridge University Press.

Fernández, R., A. Fogli, and C. Olivetti (2004): "Mothers and Sons: Preference Formation and Female Labor Force Dynamics," *Quarterly Journal of Economics*, 119(4), 1249–1299.

Fischer, C., and Y. Shavit (1995): "National Differences in Network Density: Israel and the United States," *Social Networks*, 17(2), 129145.

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

Gorodnichenko, Y., and G. Roland (2011): "Culture, institutions and the wealth of nations," University of California at Berkeley Working Paper.

Granovetter, M. (1973): "The Strength of Weak Ties," *American Journal of Sociology*, 78, 1360–1380.

——— (2005): "The Impact of Social Structure on Economic Outcomes," *The Journal of Economic Perspectives*, 19(1), 33–50.

Greenwood, J., A. Seshadri, and M. Yorukoglu (2005): "Engines of Liberation," *Review of Economic Studies*, 72(1), 109–133.

Greif, A. (1994): "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy*, 102, 912–950.

Grinstead, C. M., and J. L. Snell (1997): *Introduction to Probability*. Russell Sage, second edn.

Gudykunst, W., G. Gao, K. Schmidt, T. Nishida, M. Bond, K. Leung, and G. W. and (1992): "The Influence of Individualism Collectivism, Self-Monitoring, and Predicted-Outcome Value on Communication in Ingroup and Outgroup Relationships," *Journal of Cross-Cultural Psychology*, 23(2), 196–213.

Hall, R., and C. Jones (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?," *Quarterly Journal of Economics*, 114, 83–116.

Hamilton, W., and M. Zuk (1982): "Heritable True Fitness and Bright Birds: A Role for Parasites?," *Science*, 218, 384–387.

Hofstede, G. (2001): *Culture's consequences : comparing values, behaviors, institutions, and organizations across nations*. Sage Publications, second edn.

Jackson, M. (2008): *Social and Economic Networks*. Princeton University Press, first edn.

Kashima, E., and Y. Kashima (1998): "Culture and Language: The Case of Cultural Dimensions and Personal Pronoun Use," *Journal of Cross-Cultural Psychology*, 29(3), 461–486.

Loehle, C. (1995): "Social Barriers to Pathogen Transmission in Wild Animal Populations," *Ecology*, 76(2), 326.

Lucas, R., and B. Moll (2011): "Knowledge Growth and the Allocation of Time," NBER Working Paper 17495.

Munshi, K. (2004): "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 73, 185–213.

Murray, D., and M. Schaller (2010): "Historical Prevalence of Infectious Diseases Within 230 Geopolitical Regions: A Tool for Investigating Origins of Culture," *Journal of Cross-Cultural Psychology*, 41(1), 99–108.

Newman, M. (2010): *Networks: An Introduction*. Oxford University Press.

Oberfield, E. (2013): "Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture," Chicago Federal Reserve working paper.

Perla, J., and C. Tonetti (2011): "Endogenous Risk and Growth," NYU working paper.

Rauch, J., and A. Casella (2001): *Networks and Markets*. Russell Sage, first edn.

Rodenwaldt, E., and H. Jusatz (1961): *World Atlas of Epidemic Diseases, 1952-1961*. Falk Verlag, Hamburg.

Simmons, J., T. Whayne, G. Anderson, and H. Horack (1945): *Global Epidemiology*. Lippincott, first edn.

Smith, K., D. Sax, S. Gaines, V. Guernier, and J.-F. Gugan (2007): "Globalization of Human Infectious Disease," *Ecology*, 88(8), 1903–1910.

Society, N. G. (2005): *Atlas of the World*. Eighth edn.

Spolaore, E., and R. Wacziarg (2009): "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2), 469–529.

Tabellini, G. (2010): "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, 8(4), 677–716.

Taylor, C., and M. Hudson (1972): *World Handbook of Political andSocial Indicators*. Yale University Press (New Haven), first edn.

Thornhill, R., C. Fincher, D. Murray, and M. Schaller (2010): "Zoonotic and Non-Zoonotic Diseases in Relation to Human Personality and Societal Values: Support for the Parasite-Stress Model," *Evolutionary Psychology*, 8(2), 151–169.

Watts, D., and S. Strogatz (1998): "Collective dynamics of small-world networks," *Nature*, 393, 409–410.

Young, P. (2009): "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *The American Economic Review*, 99(5), 1899–1924.

A Proofs of Propositions

Proof of Result 1 *Average lifetime.* Suppose $\psi_k(0) = 1$ for some k and $\psi_j(0) = 0 \, \forall \, j \neq k$. For a person living in location j, the sick person lives s_{jk} steps away. Since the probability of contagion is equal to 1, person j will be sick in s_{jk} periods and then die, i.e. $\Psi_j(0) = s_{jk}$. Averaging over all locations j, we have that the average lifetime is equal to the average path length from k to all other nodes: $E_j[\Psi_j(0)] = E_j[s_{jk}]$.

What if the probability of transmission is less than one? Note that if the probability of disease transmission is less than one, then there is a positive probability that the disease dies out before it is spread to anyone. Since there is no other source of death, this implies that lifetime is infinite. With a positive probability of infinite lifetime, $E_j[\Psi_j(0)] = \infty, \, \forall \, \pi < 1$.

Average discovery time. Analogously, suppose that a new idea is introduced by person k in period 0. Since the idea is transmitted with probability 1, the number of periods it takes to reach person j is given by $\alpha_j(0) = s_{jk}$. Thus the average discovery time is equal to the average path length from k to other nodes, $E_j[\alpha_j(0)] = E_j[s_{jk}]$.

Proof of Result 2 A new technology shock advances the technological frontier if it arrives to an agent that has a technology level that is as high as any other agent in the network. Suppose that at t, the technology of each agent is the same in both types of networks and agent j (and only him[12]) is at the technological frontier. In the next period, with probability $1-(1-p)^4$, agent j transmits his technology to at least one of his connections and the expected number of people that have the latest technology in $t+1$ is $1+4p$. That probability is the same in both networks. Each agent has an identical probability λ of inventing a new technology. Thus, the probability that a technology shock hits an agent who has the highest technology level at $t+1$ and advances the frontier is $(1+4p)\lambda$, in either network.

Now consider time $t+2$. In expectation, $1+12p$ people have the latest technology in N2 but only $1+8p$ in N1. Thus the probability of moving the frontier is $\lambda(1+12p)$ in N2. That probability is larger than the same probability in N1, which is given by $\lambda(1+8p)$. Continue in this fashion until every agent in the network has acquired such level of technology. At that point, all agents have the same level of technology and the probability of advancing the frontier is again equal in both networks. In every period, we find that the probability of advancing the technological frontier is weakly higher in N2 than in N1, with strict inequality in at least one period. Therefore, we conclude that the probability of a technology shock moving the frontier in N2 is than the probability of moving the frontier in N1.

Results for networks 1 and 2 *In a collectivist network, where $\gamma = 4$, there are n unique collectives.*

Claim 1: Any three adjacent nodes are a collective.

Proof: Consider nodes j, $j+1$ and $j+2$. Since every node is connected to its adjacent nodes, $j+1$ is connected to j and $j+2$. And since every node is also connected to nodes 2 places away, j is connected to $j+2$. Since all 3 nodes are connected to each other, this is a collective.

Claim 2: Any sets of 3 nodes that are not 3 adjacent nodes are not a collective.

Proof: Consider a set of 3 nodes. If the nodes are not adjacent, then two of the nodes must be more than 2 places away from each other. Since in a collectivist network with $\gamma = 4$, nodes are only connected with other nodes that are 2 or fewer places away, these nodes must not be connected. Therefore, this is not a collective.

Thus, there are n unique sets of 3 adjacent nodes (for each j there is one set of 3 nodes centered around j: $\{j-1, j, j+1\}$). Since every set of 3 adjacent nodes is a collective and there are no other collectives, there are n collectives in the network. □

In an individualistic network, where each person i is connected to $i-\psi$, $i-1$, $i+1$, and $i+\psi$, where $\psi > 2$, there are zero collectives.

Proof: Consider each node connected to an arbitrary i, and whether it is connected to another node, which is itself connected to i. In addition to being connected to i, node $i-\psi$ is connected to $i-2\psi$, $i-\psi-1$, and $i-\psi+1$. None of these is connected to i. Node $i-1$ is also connected to $i-2$, $i-\psi-1$ and $i+\psi-1$. But none of these is connected to i. Node $i+1$ is also connected to $i+2$, $i-\psi+1$ and $i+\psi+1$. But none of these is connected to i. Finally, node $i-\psi$ is also connected to $i+\psi-1$, $i+\psi+1$ and $i+2\psi$. But none of these is connected to i. Therefore, there are no collectives among any connections of any arbitrary node i. □

Proof of result 3 We start by deriving the path length in each network and then compare the two.

Average path length in network 1. Consider the distance from the last node, n. n can be connected to nodes 1 though $\gamma/2$ and $n-1$ through $n-\gamma/2$ in 1 step. More generally, it can be connected to nodes $(s-1)\gamma/2+1$ through $s\gamma/2$ and $n-(s-1)\gamma/2-1$ through $n-s\gamma/2$, in s steps. For each s, there are γ nodes for which the shortest path length to n is s steps. We know from result 1 that when γ is even and n/γ is an integer, the longest path length (the diameter) is n/γ. Thus, the average length of the path from n to any other node is $1/n\sum_{s=1}^{n/\gamma}\gamma s$. By symmetry, this is the same average distance from any node to others. Using the summation formula, this is $(\gamma/n)(n/\gamma)(n/\gamma+1)/2 = 1/2+n/(2\gamma)$.

Average path length of network 2. Consider the path length from node n to any other node in the network between 1 and $n/2$. By symmetry, the path length starting from any other node is the same and average the path length to the nodes between $n/2$ and n is the same as for the nodes in the first half. Consider taking steps on length m until one reaches or passes the node $n/2-m/2$. The number of steps in this path is $\tilde{m} \equiv \text{round}(n/(2m))$, where round is the nearest integer value. All points not on this path (interior nodes) can be reach by steps of length 1 from the nearest multiple of m. To reach all of these interior nodes with a step of length 1, from the path of m multiples requires $m/2$ steps. Thus, one can reach all the nodes between $(s-1/2)m$ and $(s+1/2)m$ in, at most, $s+m/2$ steps.

This implies a sequence of path lengths of the following form:

$$\{1, \ldots, \frac{m}{2}\}$$

[12]The reasoning is analogous if more than one agent receives the original shock at the same time.

$$1 + 2\{1, \ldots, \frac{m}{2}\}$$

$$\vdots$$

$$\tilde{m} + 2\{1, \ldots, \frac{m}{2}\}$$

This is an upper bound on the total path lengths of the network because $\tilde{m}$ may be greater than $n/2 - m/2$. The average path length is the sum of all path lengths, divided by the number of nodes. In this case, that is

$$PL \leq \frac{1}{n/2}\left[\sum_{i=1}^{\tilde{m}} i + (2\tilde{m}+1)\sum_{i=1}^{m/2} i\right]$$

We can use the summation formula to replace the sums.

$$PL \leq \frac{2}{n}\left[\frac{\tilde{m}(\tilde{m}+1)}{2} + (2\tilde{m}+1)\frac{m/2(m/2+1)}{2}\right]$$

Note also that any number is rounded down by, at most, 1/2. Therefore, an upper bound on $round(x)$ is $x+1/2$. Similarly, we know that $\tilde{m} \leq n/(2m) + 1/2$. Since the path length expression is increasing in $\tilde{m}$,

$$PL \leq \frac{1}{4n}\left[\frac{(n+m)(n+3m)}{m^2} + \frac{n+2m}{m}m(m+2)\right]$$

Comparing path lengths. A sufficient condition for the individualist path length to be smaller is

$$\frac{1}{4n}\left[\frac{(n+m)(n+3m)}{m^2} + \frac{n+2m}{m}m(m+2)\right] < \frac{n}{8}$$

Rearranging, this implies that

$$\left(\frac{1}{2} - \frac{1}{m^2}\right)n^2 - \left(\frac{4}{m} + m + 2\right)n - 2m(m+2) - 3 > 0$$

Since we assumed that $m > 2$, the coefficient on the n^2 term is positive. Therefore, there is a sufficiently large n such that the inequality holds.

Proof of result 4 (higher degree speeds diffusion) Take a network and its matrix of shortest path lengths $\{p_{ij}\}_{i,j=1}^{N}$. For one node i, decrease its degree γ by 2, by breaking the two farther links, the links to nodes $j \pm \gamma/2$. Then the shortest path length between nodes i and $j \pm \gamma/2$ increases by one. Furthermore, breaking these two links can only increase the shortest path length(s) for any other node $j \neq i$. Therefore, for the new matrix of shortest path lengths $\{\tilde{p}_{ij}\}_{i,j=1}^{N}$, $p_{ij} \leq \tilde{p}_{ij}$ for all i, j. The average path length (1) increases.

Proof of result 5 (Link stability slows diffusion) **Step 1:** We first prove the following lemma which will be an important step in proving the result.

Lemma 1 *For any argument z,*

$$\frac{\partial tanh^{-1}z}{\partial z} = \frac{1}{1-z^2}$$

Proof: Note that if $y = tanh^{-1}z$, then $z = tanhy$. It is a standard result that, $dz/dy = sech^2 y$ (See e.g., CRC Standard Mathematical Tables (1964) p.338). Using the inverse function rule, this implies that

$$dy/dz = 1/sech^2 y.$$

Note that a property of sin and cos is that $cosh^2 y + sinh^2 y = 1$. Dividing this equality by $cosh^2$ on both sides yields $sech^2 y = 1 - tanh^2 y$. Therefore, we have

$$dy/dz = 1/(1 - tanh^2 y)$$

But since we know that $x = tanhy$, $tanh^2 y = x^2$. Thus,

$$dy/dz = \frac{1}{1-z^2}.\square$$

Step 2: Next, we prove the following result for the static small world network and then prove an equivalence between our dynamic network 5 and the small world network 4.

Result 9 *If $n\gamma\tilde{p} \geq 4\sinh^2(1)$, then at any given time $t > 0$, the expected average path length of network 4 is a decreasing function of the rewiring probability.*

Newman (2010) considers a static network with n nodes, where each node is connected to its γ closest neighbors. In addition, for each link, there is a probability $\tilde{p}$ that the link is broken and an additional link is formed. This new link connects each unconnected pair of nodes with equal probability. Thus, the expected number of links is $x \equiv n\gamma\tilde{p}$. Newman (2010) proceeds to argue that a mean-field approximation to the path length Ω of this network is

$$\Omega = \frac{n}{\gamma}\frac{2}{\sqrt{x^2+4x}}tanh^{-1}\sqrt{\frac{x}{4+x}} \tag{8}$$

which is a good approximation numerically to the true path length, for small rewiring probabilities p. Since the presence of links between neighboring nodes has little effect on the average path length, the behavior of the small-p Watts and Strogatz (1998) model is almost identical to the other commonly used formulation, where links are not broken. Instead, new random links are added to the ring network (Newman, 2010).

Using lemma 1 and the product rule, we can compute the first derivative of the path length in the rewiring probability:

$$\frac{\partial\Omega}{\partial x} = \frac{-2n}{\gamma}\frac{(2x+4)}{2(x^2+4x)^{3/2}}tanh^{-1}\sqrt{\frac{x}{x+4}} + \frac{2n}{\gamma\sqrt{x^2+4x}}\frac{1}{1-\frac{x}{x+4}}\frac{1}{2}\left(\frac{x}{x+4}\right)^{-1/2}\frac{4}{(x+4)^2} \tag{9}$$

$$= \frac{-n(2x+4)}{\gamma(x^2+4x)^{3/2}}tanh^{-1}\sqrt{\frac{x}{x+4}} + \frac{n}{\gamma(x^2+4x)} \tag{10}$$

$$= \frac{n}{\gamma(x^2+4x)}\left[1 - \frac{2x+4}{\sqrt{x^2+4x}}tanh^{-1}\sqrt{\frac{x}{x+4}}\right] \tag{11}$$

This is negative iff

$$(2x+4)tanh^{-1}\sqrt{\frac{x}{x+4}} > \sqrt{x^2+4x} \tag{12}$$

Solving this explicitly for x is not feasible. But we can easily derive a sufficient condition for it to hold. Note that n, $\tilde{p}$ and γ are all positive variables and $tanh^{-1}$ is positive for all positive arguments. Thus $(2x+4)tanh^{-1}(\cdot) > (x+4)tanh^{-1}(\cdot)$. Similarly, since $x > 0$, the fraction $\sqrt{(x+4)/x} > 1$. Therefore, $\sqrt{(x+4)/x}\sqrt{x^2+4x} > \sqrt{x^2+4x}$. Note that the left side is equal to $(x+4)$, implying that $x+4 > \sqrt{x^2+4x}$.

So, if $(x+4)tanh^{-1}\sqrt{\frac{x}{x+4}} \geq x+4$, then since the left side is strictly less than the left side of (12) and the right side is strictly greater than the right side of (12), then this is a sufficient condition for (12) to hold. Dividing by $x+4$ on both sides yields

$$tanh^{-1}\sqrt{\frac{x}{x+4}} \geq 1.$$

Since the $tanh^{-1}$ function is monotone increasing, this implies

$$\sqrt{\frac{x}{x+4}} \geq tanh(1)$$

Solving for x delivers

$$x \geq \frac{4tanh^2(1)}{1-tanh^2(1)}.$$

Recall from lemma 1 that $1 - tanh^2(1) = sech^2(1)$. Tangent and secant functions are defined as $tanh = sinh/cosh$ and $sech = 1/cosh$. This means that $tanh/sech = shinh$. Furthermore, recall that $x = n\gamma\tilde{p}$. Thus, the sufficient condition becomes

$$\text{If} \quad n\gamma\tilde{p} \geq 4sinh^2(1) \quad \text{then} \quad \frac{\partial\Omega}{\partial\tilde{p}} < 0.$$

Step 3: Map network 5 onto network 4. In network 4, the expected number of shortcuts s is the number of links $2n$ times the probability $\tilde{p}$ that each link generates a shortcut: $s = 2n\tilde{p}$.

The steady state of network 5 is the state where the expected number of shortcuts is constant. The expected number of new shortcuts formed in each period is the number of existing links, which includes the ring lattice links, plus existing shortcuts, $2n + s$, times the link formation probability: $(2n + s)p$. The number of shortcuts lost each

period is the current number of shortcuts s, times the rate of shortcut decay: sz. Equating these two yields the expected steady-state number of shortcuts: $s = 2np/(z-p)$.

Equating these two expressions, we find that the expected number of links is the same when $p = \tilde{p}z/(1-z)$. If this equality holds, then the static small-world network is a ring lattice, with a uniform distribution of all possible shortcuts. At each date t, our network 5 is also a ring lattice with a uniform distribution over all possible shortcuts. Each network has the same uniform probability of forming a shortcut. Thus, these networks are equivalent, in the sense that they are drawn from the same distribution of random networks. Therefore, they must have the same average path length. Since this average path length is a decreasing function of $\tilde{p}$ and p is a linear, increasing function of $\tilde{p}$, the average path length must also be a decreasing function of p.

Proof of result 6 (Factions slow diffusion) The first part of the proof is a lemma which considers what happens to the expected average path length of a network if we start from a small world network and rewire one link. Rewiring means breaking one shortcut and forming a new shortcut somewhere else. Suppose the shortcut that is broken is a long link (meaning that without the shortcut, the path length is long) and the new link that is created is a short link (meaning that before the shortcut is formed, the path length was not as long). Then the lemma shows that the rewiring increases the expected average path length of the network.

Lemma 1: Consider two random networks. Both are small world networks, meaning that they are a ring lattice with degree γ, with additional links (shortcuts) uniformly distributed among all nodes not connected by the ring. In network N^A, nodes i and j are linked $n^A_{ij} = 1$, but nodes i and k are not $n^A_{ik} = 0$. In network N^B, all links are identical to N_A, except that $n^B_{ij} = 0$ and $n^B_{ik} = 1$. If the path length $p^A_{ik} < p^B_{ij}$, then the expected average path length in network B is longer than in network A: $\bar{p}^B > \bar{p}^A$.

The average path length is $\bar{p} = 1/N^2 \sum_{i,j} p_{ij}$. Since path lengths are symmetric $p_{ij} = p_{ji}$, and $p_{ii} = 0\ \forall i$, we can rewrite $\bar{p} = 2/N^2 \sum_{j>i} p_{ij}$.

Severing one link between i and j affects the path length p_{ij} as well as the lengths of all the paths p_{mn} that passed through i and j. Severing link i,j increases p_{ij} from 1 to p^B_{ij}. It increases $p_{i,j+1}$ and $p_{i,j-1}$ from 2 (or 1 with a small probability p) to at least $p^B_{ij} - 1$. If the network B path length was less, then there would exist a path through $p_{i,j+1}$ or $p_{i,j-1}$ that is shorter than p^B_{ij}, which would contradict p^B_{ij} being the shortest path length. By the same argument, the path length of all links $p_{i,j+q}$ and $p_{i,j-q}$ increases to be at least $p^B_{ij} - 2q/\gamma$. Otherwise, there would exist a path from i to j shorter than p_{ij}.

The number of links that connect to i and that increase in path length when link i,j is eliminated is at least $(\tilde{q}-1)$ nodes on each side of j, where $\tilde{q} \equiv \min q : 1 + 2q/\gamma = p^B_{ij} - 2q/\gamma$. On one side, this is exactly the number of links that increase in path length, to the nearest integer. On the other side, the number of paths than lengthen may be longer, depending on the location of the nearest shortcut. In other words, the number of nodes that lengthen their path on one side is $(\tilde{q}-1)$, while the number of nodes on the other is $(\tilde{q}-1)+\epsilon^B$. Notice that this is increasing in p^B_{ij}.

Conversely, when the new link is formed between i and k, the path length between i and k falls from p^A_{ik} to 1. The path length to the neighboring nodes $k+1$ and $k-1$ falls from $[p^A_{i,k}+1, p^A_{i,k}-1]$ to 2, with probability 1-p that there is no shortcut between i and that link and otherwise to 1. By the same argument, the path length of all links $p_{i,j+q}$ and $p_{i,j-q}$ falls from at least $p^A_{ik} - 2q/\gamma$ to some length not longer than $1 + 2q/\gamma$.

The number of links that decrease in path length when link i,k is added is at least $2(\dot{q}-1)$ where $\dot{q} \equiv \min q : 1 + 2q/\gamma = p^A_{ik} - 2q/\gamma$. On one side, $(\dot{q}-1)$ will be the number of paths that decrease in length, up to an integer. On the other side, it will be $(\dot{q}-1)+\epsilon^A$. Notice that the number of paths that shorten is increasing in p^A_{ik}. Because links are uniformly distributed, the probability that the next shortcut is ϵ spaces away is $(1-p)^\epsilon$ for both networks. Thus, $E[\epsilon^A] = E[\epsilon^B]$.

Since we assumed that $p^A_{ik} < p^B_{ij}$, it means that when we switch from network A to B, there are more links that increase in path length, then the number that decrease in path length, in expectation. Furthermore, because $p^A_{ik} < p^B_{ij}$, for every path that decreases in length ($p^B_{il} - p^A_{il} < 0$), there is a path that increases in length by more: $p^A_{im} - p^B_{im} < p^B_{il} - p^A_{il}$. Therefore, $\bar{p}^A - \bar{p}^B = 1/N^2 \sum_{i,j} p^A_{ij} - p^B_{ij} < 0$. This proves that $\bar{p}^B > \bar{p}^A$.

Step 2: Consider a small world network. Show that the expected path length between two uniformly-chosen nodes inside the same faction is smaller than the expected path length for nodes chosen uniformly from the entire network.

Consider two random nodes i and j, $i \neq j$, each chosen with a uniform probability from the ring of N nodes. With probability $\gamma/(n-1)$, the nodes are linked directly through the ring lattice. With an additional probability $p/(n-\gamma-1)$, the two nodes are linked by a shortcut. Thus $Pr(p_{ij} = 1) = \gamma/(n-1) + p/(n-\gamma-1)$. Similarly, with

probability $\gamma/(n-1)$, the nodes are two steps away on the ring lattice. Additionally, if i has a shortcut to any of j's γ neighbors or j is connected by a shortcut to any of the γ neighbors of i, then the path length between i and j is also not larger than 2. Thus $Pr(p_{ij} \leq 2) = \gamma/(n-1) + 2\gamma p/(n-\gamma-1)$. We can continue in this fashion to compute the probability of each path length between i and j.

Now, consider two random nodes i and k, $i \neq k$, each chosen with a uniform probability from the same faction $f(i)$. The probability of shortcuts is uniform on the ring and is therefore the same as before. But, conditional on being in a faction of size n/F, the probability of being linked by the ring lattice is approximately $\gamma/(n/F-1)$. This is an approximation because we are ignoring the small probability that i is on the boundary of faction F and therefore has fewer than γ neighbors in the same faction. For large n/F, this probability goes to zero. Since the size of the faction must be smaller than the size of the ring, $n/F < n$ and $\gamma/(n/F-1) > \gamma/(n-1)$. Thus, $Pr(p_{ik}=1) > Pr(p_{ij}=1)$. Similarly, the probability of being two steps away on the ring lattice is approximately, $\gamma/(n/F-1)$. Since $\gamma/(n/F-1) > \gamma/(n-1)$ and the probability of being connected in two steps by a shortcut is equal to the probability above, $Pr(p_{ik} \leq 2) > Pr(p_{ij} \leq 2)$. Continuing in the same fashion, we can sign $Pr(p_{ik} \leq q) > Pr(p_{ij} \leq q)$ for all $q < n$. Therefore, $E[p_{ik}] < E[p_{ij}]$

Step 3: Show that the expected average path length is longer in a fractionalized network with $F > 1$ than in a small world network ($F = 1$).

Start with a small world network, with shortcuts uniformly distributed over the whole ring lattice. We can construct the network with 2 factions by sequentially breaking all shortcuts that cross faction boundaries and for each broken link, creating a new link that connects two nodes in the same faction. Consider the first shortcut rewired, since the two nodes in different factions have a higher probability of having a longer path length, and for each path length, there is stochastic dominance of probabilities of a path length at least that short, this rewiring will increase expected path length $E[\bar{p}]$.

Now, the remaining network is no longer a small world network because links are no longer uniformly distributed. Instead, there is now a higher probability of a shortcut connecting two nodes inside a faction than across a faction. This lowers $E[p_{ik}]$ $\forall i,k : f(i) = f(k)$ and raises $E[p_{ij}]$ $\forall i,j : f(i) \neq f(j)$. When the next link is rewired, the probability that the path length after the link is broken exceeds the path length before the new link is formed ($\bar{p}^B > \bar{p}^A$) is higher. Therefore, the second and all subsequent re-wirings also raise $E[\bar{p}]$. Thus, the expected path length increases in the fractionalized network with $F > 1$.

Step 4: Let $E[\bar{p}^F]$ be the expected average path length of a network with F factions. Show that $E[\bar{p}^{\alpha F}] > E[\bar{p}^F]$, where $\alpha > 1$ is an integer.

Starting from a network with F factions, the network with αF factions can be created by dividing each existing faction into α new factions, breaking all shortcuts that cross the new faction boundaries, and rewiring those shortcuts so that they connect i and j: $f(i) = f(j)$ in the new faction set. Using the same argument as above, this procedure increases the expected network path length with each rewiring. Thus, $E[\bar{p}^{\alpha F}] > E[\bar{p}^F]$.

Proof of result 7 (Network becomes homogenous) Observe that the state where all agents have the same type is absorbing. We will show that such state can be reached from any state with positive probability and therefore the process will be absorbed with probability 1 (by Lemma 1).

Lemma 2 *In an finite Markov chain that is absorbing (it has at least one absorbing state and from every state it is possible to go to an absorbing state), the probability that the process will be absorbed is 1. For proof see Grinstead and Snell (1997).*

Suppose agent j is the only one whose type is different to the rest of the network. The number of j-types increases in the next period if: (i) agent j survives, (ii) all the nodes directly connected to agent j die (first tier nodes) and (iii) all the nodes connected to the nodes connected directly to agent j also die (second tier nodes). To see this, index the first tier connections with i and let $k^*(i) = argmax_{\{k:\eta_{ik}(t)=1\}} A_k(t)$. By assumption, if i dies at t, we have $\tau_i(t+1) = \tau_{k^*(i)}(t)$. Then if the three situations described happen, we have that $k^*(i) = argmax_{\{k:\eta_{ik}(t)=1\}} A_k(t) = argmax\{A_j(t), 0\} = j$ $\forall i$. Therefore $\forall i$ we have $\tau_i(t+1) = \tau(k^*(i)) = \tau_j(t)$.

Now we compute a lower bound for the probability of (i)-(iii) happening at any time. First, assume $\tau_j(t) = co$. Recall that j's own type governs the links to the right and others' types govern links to the left, so in this case the first tier connections for which $\eta_{jk} = 1$ are $k = \{j-4, j-1, j+1, j+2\}$. The second tier connections (nodes connected to j's connections that are not directly connected to j) are the following: $\{j-8, j-5, j-3, j-2, j+3, j+5, j+6\}$. Therefore, with probability of at least $(1-z)z^{11}$ node j survives and all his first and second tier connections have an

accident and die, reaching the absorbing state.[13]. Second, if we assume that $\tau_j(t) = in$, then his direct connections are $\eta_{jk} = 1$ for $k = \{j-2, j-1, j+1, j+4\}$ and the second tier connections are $\{j-3, j+2, j+3, j+5, j+6\}$. Therefore, with probability of at least $(1-z)z^9$ node j survives and all his first and second tier connections have an accident and die reaching the absorbing state.

In summary, we have shown that if there is one agent left with different type to the rest, with positive probability we can reach the absorbing state. If there are two or more agents whose type is different than the rest of the network, we can apply an analogous reasoning to reach the absorbing state in some finite number of steps. Since we can reach an absorbing state from any state with positive probability, the result follows from Lemma 1.

Proof of result 8 (Disease dies out) Observe that the state with zero infected people is an absorbing state. At any given time t, for any number of sick people $m \in \{1, ..., n\}$, with probability $(1-\pi)^m > 0$ the disease is not spread and it dies out, reaching the absorbing state. Since we can reach the absorbing state from any other state with positive probability, and the number of states is finite, by Lemma 1 the probability that the process will be absorbed is 1.

B Data Appendix

Summary statistics for each of the variables we use are described in table 7.

Table 7: **Summary statistics**

	Summary Statistics				
Variable:	Obs	Mean	Std Dev	Min	Max
Tech Diffusion	72	-0.01	.63	-2.39	.999
log(GDP per worker)	65	9.29	0.89	7.02	10.48
Individualism	77	41.47	22.26	6	90
Degree	57	-0.62	2.54	-4.12	9.86
Stability	77	-0.04	.29	-.72	.63
Fractionalization	190	.44	.26	0	.93
Latitude	169	.29	.20	0	.71
Elevation	167	5.14	9.74	0	71.99
Kashima	85	.72	.45	0	1
English	86	.10	.28	0	1
Δgerm	77	22.52	3.76	15	31
Δgerm_std	77	1.12	.67	-.37	2.89

B.1 Disease Data

Historical disease data. To assess the historical prevalence of disease, we study 9 pathogens: leishmanias, leprosy, trypanosomes, malaria, schistosomes, filariae, dengue, typhus and tuberculosis. We choose these diseases because we have good worldwide data on their incidence, and they are serious, potentially life-threatening diseases that people would go to great length to avoid.

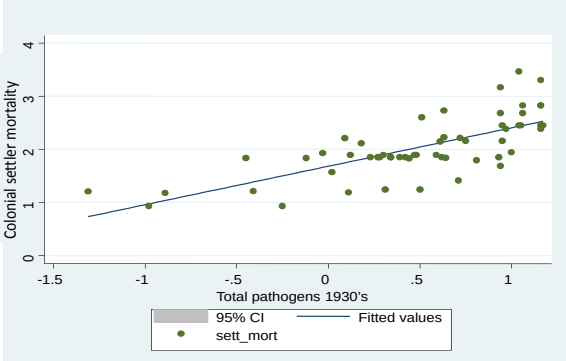
The historical pathogen prevalence data is from Murray and Schaller (2010), who built on existing data sets and employed old epidemiological atlases to rate the prevalence of nine infectious diseases in each of 230 geopolitical regions world. The nine diseases coded were leishmanias, schistosomes, trypanosomes, leprosy, malaria, typhus, filariae, dengue, and tuberculosis. For all except tuberculosis, the prevalence estimate was based primarily on epidemiological maps provided in Rodenwaldt and Jusatz (1961) and Simmons, Whayne, Anderson, and Horack (1945). Much of their data was, in turn, collected by the Medical Intelligence Division of the United States Army. A 4-point coding scheme

[13]Clearly the probability of this event is higher because of the infection process.

was employed: 0 = completely absent or never reported, 1 = rarely reported, 2= sporadically or moderately reported, 3 = present at severe levels or epidemic levels at least once. In the rare cases in which two epidemiological sources provided contradictory information, priority was placed on data provided by the older source. In cases in which the relevant maps were unavailable (this was especially true for leprosy) or insufficiently detailed (this was especially true for many of the Pacific island nations), prevalence ratings were informed also by verbal summaries found in Simmons, Whayne, Anderson, and Horack (1945). The prevalence of tuberculosis was based on a map contained in the National Geographic Society's (2005) *Atlas of the World*, which provides incidence information in each region for every 100,000 people. Prevalence of tuberculosis was coded according to a 3-point scheme: $1 = 3 - 49$, $2 = 50 - 99$, $3 = 100$ or more. For 160 political regions, they were able to estimate the prevalence of all nine diseases. The majority of these regions are nations (e.g.,Albania, Zimbabwe); others are territories or protectorates (e.g., Falkland Islands, New Caledonia) or culturally distinct regions within a nation (e.g., Hawaii, Hong Kong). Figure **??** uses a color-coded map to summarize the historical data.

One testament to the accuracy of this data is its high correlation with the historical disease data reported by Acemoglu, Johnson, and Robinson (2001). Figure 5 plots our total pathogen prevalence in the 1930's against the AJR data from the colonial period.

Figure 5: Relationship between colonial settler mortality and 1930's pathogen prevalence.



.

Contemporaneous disease data. Data were obtained from the Global Infectious Diseases and Epidemiology Online Network (GIDEON, http://www.gideononline.com) in 2011-12 and report primarily 2011 prevalence rates. The sources for data included in GIDEON currently include health ministry publications (electronic and print) and peer review journal publications. A partial listing is available at *http://www.gideononline.com/resources.htm*. The quality and frequency of data input vary by source. A total of 34 specific pathogenic diseases are coded, each on a 1-3 prevalence scale. There are some diseases that GIDEON classifies on a 6-point scale, according to the per-capita reported infection rate. The cutoff rates for each level vary by disease; for example, a "4" for rabies means an infection rate between .01 and .02 per 100,000 people, while the same range delimits a "3" for tetanus. We convert from the 1-6 scale to a 1-3 scale as follows: a 1 remains a 1, a 2 or a 3 is coded as a 2, and any number above 3 is coded as a 3. The total pathogen prevalence variable is the sum of the values for each disease within each country.

Our two pathogen prevalence indices appear to be accurate because they are highly correlated (0.77). They are also highly correlated with a similar index created by Gangestad & Buss (1993) to assess pathogen prevalence within a smaller sample of 29 regions. Correlations are 0.89 with our index from 1930's data and 0.83 with our index of 2011 data. This high correlation explains why the results with contemporaneous data are nearly identical. For example, the coefficient on the historical nine-pathogen index in table 8 is -2.73, while the analogous coefficient on the contemporaneous index is -2.72.

B.2 Measuring Individualism

Hofstede (2001) defines individualism in the following way:

> Individualism (IDV) on the one side versus its opposite, collectivism, that is the degree to which individuals are integrated into groups. On the individualist side we find societies in which the ties

Table 8: **Comparing first-stage regressions of individualism on historical and contemporary pathogen prevalence**

Dependent variable	Hofstede's individualism index (S)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Historical				Contemporary			
Total pathogens	−2.73				−2.72			
	(0.31)				(0.32)			
Human - zoonotic		-3.46	-2.15			-3.77	-2.38	
pathogens (Δgerm)		(0.44)	(0.45)			(0.57)	(0.48)	
Human - zoonotic				-5.26				-7.12
pathogens (Δgerm_std)				(2.04)				(2.90)
English			25.33	28.48			23.05	24.90
			(7.50)	(8.58)			(7.42)	(8.43)
Pronoun			-19.17	-28.33			-23.14	-30.02
			(4.83)	(4.70)			(4.35)	(4.57)
Constant	77.10	67.53	69.71	59.86	77.90	127.0	220.5	69.46
R^2	0.52	0.47	0.71	0.64	0.51	0.38	0.72	0.64
Observations	72	72	62	62	71	72	62	62

The table reports OLS estimates of the γ coefficients in $S = \gamma_1 + \gamma_2 A + \gamma_3 \bar{d} + \gamma_4 z + \eta$. The variables Δgerm and Δgerm_std are defined in equations (6) and (7). Columns (1)-(4) use historical disease prevalence data from the 1930s. Columns (5) -(8) use a more extensive set of diseases, measured in 2005. The other instruments are pronoun drop and whether is English spoken (see appendix B). Standard errors are in parentheses. All coefficients are significant at the 5% level.

> between individuals are loose: everyone is expected to look after him/herself and his/her immediate family. On the collectivist side, we find societies in which people from birth onwards are integrated into strong, cohesive in-groups, often extended families (with uncles, aunts and grandparents) which continue protecting them in exchange for unquestioning loyalty.

The Hofstede individualism index values are based on the results of a factor analysis of work goals across countries. The index was constructed from data collected during an employee attitude survey program conducted by a large multinational organization (IBM) within its subsidiaries in 72 countries. The survey took place in two waves, in 1969 and 1972 and included questions about demographics, satisfaction and work goals. The answers to the 14 questions about "work goals" form the basis for the construction of the individualism index. The individual answers were aggregated at the country level after matching respondents by occupation, age and gender. The countries mean scores for the 14 "work goals" were then analyzed using factor analysis that resulted in the identification of two factors of equal strength that together explained 46% of the variance. The individualism factor is mapped onto a scale from 1 to 100 to create the individualism index (hereafter IDV) for each country. The highest IDV values are for the United States (91), Australia (90), and Great Britain (89); the lowest are for Guatemala (6), Ecuador (8) and Panama (11). Subsequent studies involving commercial airline pilots and students (23 countries), civil service managers (14 counties) and consumers (15 countries) have validated Hofstede's results.

Figure **??** summarizes the findings of Hofstede's survey in a color-coded map. The most individualist countries (with an index between 80 to 91) are the Netherlands, Canada, Hungary, the United Kingdom, Australia and the United States. The most collectivist countries (with an index between 6 and 14) are Guatemala, Ecuador, Panama, Venezuela, Colombia, Pakistan, and Indonesia.

IBM survey text (a subset). The original Hofstede survey is too lengthy to include in its entirety. Below, we list a subset of the questions asked. We categorize questions according to which aspect of collectivism they measure, as described in section 3.2. That grouping is not in the original survey. The survey instructions read as follows:

We are asking you to indicate how important each of these is to you. Possible answers: of utmost importance to me (1), very important (2), of moderate importance (3), of little importance (4), of very little or no importance. How important is to you to:

Category 1: Questions about the importance of personal freedom and individual benefits from the organization

1. Have considerable freedom to adopt your own approach to the job (I)
2. Have a job which leaves you sufficient time for your personal or family life (I)
3. Have challenging work to do (I)

   In contrast, the last example question emphasizes the opposite, how the organization benefits from the individual's skills:

4. Fully use your skills and abilities on the job (C)

Category 2: Value of cooperation

1. Work with people who cooperate well with each other (C)
2. Have training opportunities (C)

Category 3: Willingness to change job or location

1. Live in an area desirable to you and your family (I)

We have followed the question with (I) when high importance (a low numerical score) indicates more individualism. When the higher importance indicates less individualism (more collectivist) we denote that with (C). We report these particular questions because all have factor loadings of 0.35 or more in absolute value.

Theories linking questions to network structure These questions reflect two views of a collectivist society: one where ties are strong, and one where ties are shared. In a widely cited paper, Granovetter (1973) provides the bridge between shared ties and strong ones; he argues, "the stronger the tie between A and B, the larger the proportion of individuals [that either of them knows] to whom they will both be tied." Granovetter goes on to give three theoretical reasons to believe this is true: (1) Time. If A and B have strong ties, they will spend a lot of time together. If A and C also have strong ties, they will also spend a lot of time together. If these events are independent or positively correlated, this necessarily implies B and C will spend a lot of time together, giving them a chance to form a strong tie. (2) The tendency of an individual to interact with others like himself. If A and B have strong ties, chances are good that they are similar; the same holds for A and C. Transitivity implies B and C will be similar, and will therefore get along. (3) The theory of cognitive balance. If A is good friends with B and C, then B will want to develop a good relationship with C, in order to maintain his relationship with A. Thus, Granovetter's theory explains why Hofstede's survey questions, many of which are about the strength of social ties, are informative about the prevalence of collectives, as defined in the model.

Other questions in Hofstede's survey assess the strength of cooperation, social influence and individuals' weight on social objectives. One example of such a question is " How important is it to you to work with people who cooperate well with each other?" Coleman (1988) explains why cooperative behavior is also linked to the presence of network collectives. He shows that effective norms depend on the presence of collectives because people enforce norms through collective punishments of deviators. If j observes i deviating from a social norm, then j can directly contact other friends of i to enact some joint retribution for the misdeed. When collective punishments are implementable, cooperation and conforming behavior is easier to sustain than if punishments must be implemented in an uncoordinated way.

A third category of questions in Hofstede's survey are about mobility, specifically one's willingness to move or change jobs. The essence of strong social ties is that the people involved are averse to breaking those ties. Thus an unwillingness to change one's social environment is indicative of strong social network ties. In the survey, the individualism index loads positively on one's willingness to move, which is consistent with the interpretations of individualism as a society with fewer collective and thus weak ties.

Cross-Country Network Analysis There is a small literature that analyzes and compares social network structures across countries. It is summarized and extended by Fischer and Shavit (1995). Surveys typically ask respondents to name people with whom they confided, were friends, asked for help, ect. The survey takers would then interview the named friends to find out their networks and interview the friends they named as well. By repeating this process many times, the researchers could map out fairly complete social networks in specific geographic locations. For our purposes, the key finding from these studies is that the frequency of network collectives varies greatly across countries. These studies do not typically report the number of collectives. They report a related measure, network density. Density is the fraction of possible links between individuals that are present. Importantly, a network that is fully dense also has the maximum possible number of collectives. Because this research design involves lengthy interviews of many respondents, it has been done only on a handful of countries. But it is useful to see how the prevalence of network collectives correlates with Hofstede's individualism index.

Table 9: **Measures of network interdependence and individualism**

Region	Country	Network interdependence	Individualism (for country)
Haifa	Israel	0.57	54
N. California	U.S.	0.44	91
all	U.S.	0.40	91
E.York, Toronto	Canada	0.33	80
London	U.K.	0.34	89
Taijin	China	0.58	20
West Africa		0.45-0.77	20

The theory predicts a negative relationship between network interdependence (closely related to collectivism) and the individualism index. Interdependence is measured as the fraction of all possible links in a social network that are present. It is also referred to as "network density." West Africa here includes Ghana, Nigeria and Sierra Leone.

Correlation of individualism with other measures of culture. To better understand what Hofstede's individualism index (IDV) measures, we examine related cultural measures that are highly correlated with the index.

Family structure. In a collectivistic society, people grow up with members of an extended family and sometimes also neighbors, housemates, other villagers, lords and servants. Collectivists have strong ties and frequent contact with family members. In individualistic societies, people grow up in nuclear families. Their family ties are weaker. Extended family live elsewhere and visit infrequently.

Group identity. In collectivist societies, people learn to think about themselves as part of collective, with a group identity. That identity is determined by birth. Similarly, friendships come from existing group ties. Members of the collective are distinct from non-members. In the individualistic society, people learn to think about themselves as an individual, not a member of a group. There is no distinction between group members and and non-members. Gudykunst, Gao, Schmidt, Nishida, Bond, Leung, and and (1992) surveyed 200 students in each of 4 countries: Australia and US (high IDV) and Hong Kong and Japan (lower IDV). Half of the respondents were asked to imagine a group member; the others were asked to imagine a non-member. They were then asked to report if they would: talk about themselves with the person, ask about the other, expect *shared attitudes and networks*, and have confidence in the other. The differences between how respondents viewed group members and non-members correlated exactly (negatively) with their country's IDV scores.

Other ways of modeling individualism and collectivism in networks. Weak vs. strong ties: Granovetter (1973) introduced the idea of strong ties and weak ties in networks. Strong ties are close friends, while weak ties are acquaintances. Granovetter argues that more novel information comes from weak ties than from strong ties. The reasoning is very similar to that in our model. Because people who are very closely socially related have similar information sets, they are more likely to convey redundant information and are less likely to have novel information. Weak ties are more likely to be connected to people that we do not know and therefore are possible

conduits for new information. Granovetter argues that people with few weak ties are at an informational disadvantage because they have difficulty accessing information in other parts of the social network. Thus, a society comprised of agents with mostly strong ties and few weak ties will not transmit information (or disease) as easily. Thus another way to formulate our model that would lead to the same conclusions would be to characterize collectivist societies as ones with strong ties and individualist societies as one with weak ties.

B.3 Measuring Link Stability

Using the hypothesis that people break social network ties when they move from one community to another, we construct the following proxy for social link stability. The data on link stability comes from the General Social Survey (GSS) the web address is *http://www3.norc.org/gss+website/*. The data consists of 4 variables that measure the length of time a respondent has spent in his current community. These variables and their descriptions are:

1. **loclived**: How long has respondent lived in community? *Asked in 1987.*
2. **livecom**: *Literal Question 1215, asked in 1986.* How long have you lived in the city, town or community where you live now?
3. **comyear**: *Literal Question 1248, asked in 2002.* How long have you lived in the city, town or local community where you live now?
4. **livecom1** : *Literal Question 1449, asked in 1996.* How long have you lived in the community where you live now?

Ideally, we would observe at least one of these variables for many GSS respondents. However, each of these variables was observed in only one year; this combined with the relatively small sample size of the GSS means that it is necessary to combine these variables into one "*location*" variable. To make the levels of these variables comparable, we divide each variable by its mean.

Next, select out all the first-generation immigrants. We separate these into the ethnicity, using the GSS variable *ETHNIC*. Ethnicities are supposed to be listed in order of importance. Thus, in cases where multiple ethnicities are reported, we use only the first one. Sometimes respondents report regions, rather than countries as an ethnicity. We map regions into countries as described in the following section.

Finally, we want to control for obvious demographic differences that likely affect mobility. For example, poor and young people are more likely to have to move for job-related reasons. To control for such confounding effects, we regress **location** on: (x_1) age, (x_2) income, (x_3) marital status, ($x_4 - x_9$) dummies for education level, and (5) the proportion of the population constituted by the respondent's nationality. The education levels are less than high school, high school, associate, bachelor, graduate, and other. The GSS names for the first four of these variables are, respectively age, coninc, marital, and degree. The proportion variable was constructed using country of origin data from the GSS variable natborn. Including each person i, born in country j, we estimate:

$$location(i,j) = \alpha + \sum_{k=1}^{9} \beta_k x_k(i,j) + e(i,j)$$

Then, we construct the *stability* variable for each country as an average of the residuals e:

$$stability(j) = \frac{1}{n_j} \sum_{i=1}^{n_j} e(i,j)$$

where n_j is the number of respondents who were born in country j.

B.4 Measuring Network Degree

The data on network degree comes from the General Social Survey (GSS) the web address is *http://www3.norc.org/gss+website/*. The variable is called *frinum* and it asks the respondent " Thinking now of close friends - not your husband or wife or partner or family members - but people you feel fairly close to, how many close friends would you say you have?" The modal response is 4. Roughly 60% of respondents report between 2 and 6 close friends. All respondents reside in the United States.

We first select out all the first- and second-generation immigrants. We separate these into the ethnicity, using the GSS variable *ETHNIC*. Ethnicities are supposed to be listed in order of importance. Thus, in cases where multiple

Table 10: Average Path Length, Network 5

Factions	No breaks	Breaks
$F = 2$	17.1324	17.0427
$F = 4$	17.5235	18.3141
$F = 8$	21.3676	20.9381
$F = 20$	23.9995	24.0956
$F = 40$	24.7444	25.0653

100 runs. $p = 1/70$. 200 nodes.

ethnicities are reported, we use only the first one. We are using the number of friends of US residents who are first- or second-generation immigrants from a country as a proxy for the number of friends of residents of that country.

One problem is that many times, respondents do not list a country of origin, but instead list a region. We use a one-to-many mapping to impute a network degree measure for the countries in this region. Specifically, for Argentina, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Panama, Peru, Uruguay, and Venezuela, we assign al these countries the average number of reported friends of respondents whose origin is listed as "other Spanish." Bangladesh, Indonesia, South Korea, Malaysia, Pakistan, Singapore, Thailand, and Vietnam are assigned the value from the region reported as "Other Asian." Estonia, Luxembourg, Malta, and Switzerland are assigned the value from the region reported as "Other European." Iraq, Kuwait, Lebanon, Libya, Morocco, Saudi Arabia, UAE, and Egypt are assigned the value from the region reported as "Arabic." Ethiopia, Ghana, Kenya, Nigeria, Sierra Leone, South Africa, Tanzania and Zambia are assigned the value from the region reported as "Africa." Czech Republic and Slovak Republic are both assigned the value from "Czechoslovakia." Jamaica and Trinidad are assigned the value from "Non-Spanish West Indies." And finally, Hong Kong is given the same value as China. In addition, there were a few countries for which we still missed data and therefore assigned them the degree values of the closest substitute country or region. These substitutions are: Brazil was assigned the same value as "Other Spanish." Australia, Israel, and New Zealand were assigned the same value as "Other European." Iran, Turkey were assigned the same value as "Arabic." Finally, Suriname was assigned the same value as "Non-Spanish West Indies."

To control for obvious demographic differences that likely affect social networks, we regress **frinum** on: (x_1) age, (x_2) income, (x_3) marital status, ($x_4 - x_9$) dummies for education level, and (5) the proportion of the population constituted by the respondent's nationality, as described in the previous section. Including each person i, from country j, we estimate:

$$frinum(i,j) = \alpha + \sum_{k=1}^{9} \beta_k x_k(i,j) + e(i,j)$$

Then, we construct the *degree* variable for each country as an average of the residuals e:

$$degree(j) = \frac{1}{n_j} \sum_{i=1}^{n_j} e(i,j)$$

where n_j is the number of respondents who are from country j.

B.5 Fractionalization and path length

Deriving properties of random networks analytically is often infeasible. So far, we do not have clean analytical proof of the relationship between the number of factions and the average path length of a social network. But numerical results back up the argument in the text that a large number of factions increases path length, and therefore slows diffusion. Using the other parameter values from Table 1 and just varying factions yields the results in Table 10.

B.6 Other Control Variables and complete results for output per worker

An inevitable question arises: "What if you also control for X?" We would like to know if individualism is highly correlated with and thus proxying for some other economic phenomenon. The problem with answering this question is that what we would like to control for is likely an endogenous variable. We could treat it as such and instrument

for it. But in most cases, our instruments are not strong predictors. Or, we could just, suspend disbelief, assume that these are exogenous variables, abandon any pretense of saying anything about causality, and just see what statistical relationship they have with the other variables in the estimation. We take the second approach. Each row of table 12 reports the coefficients of a second stage regression of technology diffusion on the Hofstede individualism index, one other control variable, and a constant. Since we have assumed that the control variable is exogenous, we use it as an instrument in the first stage, in addition to a constant and our standard instruments: pronoun, english and the standardized difference in pathogens variable, Δgerm_std.

Table 11: **Controlling for other economic variables**

Dependent variable	Technology Diffusion						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Individualism (S)	0.59	0.69	1.23*	1.35*	1.02*	1.24*	1.46*
	(0.34)	(0.39)	(0.30)	(0.36)	(0.27)	(0.36)	(0.31)
Life expectancy	4.29*						
at birth (LEB)	(0.78)						
Social		112.2*					
Infrastructure (SocI)		(30.17)					
Ethno-linguistic			−1.08*				
fractionalization (EFL)			(0.21)				
Latitude				0.21			
				(0.26)			
Disease-adj					−0.0030*		
life expectancy (DALY)					(0.0006)		
Capitalist						5.89	
(EcOrg)						(4.44)	
Population							0.040*
Density							(0.010)
Constant	-300.7	-98.51	-15.40	-67.68	7.05	-76.77	-72.64
R^2	0.58	0.47	0.52	0.33	0.63	0.34	0.43
Observations	62	60	55	61	61	61	62

2SLS estimates of $100 * \gamma$ coefficients in Diffusion $= \gamma_1 + \gamma_2 \mathrm{S} + \gamma_3 x + \eta$, where the x variables are listed in the first column of the table. The first stage regression is $S = h_1 + h_2 x + h_3 \Delta\text{germ\_std1930} + h_4 pronoun + h_5 english + e$. Standard errors in parentheses. $*$ denotes significance at 5% level.

The control variables are social infrastructure, a measure of the efficient functioning of political and social institutions, constructed by Hall and Jones (1999); ethno-linguistic fractionalization, a measure of the probability that two randomly-chosen people in the country will belong to different ethnic or linguistic groups, constructed by Taylor and Hudson (1972); latitude, which is the absolute value of the country's latitude, divided by 90; disability-adjusted life expectancy, which is the expected length of time an individual lives free of disability, is measured by the World Health Organization in 2004 (*http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/index.html*); capitalist, which is the "economic organization" variable constructed by Freedom House, scores more capitalist countries higher and more socialist countries lower; and population density is the 1970 population per square mile, as reported by the World Bank.

Table 12: **Controlling for other economic variables**

	Labor Productivity							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instrument:	Δgerm	Δgerm_std	Δgerm	Δgerm_std	Δgerm	Δgerm_std	Δgerm	Δgerm_std
Individualism	2.07	1.53						
	(0.45)	(0.52)						
Degree			43.19	41.14				
			(13.23)	(12.79)				
Stability					-4.64	-1.97		
					(2.14)	(1.48)		
Fractionalization							-2.89	-3.52
							(0.92)	(0.68)
Over-ID p-val	0.65	0.13	0.20	0.03	0.94	0.08	0.08	0.46
	Accept	Accept	Accept	Reject	Accept	Reject?	Reject?	Accept
N	63	63	48	48	63	63	63	64

2SLS estimates of $100 * \gamma$ coefficients in $Y/L = \gamma_1 + \gamma_2 \mathrm{S} + \eta$, where Y is real GDP and L is population. The first stage regression is $S = h_1 + h_2 \Delta\text{germ} + h_3 pronoun + h_4 english + e$. Standard errors in parentheses.

Table 13: **Classification of Diseases**

Disease type is H if we classify the disease as human, M for multi-type and Z for zoonotic. The classification is based on the disease reservoir. Historical is Y if the disease is present in our 1930's historical data. For disease categories with multiple sub-types (i.e. Filaria - Bancroftian and Filaria - Brugia Timori), there is one combined value in the historical data. That value is the maximum of the value for each strain. It lists 4 if any strain is epidemic, 3 if any strain is endemic, ect. All listed diseases are present in the 2011 data. Different strains are treated like different diseases. All 2011 results are robust to combining strains (results available on request). Source: GIDEON database.

Disease	Agent	Reservoir	Spread By	Type	Historical
Diphtheria	Bacteria	Man	Droplet, Contact, Dairy, Clothing	H	N
Filaria - Bancroftian	Nematoda	Man	Mosquito	H	Y
Filaria - Brugia Timori	Nematoda	Man	Mosquito	H	Y
Measles	Virus - RNA	Man	Droplet	H	N
Meningitis - Bacterial	Bacteria	Man	Air, Secretions	H	N
Meningitis - Viral	Virus - RNA	Man	Fecal-oral, Droplets	H	N
Pertussis	Bacteria	Man	Air, Secretions	H	N
Poliomyelitis	Virus - RNA	Man	Fecal-oral, Food, Water, Flies	H	N
Smallpox	Virus - DNA	Man	Contact, Secretions, Fomite	H	N
Syphilis	Bacteria	Man	Sexual Contact, Secretions	H	N
Typhoid fever	Bacteria	Man	Fecal-oral, Food, Flies, Water	H	N
Dengue	Virus - RNA	Man, Monkey, Mosquito	Mosquito, Blood (rare)	M	Y
Filaria - Brugia Malayi	Nemotoda	Man, Primate, Cat, Civet	Mosquito	M	Y
Leishmania - Cutaneous	Protozoa	Man, Rodent, Other Mammals	Fly	M	Y
Leishmania - Mucocutaneous	Protozoa	Man, Rodent, Sloth, Marsupial	Fly	M	Y
Leishmania - Visceral	Protozoa	Man, Rodent, Dog, Fox	Fly, Blood	M	Y
Leprosy	Bacteria	Man, Armadillo	Patient Secretions	M	Y
Malaria	Protozoa	Man, Mosquito	Mosquito, Blood	M	Y
Trypanosoma - African	Protozoa	Man, Deer, Cattle, Carnivores	Fly	M	Y
Trypanosoma - American	Protozoa	Man, Dog, Cat, Other Mammals	Kissing Bug, Blood, Fruit	M	Y
Tuberculosis	Bacteria	Man, Cattle	Air, Dairy Products	M	Y
Typhus - Epidemic	Bacteria	Man, Flying Squirrel	Louse	M	Y
Anthrax	Bacteria	Soil, Water, Other Mammals	Fly, Hair, Hides, Bone, Air, Meat	Z	N
Leptospirosis	Bacteria	Frog, Cattle, Other Mammals	Water, Soil, Urine, Contact	Z	N
Rabies	Virus - RNA	Dog, Fox, Other Mammals	Saliva, Bite, Transplants, Air	Z	N
Schistosomiasis - Haematobium	Flatworms	Snail, Baboon, Monkey	Water (Skin Contact)	Z	Y
Schistosomiasis - Intercalatum	Flatworms	Snail	Water (Skin Contact)	Z	Y
Schistosomiasis - Japonicum	Flatworms	Snail, Other Mammals	Water (Skin Contact)	Z	Y
Schistosomiasis - Mansoni	Flatworms	Snail, Other Mammals	Water (Skin Contact)	Z	Y
Schistosomiasis - Mattheei	Flatworms	Snail, Other Mammals	Water (Skin Contact)	Z	Y
Schistosomiasis - Mekongi	Flatworms	Snail, Dog	Water	Z	Y
Tetanus	Bacteria	Animal Feces, Soil	Injury	Z	N
Typhus - Endemic	Bacteria	Rat	Flea	Z	Y
Typhus - Scrub	Bacteria	Rodent, Carnivores, Mite	Mite	Z	Y