WORKING PAPER SERIES

Center for the Study of *Social Organization*

Intergroup Inequality as a Product of Diffusion of Practices
with Network Externalities under Conditions of Social Homophily:
Applications to the Digital Divide in the U.S. and
Rural/Urban Migration in Thailand

Paul DiMaggio & Filiz Garip
Working Paper #2, Spring 2010

CSSO

Intergroup Inequality as a Product of Diffusion of Practices with Network Externalities under Conditions of Social Homophily: Applications to the Digital Divide in the U.S. and Rural/Urban Migration in Thailand

Paul DiMaggio
Princeton University

Filiz Garip
Harvard University

August, 2009

Draft 2.0 PLEASE DO NOT QUOTE, REPRODUCE, OR DISTRIBUTE WITHOUT PERMISSION OF AUTHORS

We are grateful to Duncan Watts for the opportunity to present a very early version of our model at his networks workshop at Columbia University. We have benefited as well from feedback on this version from participants in the Harvard MIT Economic Sociology Workshop, the Princeton University Economic Sociology Workshop, and the Labor Lunch/Social Networks Workshop at the University of Pennsylvania/Wharton School as well as from helpful advice from Miguel Centeno, Peter Marsden, Gabriel Rossman, and Bruce Western.

Abstract

Research in social stratification has tended to view intergroup inequality in one of two ways. Work in the status-attainment tradition focuses on individual outcomes and, by implication, views the reproduction of intergroup inequality as a consequence of agents with differing endowments attaining outcomes that vary depending on the level of those endowments. More recent work has deviated from this aggregationist strategy in two ways. First, researchers have introduced social structure, in the sense of ego-centered social networks, as an additional kind of resource upon which actors draw in their efforts to retain privilege or achieve social mobility. Second, other researchers have studied how collective action may alter the terms of competition by changing state policies or the practices of private actors in response to claims by mobilized groups. In this paper, we introduce a third mechanism, which we contend chronically reproduces and, under some conditions, may generate or even efface intergroup inequality. That mechanism is (a) the diffusion of goods, services, and practices that (b) are characterized by strong network externalities under conditions of (c) social homophily. When the value of a good or practice to an agent is a function of the number of persons in that agent's network who also possess the good or engage in the practice, and when networks are homophilic with respect to certain social characteristics, this mechanism will exacerbate initial individual-level differences in access to the good or practice and, under some conditions, induce persistent intergroup inequality. We illustrate this claim in two empirical contexts. For the first, the diffusion of access to and use of the Internet, we start with observed data on the relationship between cost and adoption and between adoption levels and price, and produce a computational model that permits us to predict variation in intergroup inequality over time as a function of variation in the strength of network externalities and the extent of social homophily. For the second, the practice of rural-to-urban migration by young people in rural Thailand, we use village-level data on family resources and migration patterns to explore the relationship between information sharing, homophily, and over-time change in differences among villages in migration rates, demonstrating that, as in the Internet case, homophily interacts with network externalities to exacerbate inequality. We conclude with a discussion of the scope conditions of our argument and the range of phenomena to which this mechanism may apply.

Introduction

We draw on work in several fields to introduce a class of social mechanisms that we believe influence intergroup inequality, usually by making such inequality larger and more durable than it would be in a world without these mechanisms, although potentially effacing inequality as well. These mechanisms come into play when (a) a good, service, or practice influences individual life chances; (b) that good, service or practice is characterized by network externalities, such that the costs to an actor are lower or the benefits higher if persons with whom he or she is socially tied consume the good or service or engage in the practice; and (c) actors' social networks are characterized by social homophily. We illustrate our argument with examples of two such practices --- technology adoption in the U.S. and internal migration in Thailand – to suggest the scope within which we believe these mechanisms operate.

This paper's approach is broadly consistent with the view that sociological explanation can be advanced by the identification of "social mechanisms" that entail (a) goal-directed individual actions and (b) consequent social interactions that (c) yield higher-level outcomes that (i) are emergent (i.e., that cannot be recovered simply by aggregating the individual actions that combine to produce them) and that (ii) vary depending upon the initial social structure (ordinarily depicted in terms of social networks).

Although this approach has deep roots in sociology and has become especially prominent in recent years (Hedstrom 2005; Tilly 2006; Watts 2003), it has played a less central role in empirical research on social inequality than one might have expected. Instead, research on social inequality has often described inequality as the aggregate product of individual efforts to obtain useful educations, good jobs, and adequate incomes. For many years, work in the status attainment tradition in sociology (as well as the human-capital tradition in labor economics) focused on individual attainment of education, jobs, and earnings, treating these as functions of initial endowments and life-course events (Featherman and Hauser 1978). The implication of this approach is that intergroup inequality reflects the fact that people with similar initial endowments have similar experiences that lead them to similar outcomes.

In response to the evident limitations of such imagery, students of social stratification began introducing structure into research on social stratification in the 1980s. But they often did so by converting social structure into endowments predicting individual-level success, thus reproducing the individualistic and aggregationist bias of the status-attainment tradition, albeit with sophisticated accounts of the impact of structural position on achievement. Thus measures of network position (in-degree or "popularity," centrality, occupancy of "structural holes," and so on) are treated as forms of "social capital," analogous to human capital, that advance individual life chances (Lin 1999; Burt 2000). Similarly, characteristics of the firms and other organizations in which individuals are employed --- size, profitability, market power, prestige – are often converted into individual-level variables in income-determination models (Bielby and Baron 1980).

There are, of course, many exceptions to the aggregationist tendency in research on social inequality. Classical sociology emphasized the impact of collective action (by elites or subordinate groups) on the degree of intergroup inequality. Marx and Engels (1888 [1966]) famously (if incorrectly) argued that concentrating laborers in factories generates processes of identity formation and social learning that culminate in social revolution. In this tradition, Offe and Wiesenthal (1980) argues that the number and social organization of business elites and union members, respectively, influence their strategies of collective action. Weber (1922 [1968]) described how dominant groups use their influence to establish criteria of virtue ("status honor") that benefit their members. More recently, Karabel (2005) demonstrated how U.S. Protestant elites manipulated criteria of virtue to monopolize admissions to prestigious private universities and documented the effectiveness of collective action by African-Americans in persuading these institutions to adopt affirmative action plans that enhanced opportunity for underrepresented groups. Tilly (1999) likewise emphasizes the role of collective action ("opportunity hoarding") in reproducing inequality.

Institutional research on inequality also departs from the aggregationist bias of the status-attainment tradition by focusing upon supra-individual processes that influence the distribution of opportunity. Political institutionalism has emphasized the ways in which state policies influence inequality by providing structures of opportunity, redistributing income, and establishing the terms of political competition (Western 2006; Gustaffson and Johansson 1999). The application of neoinstitutional theory to the study of inequality has demonstrated the varying efficacy of state mandates and organizational policies in reducing categorical inequality in the workplace (Kalev, Dobbin and Kelly 2006). Such research must defocalize individual choice, often because it examines cases in which higher-level mechanisms such as incarceration, income transfers, or discrimination operate in ways that constrain the scope of individual action.

Only a few studies of inequality focus on the production of higher-level patterns as a consequence of individual choice under varying structural conditions. Simon (1957) introduced the idea that inequality was a function of the depth of organizational hierarchies (itself a function of organizational size and span of control) and norms about appropriate compensation differences for employees at different hierarchical levels. White (1970) explored the impact of rates at which positions in hierarchies become vacant on several emergent properties of systems of inequality. Boudon (1973) developed a competitive model of individual decisions about investments in schooling to explain the persistence of intergroup inequality despite unprecedented levels of educational expansion (and see Breen and Goldthorpe 1997). Rosenbaum (1976) demonstrated empirically the production of educational inequality as a function of the structure of opportunities within a large urban high school. Family demographers have explored the impact of family structure and marital homophily on income inequality (Western, Bloome, and Percheski 2008).

More recently, research on neighborhoods has focused on mechanisms generating individual-level behavioral or health outcomes likely to influence socioeconomic outcomes. Such research, ably reviewed by Sampson, Morenoff and Gannon-Rowley (2002), examines neighborhood effects on social mechanisms (especially the nature of neighborhood interaction networks or place-specific activity patterns) and emergent consequences at the neighborhood level. Although not all of this research explicitly addresses intergroup inequality *per se,* much of it addresses multiple interactions among conditions associated with concentrated poverty. At the same time, the authors note that existing studies rely on cross-sectional data and, in many cases, indirect measures of neighborhood level variables, often failing to realize potential gains from multi-level designs and theoretical attention to social mechanisms.

One purpose of this paper is to contribute to the depiction and understanding of the role of mechanisms entailing the interaction of structure-dependent individual choices in generating, reproducing, and effacing intergroup inequality. This project emerged from the authors' respective efforts to solve concrete empirical problems --- whether diffusion of Internet use would eventually eliminate intergroup differences observed in the technology's first years; and why demographically and geographically similar Thai villages diverged over time in rates or urban-rural migration. These two empirical questions, we believe, entail social mechanisms that are analytically quite similar. In each case individual choices (to use the Internet or to migrate) are a function of previous choices by members of ego's social network; and intergroup inequality (in technology use or in the benefits of migration) is likely to be amplified by the extent that ego's social network is homophilous with respect to socioeconomic status.[1]

To develop this model, we combine elements from three well-developed social-science literatures: on network externalities; on innovation diffusion under conditions of interdependent choice; and on homophily in social networks. Before presenting our cases, we briefly review work in each field.

Network externalities, diffusion models, and social homophily

[1] To our knowledge, this hypothesis – that inequality is exacerbated when adoption of goods and practices that can promote status attainment is governed by network effects under conditions of homophily – has neither been posed nor tested previously. The only shard of relevant evidence we could find is Van den Bulte and Stremersch's (2004) finding that adoption processes in countries with high levels of income inequality reveal stronger choice interdependence ("contagion effects") than processes in societies with lower levels of inequality. This result admits to many alternative interpretations (including the author's own, that it reflects heterogeneous and normally distributed reservation prices). The authors contend that their result is inconsistent with a contagion account of innovation adoption, but this contention is based on an assumption, which we find less than fully plausible, that local homophily is negatively correlated with global inequality; and on an untested assumption that inequality is the cause rather than the outcome of variation in adoption processes. In any case, the fact that the contagion parameter is based on overall adoption and not on adoption by network peers makes the finding a weak, albeit supportive, test of our argument.

In 1892, John F. Parkinson, owner of a hardware and lumber business, became the first telephone subscriber in Palo Alto, California. (This account is taken from Fischer 1992: 130-34.) Parkinson, an entrepreneur who contributed much to the economic development of the town he would eventually serve as mayor, placed the phone in his business. A line to the Menlo Park telephone exchange a few miles away connected him to other businesses in the Bay Area. By 1893, he was joined by a realtor and a butcher. Shortly thereafter, the local pharmacy took a subscription, placing their phone in a quiet room and permitting residents to use it on a pay basis. By 1897, after monthly charges fell, Palo Alto had nineteen telephone subscribers, including several – Parker, two physicians and two newspaper editors – with telephones in their homes.

It is no accident that early subscribers were businessmen and professionals for whom the telephone was a means of staying in contact with suppliers, customers, and clients – nor that many citizens used telephones (when they used them) at locations outside their home or place of business (just as many early Internet users inhabited Internet cafes). It did not make much sense to get a phone for social reasons unless you could call your family and friends. And it was not at all obvious how social (as opposed to business) telephone use would reach critical mass for takeoff. Indeed, it was years before even the telephone companies recognized the potential of the telephone as an instrument of sociability (Fischer 1992: chapter 3). Not until 1920 percent did telephone subscription rates approach 50 percent even in prosperous Palo Alto (with rates considerably lower in neighboring communities) (*ibid.*: 141). Naturally, telephones were more common in professional and business households, whose members were more likely to have friends and relatives who also resided in telephone households. By 1930, blue-collar households caught up in Palo Alto (where many blue-collar workers were independent tradesmen whose clients had phone service) but not in neighboring towns where blue-collar residents were more likely to work in factories (*ibid.* 146-47).

Even after the takeoff of the telephone, inequality in access was tenacious. The United States did not approach universal service (*i.e.*, approximately 90 percent household penetration) until 1970. And even in 1990, the least advantaged Americans often went without, with service close to universal at incomes of $20,000 or more (in 1990 dollars) and declining precipitously below that. Thus 50 percent of mothers living below the poverty line had no telephone service, as did 35 percent of all families receiving public assistance.[2] Race and ethnicity had independent effects on telephone service even controlling for income: overall, 16 percent of African-American households and 18 percent of Hispanic households did not have telephones (these figures are all from Schement 1995). The independent effect of race, and the nonlinear effect of poverty, suggest (but do not prove) that something other than income – perhaps interaction effects related to network composition -- may be driving these differences.

Now consider a very different case: inequalities by social class and race in advanced placement (AP) course enrollments in contemporary high schools (Kao and Thompson 2003; Klopfenstein 2004; Brown 2005). Imagine a high school class with three hundred students, divided into three hierarchically arrayed tiers. Each of these students must decide whether or not to take an advanced placement class. Each knows that advanced placement courses will help her or him gain admission to selective colleges

[2] The spread of cellular telephony rendered data on household use less informative, because by the nid-1990s, some Americans, including young, affluent Americans, were beginning to substitute cell phone service for land lines. Thus by 2000, household penetration actually declined for the first time since the Depression, but the decline appears to have been the result of cell-phone substitution, not a reduction in telephone access.

and may even reduce the time it takes to earn a college degree. Each also knows that to take advantage of this benefit, he or she will have to learn enough to pass the AP exam, and that this will require a substantial investment of time.

Several factors will influence the students' decisions: whether they expect to go to college or can afford to attend a selective college, whether they have the preparation to be admitted into AP courses or to have a realistic chance of passing them, whether they have competing demands (for example team sports or after-school work) that reduce the hours available for study, and so on. Based on inequality in financial and other endowments, imagine that the probability of choosing to enroll in an AP course is 80 percent for those in the top tier, 50 percent for those in the middle tier, and 20 percent for those in the bottom tier.

Now, consider the network externalities that are associated with taking an AP course. If your friends also take the course, it may make more sense for you to take it as well: You can study together (which is fun [increasing the benefits] and efficient [reducing the costs]); and perhaps collaborate on homework; and help each other review; and share the cost of a scientific calculator, or use your friend's high-speed Internet connection to prepare presentations. Such considerations will increase the benefit to you of taking the course (you will learn more), reduce the rate at which you discount expected benefits (you will be more likely to pass the AP exam), and also reduce the cost of taking the course in both time and money.

If friendships are distributed at random – if a top-tier kid is as likely to have a friend in the bottom tier as one in his own group – then, on average half of each student's friends will plan to take an AP course. If kids in each tier have the same number of friends, and if the effect of friends taking AP courses on your own decision is additive, then a lot more students will take AP courses than if their decision were entirely isolated. Moreover, because lower-tier kids will know as many AP-course-takers as upper-tier kids, there is a good chance that the outcome will be a little more equal than if the choices did not interact.

The problem with this scenario is that high-school friendships are never randomly distributed. Real-world nets are homophilic: kids in each tier hang out with one another more than with kids from other groups (Kandel 1978; Shrum, Cheek and Hunter 1988; Quillian and Campbell 2003). If we build such homophily into our example, externalities will exacerbate inequality, not reduce it. Most of the positive effect of mutual influence will accrue to students in the top tier, because their friends disproportionately enroll in AP courses. By contrast, kids in the bottom tier will have only a few friends in AP courses, so externalities will affect their choices less. Put these influences together – big effects on the most advantaged students' choices, small ones on the choices of the least advantaged – and the result amplifies initial inequalities.

The point of these examples is to convey the intuitions behind, and to suggest the scope of applicability of, the models we describe more systematically in the remainder of the paper. Each example has three elements: a *choice* (purchasing telephone service, signing up for an AP test); *network externalities* (your telephone is more valuable if you can all your friends with it; you can get more out of the AP course with less effort if your friends take it, too); and *social homophily* (which makes the benefits of externalities redound more decisively to those groups that possess an initial advantage). Put these three elements – choice, externalities, and homophily – together in a diffusion process (which is to say a lot of people observing one another making choices over time) and the result is a pattern of interactions that often exacerbates social inequality. In the remainder of this section, we look more closely at these elements (externalities, homophily, and systems of choice) and describe some literature that has informed or anticipated our work.

Network Externalities

A product, service, or behavior possesses network externalities in so far as its value to an actor is conditional upon the number of other actors who consume the product or service or engage in the activity. The term derives from work in communications economics and the economics of innovation, which has focused on the aggregate value of a network as a nonlinear function of scale (and the difficulty of internalizing that value) (Katz and Shapiro 1985; Arthur 1989; Shy 2001).[3] By contrast, we focus on network effects from the standpoint of individual decision-makers (to whom the value of action increases with network size). DiMaggio and Cohen (2004) distinguish between *general* network externalities (when each user's perceived benefit from adoption of a good or practice is influenced only by the number and not by the identities of other adopters) and *specific* network externalities (where each user's benefit is a function of the identities of those who adopt). Specific, or identity-based, network externalities can be *status-based* (*e.g.*, positive when users are higher-status than oneself and negative when they are lower-status) or *network-based* (when changes in perceived benefit are a function of the number (or percentage) of members of one's network who have already adopted a good or practice). Whereas the economic literature has emphasized general network

[3] The idea, if not the term, goes back considerably further, at least to Leibenstein (1950: 189), who described a class of goods for which "the utility derived from the commodity is enhanced or decreased owing to the fact that others are purchasing and consuming the same commodity,"

externalities, we are exclusively concerned with *specific* externalities, because these have especially significant implications for social inequality. In order to make the notion of network externalities useful for the study of social inequality, then, we *narrow* it by focusing upon identity-specific externalities but *broaden it* by applying to practices and behaviors as well as to goods and services.

As our example of telephone use indicates, network are endemic to communicat-ions technologies and utilities: telephones, fax machines, e-mail clients, instant messag-ing, and so on. Network effects also characterize any kind of software that produces files that consumers want to exchange with one another (text files, spreadsheets, graphic files, and so on). “Social networking” sites like Facebook or Twitter similarly increase in value in proportion to the number of one’s associates who are already on them. But as our AP-course example suggests, we hypothesize that network externalities characterize a much broader range of choices: choice between public, secular private, and religious se-condary schools (kids like to go to the same schools as their friends and parents benefit from car-pooling and information-sharing with other parents); fertility (raising children is a lot easier if your friends have kids [or if you can make friends with people who do]) (Buhler and Fratczak 2007); divorce (the more divorces in one’s social network, the more people will be available for support and companionship and the larger the pool of potent-ial mates) (Booth, Edwards and Johnson 1991); and migration (the more migrants one knows, the easier it is to get vital information about the place to which one is migrating and the more available will be social support once one gets there) (Curran, Garip, Chung and Tangchonlatip 2005; Uhlig 2006); religious choice (Shy 2007); and bilingualism (Church and King 1983).

Other things equal, network externalities tend to produce dual equilibria in diffusion: very low levels of adoption until some takeoff threshold, after which adoption increases rapidly (Shy 1971). But in many cases the takeoff threshold is easily reached because a critical mass of initial adopters is motivated by intrinsic returns. Under these conditions, network externalities tend to increase adoption levels across the board. But where externalities are identity-specific and ego-network-based, and where ego networks are homophilous, externalities can become a potent source of increased inequality, amplifying the initial effects of varying endowments.

Social Homophily

Social networks are homophilous with respect to a social trait to the extent that connected pairs of actors in a network share the characteristic in question.[4] In other words, social homophily exists to the extent that people are similar to their friends and associates. Since Lazarsfeld and Merton (1954) coined the term, research has found homophily to be ubiquitous. Studies in numerous societies and organizations demonstrate pervasive homophily with respect to race and ethnicity, gender, age, educational attainment, occupation, religion, and other factors (McPherson, Smith-Lovin, and Cook 2001).

Many researchers have studied the impact of homophily on diffusion processes and information flows. Rogers (2003:307) depicts homophily as a barrier to the diffusion of innovation, arguing that where homophily is strong, adoption of innovations is more

[4] McPherson, Smith-Lovin and Cook (2001) note that observed homophily reflects both "baseline homophily," which produces homophily even in random networks based on the distribution of characteristics of the nodes, and "inbreeding homophily," which reflects patterns of similarity between tied agents above what one would expect if ties were distributed at random. These authors also usefully distinguish between "inbreeding homophily" and "choice homophily," noting that greater-than-random levels of homophily may result not only from choice but also from institutional factors (*e.g.,* occupational gender segregation) and from correlations among different personal characteristics. These are critical distinctions if one is concerned with the causes of homophily, but they are less critical given our focus on homophily's consequences.

likely to be restricted to elites. We suggest that, when returns to adoption are high, this tendency will reinforce inequality. To our knowledge, diffusion researchers have not considered the implications of homophily in diffusion processes for social inequality.

A few scholars *have* addressed the relationship between homophily and inequality (especially the impact on educational or occupational attainment), but have not placed this relationship in the context of an explicit model of diffusion or behavioral change. Thus Buhai and van der Leij (2008) model occupational segregation as a result of social homophily combined with network effects on access to jobs. Quillian (2006) argues that friendship homophily reduces the academic achievement of high-achieving Black and Hispanic students relative to that of their equally able white peers. From the group perspective, homophily facilitates sustaining monopolies over scarce resources (Tilly 1999). From an individual perspective, however, it is precisely *heterophilous* ties (especially to persons higher status than oneself) that enhance mobility (Granovetter 1973; Lin, Ensel and Vaughn 1981; Portes 1998). We apply these insights to the system level to contend that homophily, which facilitates collective action to sustain monopoly, increases intergroup inequality, whereas heterophily, which facilitates mobility, tends to reduce it.

The systematic effects of homophily on inequality in a real social system are more complicated than individual-level models imply due to the fact that individuals have multiple identities (*e.g.*, identities based on gender, race, age, religion, occupation, nationality, educational background, avocational interests, and so on). As Blau (1977) demonstrated, as long as these identity dimensions are only moderately correlated with one another ("intersecting" rather than "consolidated" in his terminology), and as long as homophily operates with respect to more than one identity dimension, an interaction choice that is homophilous with respect to one dimension (for example, occupation) may introduce heterogeneity into a network with respect to another dimension (*e.g.*, gender or religion). So a male Episcopalian banker who acquires a useful financial tip at a business lunch (homophily by occupation) or alumni gathering (homophily with respect to education) may spread it to shop owners and Catholics at a neighborhood poker game (homophily by residence and probably gender) and to women and Ph.D.s (or college dropouts) over Thanksgiving dinner (homophily by kin relation). Watts (1999; and Watts, Dodds and Newman 2002) develop this insight formally in the context of the small world problem, but it is likely to apply to network-generated inequality as well.

Because we focus on inequality induced by patterns of choice (rather than by collective action), we develop models in which homophily only induces inequality in the presence of network externalities -- that is, when individual choices are interdependent. We argue that *the interaction between network externalities and social homophily is a critical mechanism for the production and reproduction of social inequality* in access to new goods and engagement in innovative practices. To the extent that adoption of new goods and practices characterized by such externalities contributes to individual life chances, this mechanism may also increase overall intergroup social inequality. In order to understand the implications of research on externalities and homophily for macro-level intergroup inequality, we must specify the mechanisms by which individual choices interact. The most useful instruments to this end are models of diffusion and contagion.

Models of Diffusion

There are innumerable models of diffusion (see Rogers 2003 for an exhaustive review). but we are interested only in those that rely on interdependence of choice between agents and their network alters as a primary mechanism predicting adoption (and sometimes disadoption). These latter are ordinarily referred to as "contagion models" (Burt 1987), because they depict choices as spreading through a social network. We are particularly interested in contagion models that focus upon individual choice (as opposed to imitation or social learning).[5] Many of these posit a distribution of adoption thresholds within the at-risk population, often as a function of individual attributes and network characteristics; and in some cases, with updating on the basis of changes in the behavior of network alters. Our own model is of this kind, but diverges from the existing literature by modeling *differences in diffusion rates* within stratified populations. We focus here on studies that influence or anticipated aspects of our approach.

Although Gabriel Tarde (1903) may be considered the father of diffusion modeling, an early formal statement of ideas influencing the class of models in which we are interested was produced by Harvey Leibenstein (1950), in an essay on interdependence in consumer demand.[6] Leinbenstein distinguished between positive and negative externalities in consumption, which he termed, respectively, "bandwagon" effects and "snob" effects.[7] He also distinguished between the case in which ego's demand was a function of aggregate consumption and the case in which ego was more influenced by the consumption decisions of some consumers than others. Leibenstein recognized that a major limitation of his analysis was its reliance upon comparative statics, the assumption that

[5] We do not presume that choices are "rational" either phenomenologically or in effect. People may adopt courses of action unreflectively; or they may adopt them on the basis of poor information about costs, benefits, and risks.

[6] Leibenstein credits the insight that his paper develops to Oskar Morgenstern (1948). Morgenstern, one of the developers of game theory, contended that the interactive character of demand called for the replacement of dominant approaches in economics with models derived from game theory – a point on which Leibenstein declared himself agnostic. .

[7] Leibenstein also described "Veblen effects" – cases in which a high price makes the good or service especially attractive because of its prestige value – but these need not detain us here.

"the order of events is of no significance" (*ibid.*: 187). Subsequent progress would come in relaxing this condition and taking account of time.

A step in this direction was taken in 1957, when Coleman, Katz and Menzel introduced diffusion models based on network-driven interdependence in their study of the adoption of tetracycline by physicians in four Midwestern cities. The authors reported that the rate of adoption was brisker and penetration after 17 months more complete among physicians with three or more friendship ties to other doctors than among less connected practitioners. This difference, they argued, represented interdependence of choices among the well-connected, which led to a "snowball" or "chain-reaction" pattern as use of the new drug spread.

Mark Granovetter's (1978) paper, "Threshold Models of Collective Behavior," represented a significant advance towards the class of models employed here. The models that Granovetter developed apply to situations in which (a) agents are required to make a binary choice and can do so over a number of successive time points; (b) "The costs and benefits to the actor … depend in part on how many others make which choice" (1422) (i.e. that there are network externalities); (c) each agent has a threshold (understood as a tradeoff of costs and benefits) at which she or he will choose to act[8]; (d) agents may respond more directly to the actions of their friends and associates than to those of strangers; and (e) outcomes (the proportion of agents choosing to act and, especially, whether this proportion reaches the critical mass necessary to sustain some collective behavior) are dependent upon the distribution of thresholds and not just the mean. Granovetter describes this model's applicability to a range of cases that have

[8] Granovetter credits Schelling's model of residential segregation (1971) as source of the notion of threshold.

choice and externalities in common: the adoption of birth control, the spread of rumors, the epidemiology of disease, strikes, riots, voting behavior, college attendance, and migration. Granovetter essentially assembled all the parts necessary for the models developed in this paper, with two exceptions: thresholds are exogenous and actors are homogeneous (except with respect to thresholds and network position). Granovetter and Soong (1988) extended this model to cases of heterogeneous populations whose members respond differently to the number of adopters from different groups. In a series of papers, Abrahamson and Rosenkopf (1993; 1997; Rosenkopf and Abrahamson 1999) extend the models to organizational behavior and build in heterogeneity with respect both to network position and adopter reputation. Bruch and Mare (2006) compare threshold to other forms of diffusion models and argue that the special properties of threshold decision-making generate high and stable levels of residential racial segregation.

Note that different mechanisms may account for similar forms of diffusion (Van den Bulte and Lilien 2001). Even when decisions to adopt reflect perceived self-interest influenced by the actions of members of each potential adopter's network, any mechanism that increases the benefit, reduces the cost, or reduces the perceived risk of a practice is likely to increase it.[9] Such changes can reflect tangible benefits (*e.g.*, in our cases below, the size of the network to which an adopter has access, or the ability of high-quality information to increase returns); information about the rewards of the innovative practice from trusted or high-reputation sources (as in Rosenkopf and

[9] Processes that are entirely driven by factors external to the system can be distinguished from those driven by internal dynamics by the shape of the diffusion curve (Rossman, Chiu and Mol 2006); those driven by normally distributed heterogeneous propensities to adopt (Van den Bulte and Stremersch 2004), or those characterized by general externalities can be distinguished from those driven by alter-specific ones on the basis of whether individual choices are responsive to those of network alters, specific population subgroups (as in snob effects), total numbers of adopters (as in pure bandwagon effects), or none of the above. Van den Bulte and Stremersch (*ibid.*) and Rossman, Chiu and Mol (2006) introduce innovative ways to use information on multiple innovations to distinguish among varying mechanisms.

Abrahamson 1999); or (with some modification of existing models) information cascades based on local (network-specific) imitation (Bikhchandani, Hirshleifer and Welch 1992). Distinguishing among explanations at this level ordinarily requires qualitative informat-ion about the phenomenology of decision-making within the at-risk population.

We draw on the research we have reviewed, but diverge from many modelers by treating externalities as specific (i.e. decisions are driven only by members of one's social network and not by the number of adopters in general) and innovate by including status homophily as a key variable and by focusing on group-specific diffusion paths (and their implications for social inequality) rather than rates for the population as a whole.

Summary

In the analyses that follow, we combine insights from research on network effects, status homophily in social networks, and threshold models of diffusion to argue that diffusion processes of goods and practices with strong and identity-specific network externalities under conditions of status homophily tend to exacerbate social inequality. To do so we present two cases. First, we develop a computational model to predict group-specific diffusion paths for Internet use, varying the extent of homophily and the strength of network effects to assess the importance of each. Second, we present an empirical analysis of variation in rates of rural-urban migration in Thailand, to test the hypothesis that variations among similar villages in status homophily combined with network externalities will generate heterogeneous migration patterns from small initial differences across villages over time.

These cases are very different. The diffusion of Internet use is a conventional instance of new-product adoption where the network effects act directly on the value of the product for the agent, for whom the value of the network to which the technology provides access is a function of the number of friends and associates accessible through it. Rural-urban migration is a longstanding practice in Thailand that became much more widespread in the 1980s; network effects are indirect, in the sense that connections to prior adopters are not the source of greater value in themselves, but rather are posited to provide information that increases the returns to and reduces the risks of migration. In the Internet case we care about inequality among social groups; in the migration case, we focus on inequality among villages. Nonetheless, the two cases share the requisite characteristics for the model to apply: network effects and interdependent choice; status homophily within networks; and sequential choice by large numbers of agents.

Case 1: The Internet: Transitional Inequality or Permanent Divide?

For the first of our two cases, intergroup inequality in access to and use of the Internet, we employ computational modeling as our strategy. We do so for two reasons. First, our theory directs us to network effects, but we lack suitable network data. Second, we began this inquiry when the Internet was at a relatively early stage of penetration and we were interested in projecting its trajectory forward. Our other case, rural-urban migration in Thailand, is directly empirical, because we have detailed individual-level data from twenty-two villages over a twenty-eight year time periods (1972-2000) that enable us to track the practice from takeoff through maturity.

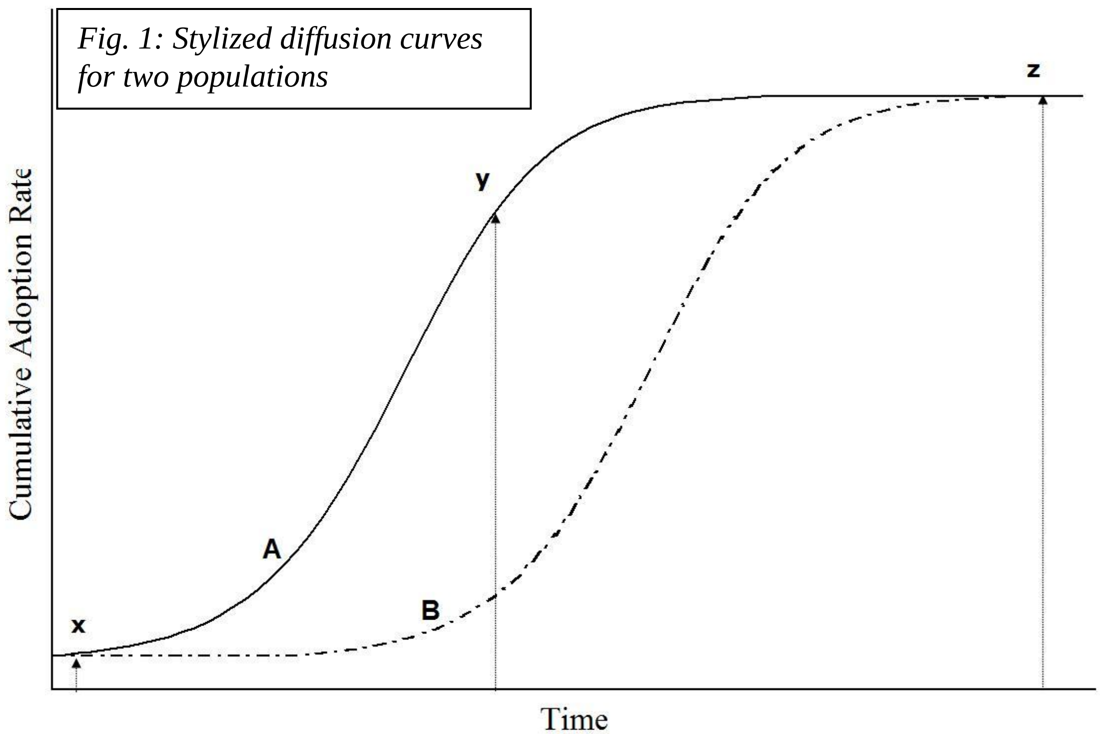
Computational models – also known as "intelligent-agent models" because they are based on simulated actors whose choices are guided by relatively simple rules – have become an increasingly important tool for theory development throughout the social sciences (Macy and Willer 2002; Watts 2003). Computational models enable the investigator to trim down the complexity of a case to the key mechanisms to facilitate theory development. In particular, computational modeling is useful for exploring interactions among individual choices when the choices themselves are amenable to theory-based predictions but the interactions among them are nonlinear and not easily apprehended through intuition. Such models indicate both the plausibility of the proposition that given mechanisms are operative (based on the fit between predicted and observed aggregate patterns) and the robustness of those mechanisms to different states of the world and ranges of key variables. Computational modeling entails a tension between realism and generalizability. We attempt to strike a balance by basing as many parameters as possible on available data, while varying two theoretically central variables as, in effect, experimental treatments.

The Problem

Although the results we report here are based on computational models, the models originated in efforts to solve a concrete empirical puzzle. At the dawn of the Internet era, some observers claimed that the technology would be equality enhancing, for two reasons: proficiency seemed associated with youth rather than socioeconomic status (Loges and Young 2001); and the technology dramatically reduced the cost to its users of acquiring many kinds of information, thus leveling the playing field (Cairncross 1997). By contrast, other observers cautioned that the rise of the Internet could simply reproduce or even exacerbate existing inequality. In this view, a variety of advantages might enable high-income, high-education persons to take advantage of the Internet more extensively and more productively than their lower-status counterparts. If this were the case, then the Internet could actually exacerbate the “knowledge gap” (and a gap in the rewards that knowledge can bring) reported by researchers studying other media (Bonfadelli 2002; DiMaggio, Hargittai, Celeste and Shafer 2004).

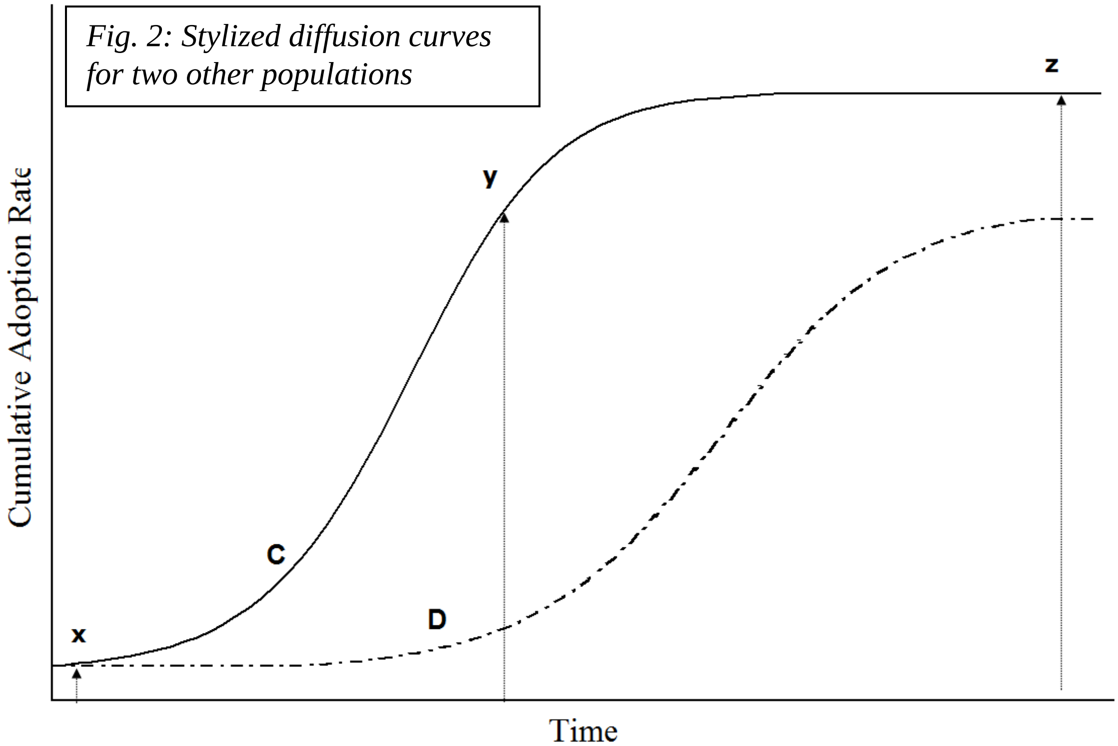
Even after a decade or so of widespread commercial availability, use of the Internet was still considerably more common among Americans with lots of education and relatively high incomes than among the less educated and the poor. This led many observers to contend that such inequality – the so-called "digital divide" – was an enduring problem and an appropriate focus of public policy. Critics of this position pointed out that intergroup inequality in adoption rates is likely whenever groups start at different baselines or reach critical mass at different times (Leigh and Atkinson 2001).

Fig. 1: Stylized diffusion curves for two populations



Consider Figure 1, which depicts stylized adoption curves for two populations, A and B, with differences measured at three points in time (*x, y,* and *z*). As is conventional, the *x* axis represents time and the *y* axis represents the population percentage that has adopted. Group B begins the adoption process and reaches takeoff after group A. If we compare rates early in the process (time *x*), we see no inequality. Because A reaches critical mass for takeoff before B, we then see increasing inequality, reaching its maximum level at *y*, just before B reaches takeoff and just as A approaches equilibrium. Trend analysis between *x* and *y* will demonstrate increasing inequality and imply a need for remedial action. Trend analysis between *y* and *z* demonstrate that inequality is declining and intervention unnecessary (which, in this stylized example, is a more accurate conclusion).

Now look at a similar figure that depicts the adoption history of two other populations. The trajectory of group C is identical to that of Group A in *Fig.* 1. Group D gets into the game at the same point as group B (in *Fig.* 1), but is slower to adopt and plateaus at a lower level of penetration. As in the previous case, inequality peaks at point *y*. But it remains considerable even at point *z.* In this case, the conclusion based on measurement at points *x* and *y* (intergroup inequality is a real problem) would be more accurate than a conclusion based on measurement at points *y* and *z* (inequality will take care of itself).



The analytic problem is that it is impossible at point *x* to know if we are in Figure 1 or Figure 2 – i.e., whether intergroup inequality will increase or decline – unless we have a theoretical understanding of the mechanisms driving diffusion to guide us. Standard diffusion research is not much help for two reasons. First, diffusion studies rarely if ever focus on group-specific trajectories or intergroup inequality. Second, they tend to focus on successful innovations, which are more likely than others to reach high levels of penetration in all populations. From the historical literature, however, we know that we cannot take the model depicted in *Figure 1* – sigmoid curves differing only in time of onset – as typical. For one thing, diffusion is not always as smooth as that model suggests: rural-urban disparities in telephone use, for example, declined in the 1920s then rose again as financially strapped farm families dis-adopted during the Depression. For another, trajectories for some populations resemble that of D in *Figure 2*. Even by the

1990s, many U.S. families lacked residential telephone service in areas of concentrated poverty (Mueller and Schement 1995).

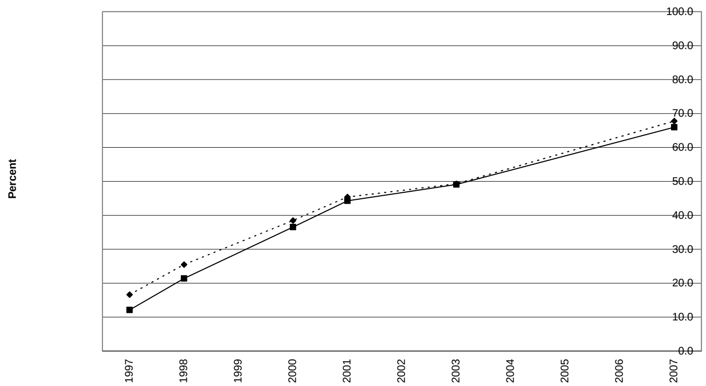
The best data on Internet use have been collected as part of the Current Population Survey (CPS) at irregular intervals from 1997 to 2007.[10] The General Social Survey (GSS) has also asked a series of questions about Internet use since 2000, which do not permit as fine-grained analysis of specific groups as the CPS (because the N is much smaller), but which are rich in covariates and serve as a useful supplement to the latter.

One can examine trends in inequality by looking at penetration rates for different groups in the years that CPS conducted surveys. We depict the same data in two ways: as actual rates in Fig. 3 and as odd ratios comparing paired groups in Fig. 4. We focus on use of the Internet at home, rather than at work, at school or library, or at other locations, because home use ordinarily provides the greatest autonomy and opportunities for learning and experimentation (DiMaggio, Hargittai, Celeste and Shafer 2004). Such autonomy is important because it takes about two years for most people to become competent in finding information on-line (Eastin and LaRose 2000) if indeed they ever do (Hargittai 2003). Moreover, Internet use at home is consequential, boosting earnings for workers even after controlling for technology use at work (DiMaggio and Bonikowski 2008).

Several features are striking. First, most types of inequality declined between 1997 and 2007. (Note that because the graphs in Fig. 4 depict change in odds ratios, inequality declines as curves moves towards 1.0 from either direction.) The advantage of men over women in Internet use at home was much greater before 2000 than thereafter.

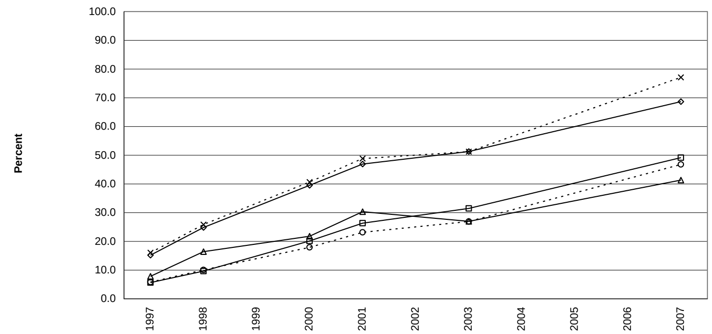
[10] Between 1997 and 2003, the items were sponsored by the National Telecommunications and Information Agency, the federal agency responsible for implementing congressionally mandated universal telecommunications service. The 2007 items were asked as part of a module on school enrollment.

Internet Use at Home by Gender



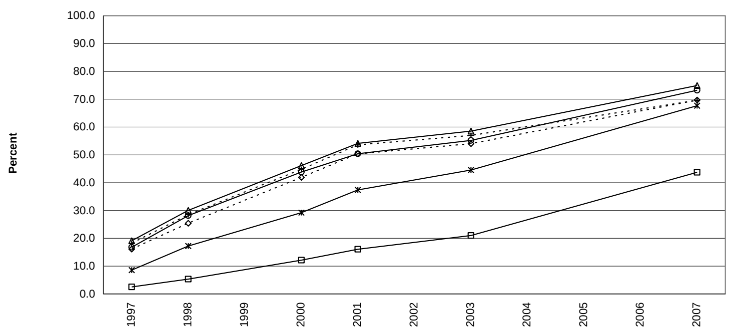
	1997	1998	2000	2001	2003	2007
Male	16.6	25.5	38.4	45.4	49.3	67.8
Female	12.1	21.4	36.5	44.3	49.1	66.0

Internet Use at Home by Race



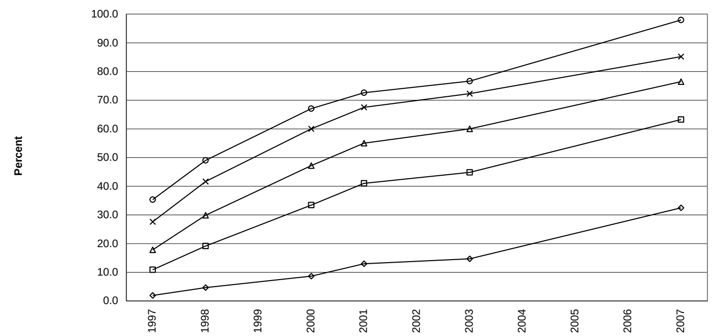
	1997	1998	2000	2001	2003	2007
White	15.2	24.8	39.5	46.9	51.3	68.6
Black	5.6	9.7	20.1	26.3	31.5	49.1
Asian	16.1	25.8	40.6	48.8	51.2	77.1
Am. Indian	7.8	16.4	21.8	30.3	26.9	41.3
Hispanic	5.8	10.1	17.9	23.2	26.9	46.8

Internet Use at Home by Age



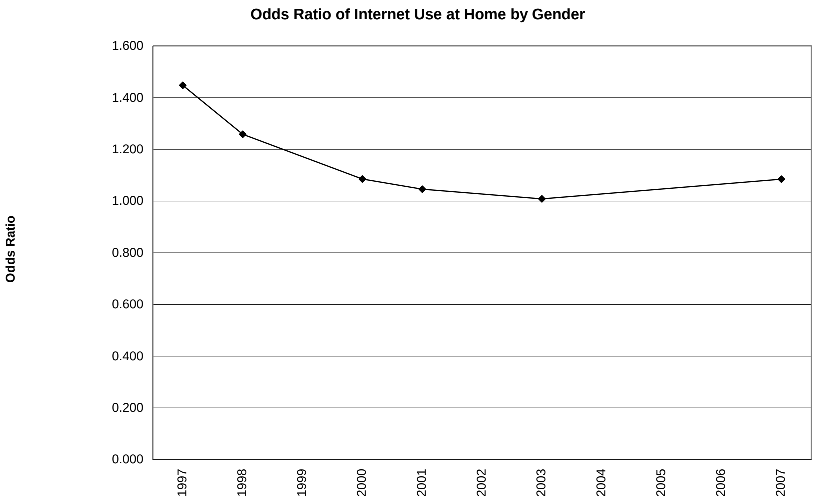
	1997	1998	2000	2001	2003	2007
18 - 25	16.1	25.4	41.9	50.4	54.0	69.7
26 - 35	18.2	28.5	44.8	53.6	57.0	69.6
36 - 45	19.1	30.0	46.2	54.1	58.5	74.9
46 - 55	16.8	28.1	43.8	50.4	55.2	73.2
56 - 65	8.6	17.2	29.2	37.4	44.5	67.6
> 65	2.5	5.3	12.2	16.0	21.0	43.7

Internet Use at Home by Education

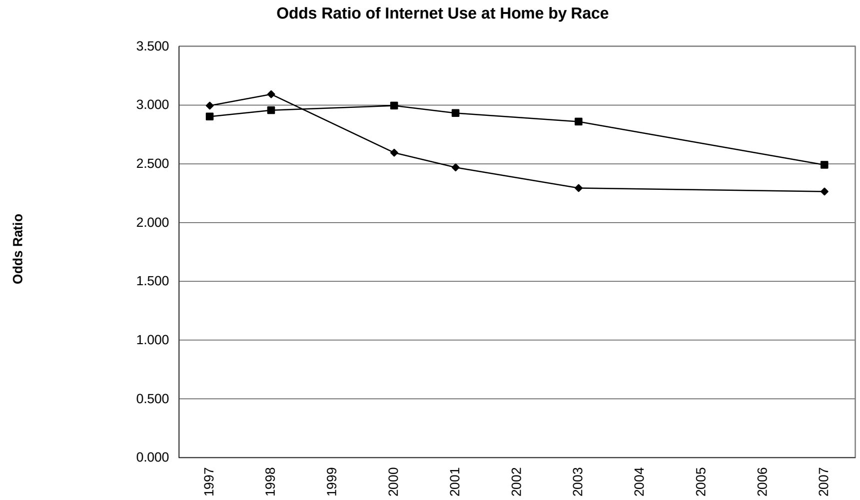


	1997	1998	2000	2001	2003	2007
Less than HS	1.9	4.7	8.7	12.9	14.7	32.5
High School	10.9	19.1	33.4	41.0	44.8	63.2
Associate	17.8	29.9	47.2	55.0	60.0	76.5
College	27.6	41.7	60.0	67.5	72.3	85.2
Advanced	35.3	49.0	67.1	72.6	76.6	98.0

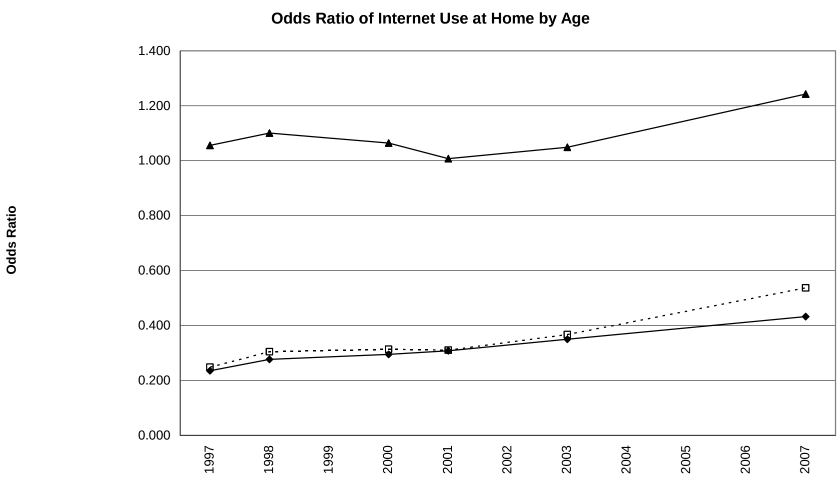
Figure 3: Percentage of Americans 18 and over with Internet service in their homes, 1997-2007. *Source:* Current Population Survey. Graphs created by Bart Bonikowski.



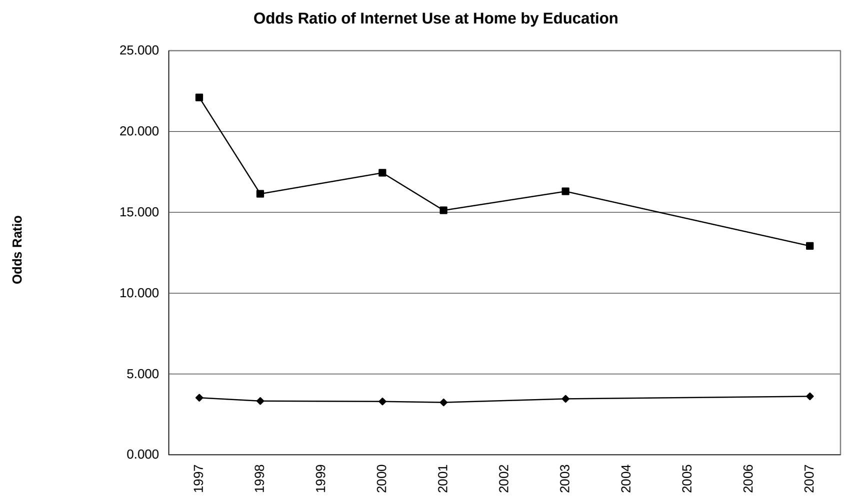
	1997	1998	2000	2001	2003	2007
Male vs. Female	1.447	1.258	1.085	1.045	1.008	1.085



	1997	1998	2000	2001	2003	2007
White vs. Black	2.994	3.091	2.594	2.469	2.293	2.264
White vs. Hispanic	2.901	2.955	2.995	2.931	2.858	2.490



	1997	1998	2000	2001	2003	2007
56+ vs. 36 - 55	0.235	0.277	0.295	0.308	0.350	0.432
56+ vs. 18 - 35	0.248	0.304	0.314	0.310	0.367	0.537
36 - 55 vs. 18 - 35	1.056	1.101	1.064	1.007	1.049	1.243



	1997	1998	2000	2001	2003	2007
College+ vs. HS	3.529	3.325	3.291	3.233	3.456	3.611
College+ vs. LHS	22.104	16.138	17.437	15.123	16.291	12.916

Figure 4: Ratios of odds that selected comparison groups of Americans 18 and over have Internet service in their homes, 1997-2007.
Source: Current Population Survey. Graphs created by Bart Bonikowski.

There was also some movement towards greater equality between non-Hispanic whites, on the one hand, and both African Americans and Hispanics on the other; and between men and women older than 55 and persons in younger cohorts. By contrast, the advantage of college graduates over high school graduates without college has been constant throughout this period.

Attempts to model odds ratios at successive cross-sections in models with a variety of controls demonstrate no trend towards greater equality on dimensions other than gender and metropolitan (as compared to rural) residence. DiMaggio and Cohen (2004), compared odd-ratios (with controls) for models based on CPS data collected in 1994 (when respondents were asked if they had modems in their home), 1997, 1998, 2000, and 2001. They reported increasing inequality between non-Hispanic whites and Hispanics, no trend (with controls) in use by African-Americans and whites, an early decline and then stability in the effects of a college education, a decline and then a rise in inequality based on white-collar *vs.* blue-collar employment, and an increase in the effect of income between 1997 and 2001.

These results suggest a paradox: If new Internet adopters are less upscale than previous adopters (as the CPS data indicate) then why are the effects of such factors as race, education, and income (as revealed by multivariate regressions) stable or rising? DiMaggio and Celeste (2004) addressed this apparent paradox by analyzing a subset of 2000 GSS respondents who were reinterviewed 18 months later. They reported that new adopters indeed had lower incomes and less education and were less likely to be non-Hispanic whites than early adopters. Yet among those at-risk for adoption in 2000, income, education and age (but not gender or race) were significant predictors of adoption by fall 2001. The authors concluded that socioeconomic queuing accounted for persistent effects of socioeconomic status on Internet use despite the relative democratization of access between 1995 and 2001. The authors also demonstrated that persistent disadvantage was reinforced by a greater likelihood that low-SES users had *dis*-adopted between 2000 and 2001.

A final means of anticipating trends in inequality is to project evidence on group-specific adoption and *dis*-adoption rates into the future until an equilibrium of adopters and disadopters is reached for each group.[11] Basing this approach on data from the 2000/2001 GSS panel, yielded the following predictions: an overall penetration rate of 75 percent of adults; persistent but modest differences between whites and African-Americans; a significant disadvantage for people with family incomes of $20,000 or less but substantial convergence above that level; converging rates for people with any amount of higher education, but persistent disadvantage for people with high school degrees only and, especially, for those who do not complete high school (DiMaggio and Celeste 2004).

So, after all this, do we know if the digital divide is a permanent or transitional problem? The analyses provide clues, but none is sufficient. We have seen why comparison of trends is unsatisfactory: as long as trajectories are right-censored, the same end point is consistent with multiple equilibria. Multivariate analyses have the virtue of acknowledging that one's technology use is a function of multiple statuses, but suffer from the same right-censorship as do simple plots. Projecting from group-specific adoption and disadoption levels overcomes right censorship, but requires the heroic assumption that rates observed over a specific 18-month period will be stable indefinitely.

[11] We are grateful to Steve Martin of the University of Maryland for suggesting this approach. We are also grateful to Professor Martin for alerting us to the advantage of odds ratios as a measure of inequality over time.

The problem with all of these approaches is that none of them rests on a theoretical account of the mechanisms that generate intergroup inequality in technology use. For that reason, we supplement these empirical approaches by developing a computational model that highlights the mechanisms we regard as most important: diffusion based on choice influenced by network externalities under condition of homophily.

The Model

Our computational model produces artificial worlds populated by 2,257 heterogeneous agents and characterized by variation (among worlds) in assumptions about the presence and type of network externality and about the degree of homophily. We use these worlds to explore how inequality in use of the Internet based on race (African-American or white), income, and education varies over time as a function of the presence and type of network effect (none, general, or identity-specific) and as a function of the strength of pressures towards homophily. To ensure that the distribution of parameters and the associations among them are realistic, we base our agents upon 2257 African-American and white respondents from the 2002 General Social Survey, which includes an item on network size, as well as data on race, education, and income.[12]

Each agent has a *reservation* price, a price at which she or he will purchase Internet service. This price, which is in effect an adoption threshold, is updated at each time period. At the end of each period, each agent compares the price of Internet service

[12] Respondents were first asked "“Not counting people at work or family at home, about how many other friends, or relatives do you keep in contact with at least once a year?” We use the network size measure generated by the follow-up to this question: “Of these friends and relatives, about how many would you say you feel really close to, that is close enough to discuss personal or important problems with?” GSS reports income as a series of ranges: We treated income as uniformly distributed within each interval and randomly assigned individuals to points in their distributions. Individuals who reported family incomes of $110K or more (about 10% of the sample) were randomly allocated incomes up to $650,000 based on CPS data on actually income distributions in that range.

(which is subject to economies of scale and therefore changes each period as a function of the number of adopters) to her or his reservation price, and either adopts or fails to adopt. The actual price of Internet service declines as more agents adopt, so that each agent may adopt because economies of scale have driven prices to or below his or her reservation price; or because more members of her or his network have adopted (raising the reservation price); or due to a combination of both. At the end of each period, the cumulative percentage of agents with different races, income levels, and educational degrees who have adopted is reported and that time point is added to a graph that records 100 periods. Our analysis focuses on the way in which two variables – the nature and extent of network externalities and the degree of homophily -- influence the extent and nature of intergroup inequality throughout this process.

Demographic characteristics. Agents have three characteristics: race, income and education. Race is Black or White, with other respondents excluded. Values are those reported in the 2002 GSS and they are stable for agents across trials and periods. Race and log income are correlated at .126; race and education at .129; and income and education at .290.

Reservation prices. Following Economides and Himmelberg (1995), we model ego's reservation price (r_{it}) as a function of ego's income (y_i), a normally distributed random perturbation with a mean of 0 introduced in the first period, and the percentage of adopters in ego's network (defined as all agents in the general-externalities condition and as those to whom ego is directly tied in the specific-externalities condition) in the previous time period (n_{it-1}). Of these parameters, only the third (adopters in ego's network) is updated in each period. We assume that the Internet has an intrinsic value in the absence of network effects and that the effect of network size varies by income.

$$(1)\ r_{it} = \underbrace{k \cdot y_i^{\gamma}}_{\substack{\text{Pure income} \\ \text{effect (i.e. the} \\ \text{innate value)}}} + \underbrace{y_i^{\gamma} \cdot \delta \cdot n_{it-1}^{\alpha}}_{\substack{\text{Network effect} \\ \text{(depends on} \\ \text{income)}}}$$

We assign the parameters *k*, *δ*, *γ*, and *α*, where *k* and *δ* are multiplicative constants for the pure income effect and for the network effects, respectively; and a and *γ* are exponents of income and the proportion of adopters, respectively. Note that this functional form implies that the reservation price increases with income and the proportion of adopters, but that the marginal effect of additional units of these variables declines so long as:

$$(2)\ \begin{array}{l} 0 \le \alpha \le 1 \\ 0 \le \gamma \le 1 \end{array}$$

We can simulate internet adoption rates by assigning reasonable values to the parameters $k, \gamma, \alpha, \delta$. We start by choosing the exponents of income and network size, α and γ:[13]

$$(3)\ \alpha = \gamma \cong 0.5$$

In order to choose the multiplicative constant of the pure income effect, *k*, we note that, in our simulations, we set the number of adopters to zero initially. So, at *t=0*, reservation prices of individuals are given by:

$$(4)\ r_{it} = k \cdot y_i^{\gamma}$$

Assuming that in the absence of any network externalities only 1 percent of the population would adopt internet use:[14]

$$(5)\ k = 0.1$$

[13] The reservation price model is too complex for the parameters to be estimated. However, we can estimate the parameters of a reduced model which assumes no pure income effect (that is, k=0). The details of this calibration exercise, based on Economides and Himmelberg (1998), are available from the authors upon request. Parameter estimates using OECD aggregate data on (off-peak) internet prices and adoption rates across countries are similar to the ones that we use for the computational models.
[14] Results are robust to a wide range of values of *k*.

Network externalities. Network externalities are varied across implementations of the model to take one of three possible states: (a) No externalities, in which case δ=0 and the reservation price is a function of income only; (b) General externalities (i.e., the percentage of the entire population that has adopted influences ego's probability of adoption, but ego is indifferent as to the identities of prior adopters), in which case δ>0; and (c) Identity-specific externalities (only a member of ego's network can influence ego), in which case δ>0 and the network of adopters for each agent is a subset of her or his personal network.

Absent network externalities, we set δ=0, and reservation price becomes a function of income only. In the case of general or specific network externalities, we must assign δ>0. We assume that at the limit of full saturation, the network effect on reservation prices should equal the pure income effect, which becomes the case if:

$$(6)\ \delta = 0.10$$

Internet price (i.e. the cost of Internet service) is a declining function of the overall number of adopters.

$$(7)\ p_t - p_{t-1} = a \cdot n_{t-1}\left(p_{\min} - p_{t-1}\right)$$

Price for each time period is the sum of the price for the last period and a price decline component that depends on the total number of subscribers in the previous time period, n_{t-1}. Note that p_{min} is the equilibrium price level and the term $a \cdot n_{t-1}$ is the speed of reversion to that equilibrium price. This model implies that the equilibrium level attracts the prices in its direction with a speed that is higher the farther away the price is from its equilibrium value. Thus prices for Internet service drop sharply at first and then stabilize once the minimum technology cost level is reached. We use internet (off-peak) price data and adoption data from OECD member countries (1998-2000) to estimate the parameters a and p_{min}[15]. This yields

$$(8)\quad \begin{aligned} p_{\min} &= 28.74 \\ a &= 3.34 \end{aligned}$$

We set the initial internet price to be:

$$(9)\quad p_0 = \$60.00$$

Using the parameter values above, the implied recursive equations of price and reservation prices are:

$$(10)\quad r_{it} = 0.1 \cdot y_{it}^{0.5} + 0.1 \cdot y_{it}^{0.5} \cdot n_{t-1}^{0.5}$$

$$(11)\quad p_t - p_{t-1} = \frac{3.34}{12} \cdot n_{t-1}\left(28.74 - p_{t-1}\right)$$

Social networks and homophily. We endow each agent with the number of ties reported by the GSS respondent on whom the agent is based (see note 12). As is typical of real-world networks, the distribution of ties resembles a power-law distribution, with most individuals having a low number of ties and a few individuals having a very high number.

Agents' networks are populated with other agents in the following way. For each agent, a social distance metric is computed to measure the distance in terms of income, education and race of each agent from every other agent in the population. We choose Euclidean distance as our metric, and assign weights to race and education based on McPherson et al.'s (2006) analysis of heterogeneity in Americans' core discussion networks.[16] Social distance (sd) between any two agents i and j is given by:

---

[15] OECD data were acquired from OECD (2003) (Table 6.6 on p. 174), OECD (2005) (Table 5.1, p. 148), and OECD (2007) (Table 5.1, p. 156).

[16] We use the authors' estimates to compute relative heterogeneity (that is, the extent to which personal networks are heterogeneous relative to the population at large), and then invert this to produce a measure of

$$(12) sd(i,j) = \|I - J\| = \sqrt{(W_I(Inc_i - Inc_j))^2 + (W_E(Educ_i - Educ_j))^2 + (W_R(Race_i - Race_j))^2}$$

where weights of income (W_I), education (W_E), and race (W_E), are:

$$(13) \begin{array}{l} W_I = W_E = 0.53 \\ W_R = 0.83 \end{array}$$

How the networks are constructed depends upon which of several homophily conditions is applicable to a particular implementation of the model. Homophily ranges in value from zero (no homophily) to perfect homophily When there is no homophily, random networks are generated such that each agent has an equal chance of establishing a tie with any other until the target number of ties for an agent is reached.

Under conditions of homophily, networks are generated as follows: Each agent, *i*, establishes a random tie with an individual *j* who is among the *n* agents closest to the individual *i* in terms of the social distance metric. We refer to this set of agents as *i'* s "in-group." In our application, *n* is arbitrarily chosen to be three times the target number of relations for that person (so that in the case of perfect homophily, all ties can be established with in-group members without violating the target number of ties for alters).

Ties between agents are established such that homophily bias occurs with probability τ. That is, the probability of having a tie to individuals from one's own ingroup, P(T), is given by:

$$(14)\ P(T) = \tau + [1- \tau].\ P_R(T)$$

---

relative homophily by education and race. (Relative, rather than absolute, network heterogeneity is used to take into account the differential variability of race and education in the overall population.) Because GSS did not collect data on alters' incomes, we set the weight of income equal to that of education.

where $P_R(T)$ is the random probability of forming a tie (Skvoretz 1990). In different implementations of the algorithm, homophily bias is assigned the following values: 0, .25, .50, .75, and 1.0 Note that even under a homophily bias of 1.0, networks will be somewhat heterogeneous because income, education, and race are only moderately correlated. Therefore individuals who satisfy the condition of being "closest" to ego and therefore eligible for selection as homophilous may not be identical to ego on each dimension.

Results

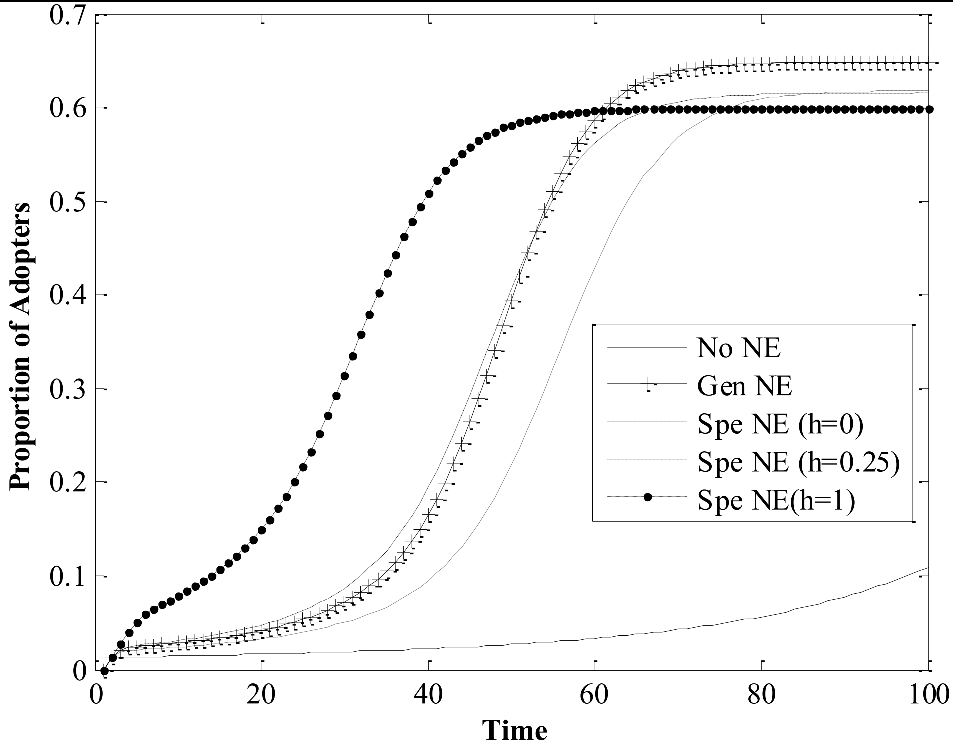
The underlying imagery of the model is one in which individual choices affect the reservation prices of close associates and adoption cascades through micro-networks, retarded both by poverty (people without much money have lower reservation prices [i.e., they are not willing to pay as much], other things equal, than people with higher incomes) and by homogeneity within ego networks resulting from homophilous choice of network alters. This intuition leaves open questions about how local adoption pressures lead to aggregate inequality. How large a difference do externalities make in speeding up adoption rates for groups with initial advantages? How much does homophily retard the spread of adoption across different types of respondents? What level of homophily is required to have an effect? And how robust are homophily effects to the presence of bridges that link Blacks and Whites, or people with different levels of income and education. The modeling apparatus is an attempt to address such questions.

We follow the adoption paths of our 2,257 agents for 100 time-periods (months) under five different scenarios (1) no network externalities, (2) general network externalities, (3) specific network externalities and no homophily (h=0), (4) specific network externalities and some homophily (h=0.25), and finally, (5) specific network externalities and total homophily (h=1.0). (More detailed analyses include cases where h=.50 and h=.75, as well.) For each condition, we undertake 1000 simulation runs, reporting the mean values for each time period.

The analysis contains four steps: First, we look at the impact of the different conditions on overall diffusion trajectories, for the population as a whole and for groups within it, focusing on the rapidity of takeoff (the slope of the curve) and on the equilibrium diffusion rates. Second, we examine variation among conditions in equilibrium inequality between pairs of groups defined on the basis of race, education, and income, respectively. Third, we undertake logistic regressions of the predictors of adoption at each time point to examine the net impact of race, education and income under different conditions and at different stages of the process. Fourth, we use regression analysis to understand how externalities and homophily interact to produce particular features of diffusion processes and different kinds of inequality.

Overall adoption rates. In the absence of externalities, Internet adoption never reaches critical mass to reduce prices, and therefore adoption proceeds very slowly, with around only 10 percent even after 100 periods (Fig. 5). Adoption rates under other conditions resemble the conventional sigmoid curve, rising slowly, then sharply, then leveling off. Adoption under general externalities (egos rate increases as a function of the proportion of adopters in the population) takes off between periods 30 and 40, rising

Figure 5: Diffusion Trajectories for Five Conditions of Network Externalities and Homophily



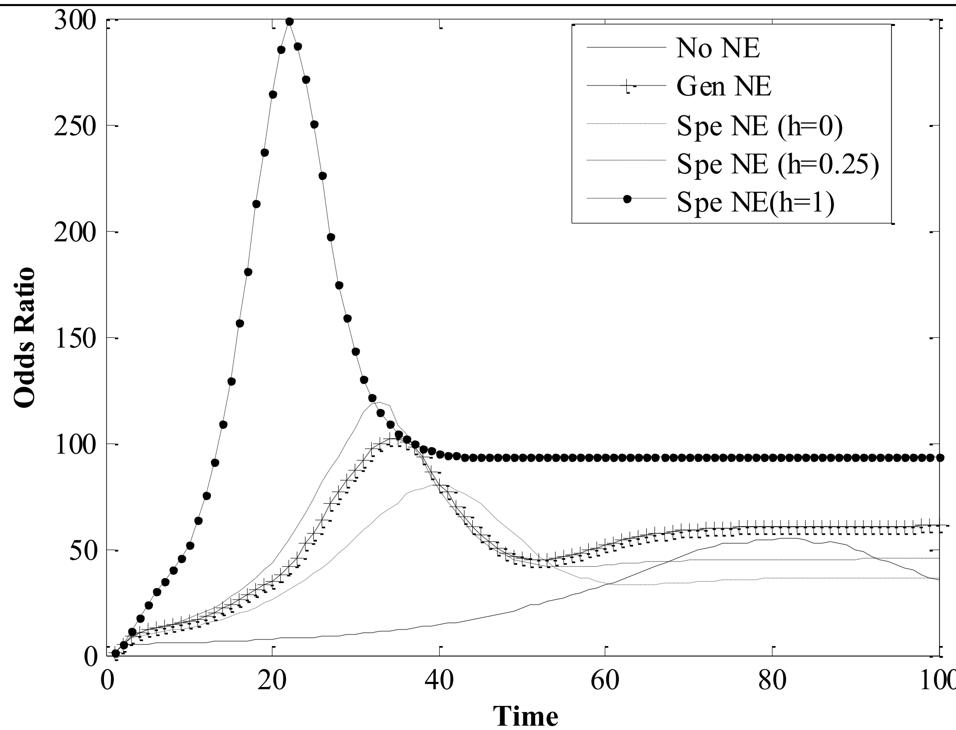
sharply to nearly 65 percent at equilibrium, the highest level of adoption. Introducing specific externalities *without* homophily scatters the effects of new adoptions at random through the population, rather than applying them to everyone as in the general-externality case. The result appears to be that some agents marooned in non-adopting networks, with adoption slowed (the takeoff begins after period 40) and reaching a lower rate at equilibrium. Combining network externalities with homophily provides a more rapid uptake, but lower penetration overall. When homophily is modest (h=.25) the trajectory is similar to that for the general-externalities condition through period 55, then plateaus more quickly, ending up at about the same penetration as specific externalities without homophily. Strong homophily (h-1.0) stimulates early adoption sharply, as adoption cascades among homogeneous networks, inducing takeoff around period 20 – but limits the reach of network effects, eventuating in an equilibrium adoption level below that of both the general-externalities condition and the other specific-externalities conditions. Middling levels of homophily (h=.50 and h=.75, not shown here) achieve takeoff in between h=.25 and h=1.0 (around periods 25 and 30) and reach equilibrium levels between those of h=.25 and h=1.0.

To summarize the results for population diffusion rates: Absent externalities, with reservation prices normally distributed and prices declining at a decreasing rate as the market expands, adoption takes off very slowly, or not at all. By contrast, network externalities of any kind induce takeoff, but do so in different ways. General externalities produce a standard S-curve, with a relatively slow uptake but relatively high penetration at equilibrium. Network-specific externalities in a random network describe a very similar pattern, but plateau at somewhat lower levels of adoption. Introducing just a modest amount of homophily in the specific network condition quickens the rapidity of diffusion, with relatively little impact on equilibrium diffusion levels. Adding homophily makes the slope steeper, while reducing penetration. Thus it appears that the optimal state for adoption (with respect to speed of diffusion and ultimate penetration) is that of general network externalities or, failing that, one in which network-specific externalities are associated with only modest tendencies towards homophily.

Forms and Degrees of Inequality. We begin with inequality in rates of Internet adoption by income and then address inequality based on education and on race. We divide agents into three income classes of equal size, high (>$55,000), medium ($30,000-$54,999), and low (<$30,000). Because the model includes a direct income effect *and* (where applicable) interactions between income and network effects, income has a very strong influence on adoption. In analyses available upon request, the top income class has an equilibrium adoption rate of more than 90 percent in all conditions with externalities, with the highest rates associated with the highest level of homophily. Rates for the middle class range from 63 to 70 percent, with rates highest for general externalities and descending with the extent of homophily. By contrast, equilibrium rates for the lowest income class range from 18 percent with specific externalities and complete homophily to 23 percent with network-specific externalities but random networks (no homophily) and 25 percent with general network effects. Penetration rates decline as homophily rises between 0 and 1.0. In other words, given specific network externalities, homophily tends to enhance adoption rates among the most prosperous and suppress adoption rates among the least privileged.

Figure 6 depicts the ratio of the odds for the highest to the odds for the lowest income group over the course of diffusion. Differences are sharpest with specific externalities and complete homophily (an odds ratio at equilibrium of almost 100:1) and, secondly, for general externalities (50:1). In the former, just after period 20, when all of the

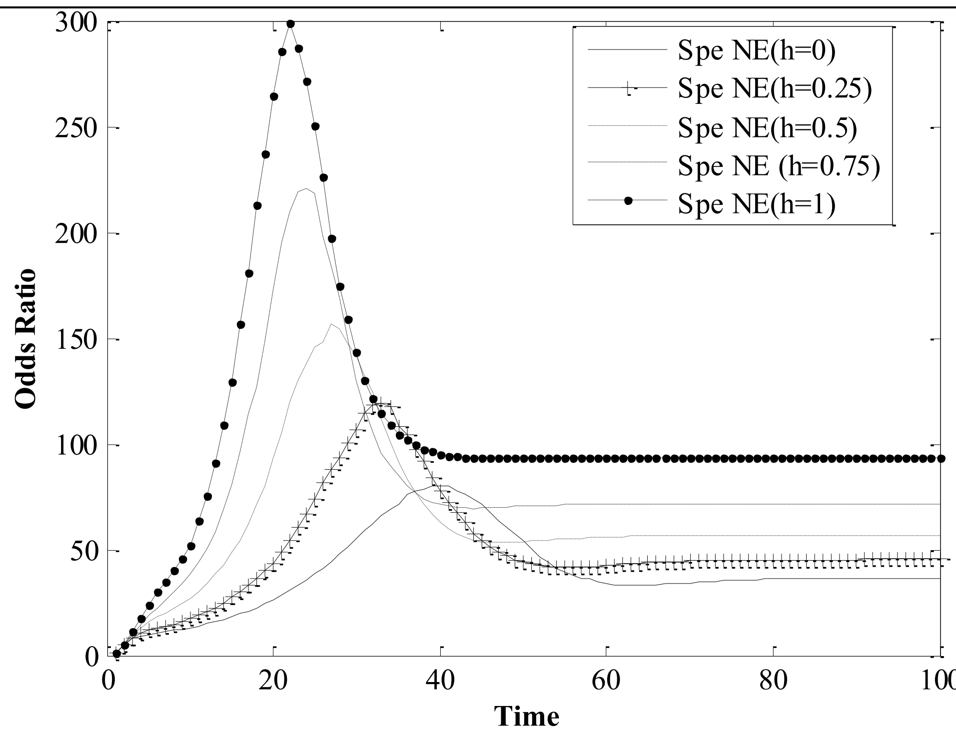
Figure 6: ***Odds Ratios of Diffusion Rates for Highest as compared to Lowest Income Classes in 5 Conditions of Externalities and Homophily***



high-income agents have adopted but none of the low-income agents have done so, the odds ratio reaches 300:1, dropping off as low-income adoption increases, but leveling off at period 40. With network-specific externalities and more modest homophily, inequality spikes up less rapidly and descends more modestly, albeit remaining substantial/

Given network-specific externalities, increments in homophily enhance the speed of uptake and the equilibrium rate of adoption for the more privileged group and increase the equilibrium advantage of the more over the less privileged group. All effects are monotonic to the degree of homophily (Fig. 7). The higher the level of homophily, the more rapidly inequality increases, the higher the rate at which it peaks and the gap between the peak and the equilibrium level, and the higher is inequality at equilibrium. A methodological implication of this is that the danger of overestimating inequality by observing it in mid-process is marked for all conditions, but greatest at high levels of homophily.

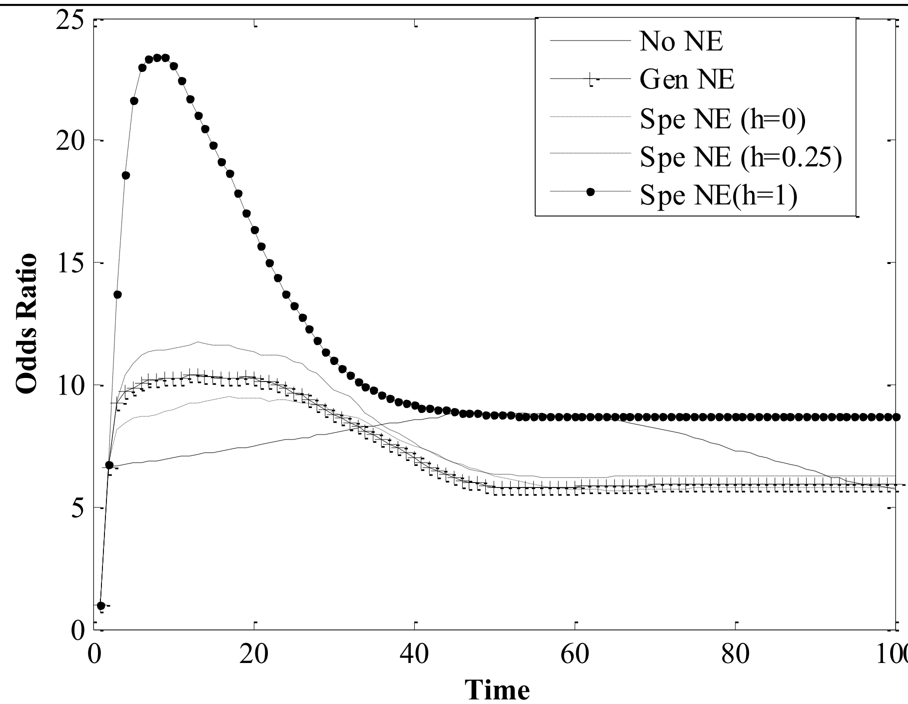
Figure 7: ***Odds Ratios of Diffusion Rates for Highest as compared to Lowest Income Classes Given Specific Externalities and Five Different Levels of Homophily***



To summarize the results for income inequality: Other things, equal, externalities promote diffusion, but homophily disproportionately benefits groups with initial advantages, and the greater the homophily, the greater the inequality.

Educational inequality. There are three education classes of equal size: High education (more than 16 years); moderate education (12 to 16 years); and low education (<12 years). Unlike income, education has no direct effect upon the reservation price. Its sole influence comes from the correlation of education with income. It follows that differences in predicted equilibrium rates for different education levels are somewhat less extreme than for comparable income levels. In the general externalities condition (where penetration is greatest), equilibrium rates (graphs available on request) are just under 80 percent for college graduates, almost has high for agents with some college, 61 percent for high school graduates without college, and 42 percent for agents lacking high school degrees. As was the case for income, the greater the tendency toward homophily, the steeper the slope of odds ratios between education levels, the more marked the difference

Figure 8: ***Odds Ratios of Diffusion Rates for College Graduates as compared to Agents without High School Degrees in 5 Conditions of Externalities and Homophily***

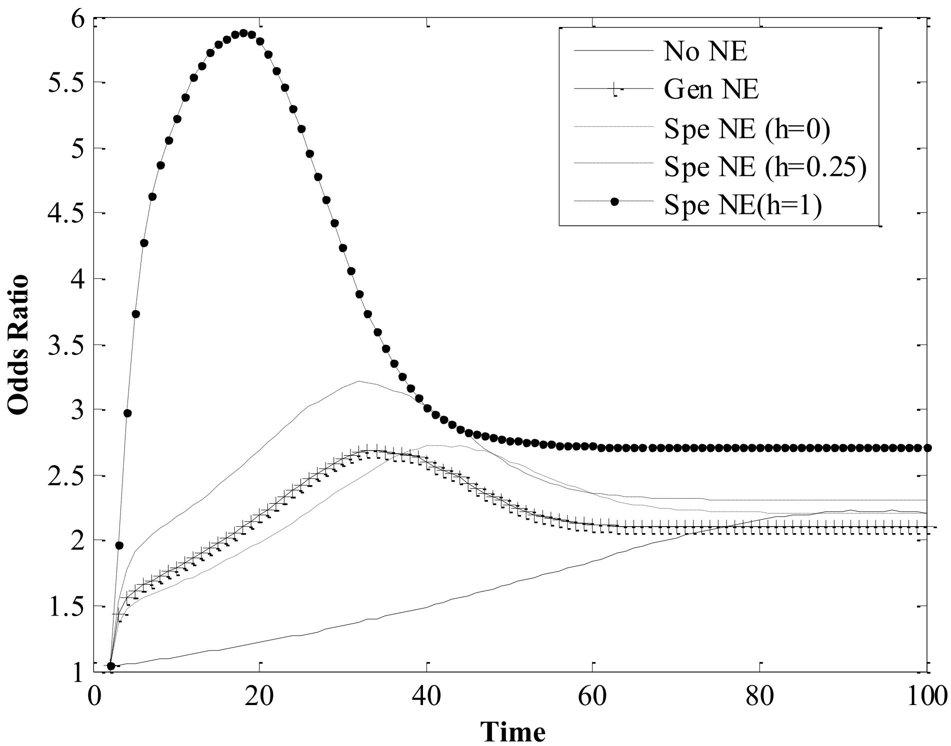


in equilibrium states by class, and the greater the difference between peak and equilibrium odds ratios.

For both the general-externality and the random-network-with-specific externalities conditions, differences between the most and least educated rise slowly, peaking between periods early and and then leveling off to an odds ratio of about 6:1 for each condition. A little homophily induces a small increase in peak and equilibrium inequality. With a larger increase, to h=1.0 (highest possible homophily bias), inequality rises far more steeply and peaks at an odds ratio of about 9:1. Once again, then, we find that externalities enhance the rate and extent of diffusion, and that strong homophily and specific externalities combine to limit absolute diffusion rates and to exacerbate intergroup inequality.

Racial inequality. The system has two races. Race has no direct effect on agents' reservation prices, but is associated with reservation prices through its correlation with income, which is lower for Blacks than for Whites. Equilibrium diffusion for whites

Figure 9: ***Odds Ratios of Diffusion Rates for Whites as compared to Blacks in 5 Conditions of Externalities and Homophily***



(graphs available upon request) is highest (68 percent) under general externalities, but essentially the same (61-62 percent) under all specific-externality conditions.

Equilibrium adoption rates for Blacks, although lower in all conditions, are, as for whites, highest (50 percent) under general externalities. In contrast to whites' rates, however, rates for Blacks decline monotonically with homophily under network-specific externalities, from 46 with random nets to 39 percent when maximum homophily bias is at its maximum (graph including more values of *h* available on request). When we measure inequality using odds ratios, racial inequality is lowest (about 2.3:1) under general externalities. In the specific-externalities conditions, the rate at which inequality grows, peak inequality and equilibrium inequality, and the difference between the peak and equilibrium inequality all rise as homophily bias increases.

Net Effects of Income, Race, and Education

Thus far we have compared diffusion trajectories and estimated degrees of intergroup inequality under different externality/homophily conditions. In this section we present results of repeated cross-sectional regressions that enable us to plot the *net* effects of income, race, and education on adoption over time.[17]

Logistic regressions were run with 2257 agents over 100 periods under five conditions (with results for each period averaged across 25 trials per condition). Figure 10 displays results from periods 10 through 80. Depending on the externality-homophily condition, each additional year of education increases the odds of adoption at equilibrium by from 1 percent (general externalities) to 6 percent (specific externalities with maximum homophily bias). Being white increases the odds of adoption by 11 percent (general

[17] We ran separate models with an added control for network size, but this did not materially affect the results.

externalities) to 50 percent (specific externalities with maximum homophily) compared to being black. A unit increase in ln(income) (for example, an increase from $20,000 to

Figure 10: Net Effects of Income, Education and Race on Adoption over 80 Time Periods

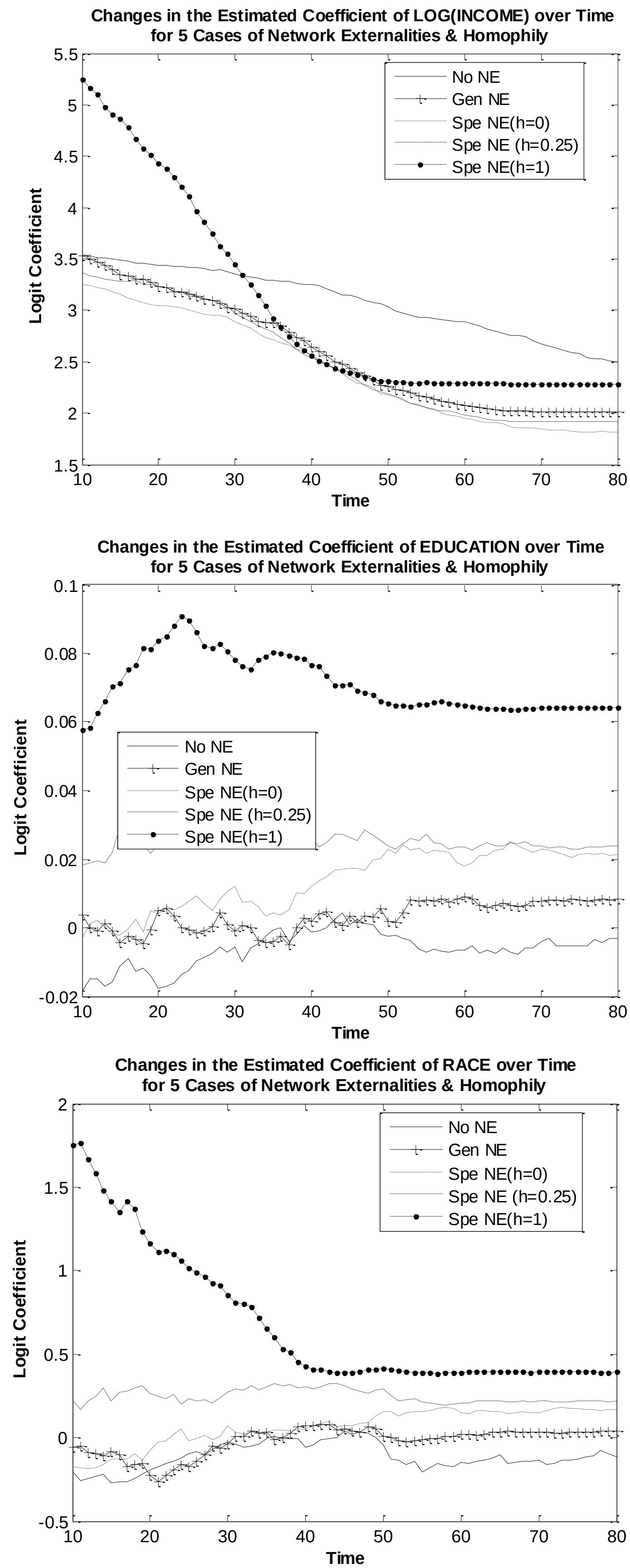


	Table 1: Logistic Regressions Predicting Adoption in Seven Conditions with Interactions													
	Conditions													
	No Externalities		General Externalities		Specific Externalities: No bias		Specific Externalities: *h*=.25		Specific Externalities: *h*=.50		Specific Externalities: *h*=.75		Specific Externalities: *h*=1.00	
Variables														
Race (White=1)	*ns*	*ns*	*ns*	*ns*	*	*Ns*	*ns*	*ns*	**	*ns*	**	*ns*	*	**(neg.)
Log (Income)	**	*ns*	**	**	**	*Ns*	**	*	**	*	**	*	**	**
Years of Education	*ns*	*ns*	*ns*	* (neg.)	*ns*	**(neg.)	*ns*	* (neg.)	*	**(neg.)	*	* (neg.)	**	*ns*
Race X Income		*ns*		*ns*		*Ns*		*ns*		*ns*		*		**
Education X Income		*ns*		*		**		*		**		*		*ns*

*p<.1,**p<.01 (two-tailed tests) Data are from runs of the computational model described in this paper. The first column under each condition reports on models that do not include interaction terms. Unless stated otherwise, all statistically significant effects are positive.

$54,364) increases the odds 5-fold (specific externalities in a random network) to 10-fold (specific externalities with maximum homophily).

The results reinforce inferences from the descriptive results presented earlier. Under the three network-externalities condition, the effects of income, education and race increase monotonically as homophily bias increases (a result confirmed by runs across other values of *h*, available upon request), providing further evidence that network-specific externalities and homophily interact to induce inequality. The patterns for education and race are similar to one another, with effects basically null absent network-specific externalities, but emerging and rising to statistical significance as homophily bias increases under specific-externality conditions. The results for income differ only in that effects are greatest without externalities and under general externalities. This result is expected because the model specifies that income is the only individual-level variable that can influence adoption decisions under those conditions. When ego-specific networks matter, the effect of income, like the effects of race and education, rises with homophily bias.

We see this more clearly in Table 1, which presents results of equilibrium models with additional conditions and interaction effects included.[18] As before (and not surprisingly given the design of the model, income is highly significant in every condition, and becomes more so as homophily increases. In models without interactions, race reaches significance only in conditions with specific network externalities, peaking when h=.75; and years of education becomes significant only with specific externalities and

[18] For the sake of economy of presentation, we report only significance levels of coefficients, for models with and without interactions between income and education and income and race. In order to calculate standard errors, we did not average across 25 runs (as we did for the coefficients plotted in Fig. 10), but used results from the implementation of each model with the median coefficient for education (as those coefficients were the most variable across runs).

only when h=.50, peaking at h=1.0. In other words, the combination of specific externalities and homophily induces inequality based on variables merely correlated with income, even though the model does not specify a direct effect on adoption choices; and the effect is stronger at higher levels of homophily. Positive interaction effects between being white and income and between education and income in most of the conditions with network externalities indicate that these network effects compound advantages associated with high income, high education, and white racial identity, and that they tend to do so more forcefully (especially in the case of race) at higher levels of homophily.

The Impact of Externalities and Homophily on Inequality

In this section we explore the effect of externalities and homophily on diffusion trajectories and different forms of inequality at equilibrium using statistical means. Data are 7000 runs of the model, 1000 for each of the seven conditions. Independent variables are dichotomous variables for each of six conditions (with a seventh, network-specific externalities without homophily). Dependent variables are indicators of equilibrium adoption levels (global and for specific population subgroups) and indicators of intergroup inequality (odds ratios comparing pairs of population subgroups).

The statistical results (depicted in Tables 2 and 3) confirm the qualitative results from the previous analyses:

1. *Externalities are necessary for takeoff to occur within 100 periods*.

2. *The highest diffusion levels occur under general-externality conditions*, when benefits from each new adopter are spread across all potential adopters.

3. *Under conditions with network-specific externalities, increases in homophily lead to declines in overall adoption levels.*

Table 2. Linear Regression of Adoption Levels on Experimental Conditions

			RACE				INCOME				EDUCATION			
	Adoption level - All		Adoption level - Whites		Adoption level - Blacks		Adoption level - High Income		Adoption level - Low Income		Adoption level - BA Education		Adoption level - Less than high school	
	coef.		coef.		coef.		coef.		coef.		coef.		coef.	
(Homophily = 0 - Reference)														
No network externalities	-0.516	**	-0.536	**	-0.399	**	-0.685	**	-0.238	**	-0.611	**	-0.351	**
General network externalities	0.030	**	0.028	**	0.043	**	0.032	**	0.017	**	0.023	**	0.030	**
Homophily = 0.25	-0.003	**	-0.001		-0.012	**	0.009	**	-0.014	**	0.005	**	-0.011	**
Homophily = 0.5	-0.005	**	-0.002	**	-0.024	**	0.017	**	-0.028	**	0.010	**	-0.024	**
Homophily = 0.75	-0.011	**	-0.006	**	-0.040	**	0.024	**	-0.046	**	0.012	**	-0.043	**
Homophily = 1	-0.019	**	-0.012	**	-0.061	**	0.029	**	-0.067	**	0.015	**	-0.068	**
Intercept	0.618	**	0.647	**	0.454	**	0.925	**	0.249	**	0.788	**	0.392	**
R^2	0.99		0.99		0.97		0.99		0.96		0.99		0.96	
N	7000		7000		7000		7000		7000		7000		7000	

**p<0.01, * p<0.05. All independent variables are binary. Variables are measured on the final period of simulations (T=100).

Table 3. Linear Regression of Logged Odds Ratios on Experimental Conditions

	RACE	INCOME			EDUCATION		
	White/Black Odds Ratio	High/Low Income Odds Ratio	High/Medium Income Odds Ratio	Medium/Low Income Odds Ratio	BA/Less than HS Odds Ratio	BA/High School Odds Ratio	Some college/ Less than HS Odds Ratio
	coef.	coef.	coef.	coef.	coef.	coef.	coef.
(Homophily = 0 - Reference)							
No network externalities	-0.02 **	-0.13 **	0.02 *	-0.15 **	-0.07 **	-0.05 **	-0.13 **
General network externalities	-0.05 **	0.51 **	0.41 **	0.10 **	0.02 *	0.01	0.06 **
Homophily = 0.25	0.04 **	0.21 **	0.15 **	0.06 **	0.08 **	0.05 **	0.06 **
Homophily = 0.5	0.09 **	0.43 **	0.31 **	0.13 **	0.16 **	0.10 **	0.13 **
Homophily = 0.75	0.14 **	0.67 **	0.46 **	0.21 **	0.26 **	0.14 **	0.22 **
Homophily = 1	0.20 **	0.93 **	0.62 **	0.31 **	0.39 **	0.20 **	0.33 **
Intercept	0.79 **	3.62 **	1.87 **	1.76 **	1.75 **	1.01 **	1.54 **
R^2	0.28	0.66	0.57	0.29	0.41	0.35	0.27
N	7000	7000	7000	7000	7000	7000	7000

**p<0.01, * p<0.05. All independent variables are binary. Variables are measured on the final period of simulations (T=100).

4. Such declines occur as a result of the differential impact of homophily on adoption patterns of privileged vs. disprivileged groups (with privilege, in this model, a function of income and attributes that are [a] a basis for homophily and [b] correlated with income). *Increases in homophily either increase (in the case of income and education) or very modestly reduce (in the case of race) adoption levels for privileged groups, while leading to marked decline in adoption levels for less privileged groups, with a net negative effect on adoption.*

5. This mechanism – homophily adding a boost to (or only very modestly retarding) adoption by high-status subgroups while reducing adoption by low-status groups – exacerbates inequality. As Table 3 indicates, the log of the ratio of the odds of adoption rby high-status groups to the odds of adoption by lower status groups increases monotonically with homophily: *The greater the homophily (under conditions of network-specific externalities) the greater the degree of inequality.*

6. *The greater the status distance between groups, the greater the extent to which homophily exacerbates inequality.* Effects for income are strongest for the top tertile as compared to the bottom tertile, with the effect on top-to-middle inequality much smaller and on middle-to-bottom smaller still but still significant. In the case of education, the effects on homophily bias inequality between college graduates and high-school nongraduates exceed the still significant effects on inequality between other subgroups..

In the absence of network externalities, inequality is less in all but one instance (the comparison of high to medium income groups) than in cases with specific externalities. Under the general externalities condition, racial inequality is at its lowest; inequality base on educational attainment is comparable to that with identity-specific externalities at very low levels of homophily bias and inequality based on income is comparable to that observed at moderate levels of homophily (h=.25 to h=.75, depending upon the comparison).

Evaluation of the Model

How well does the model capture observed patterns of diffusion? We would not expect it to do so perfectly: Our purpose has been to examine a particular set of mechanisms (diffusion of practices with network-specific externalities under conditions of homophily) and not to simulate realistically the Internet's diffusion process in every detail. Nonetheless, the model does capture some notable features of the case.

First, the predicted equilibrium adoption level (for conditions with externalities) of approximately 60 to 65 percent (household) is close to the results of the 2007 CPS (67 percent at home) and reasonably consistent with results from the tracking survey conducted by the Pew Internet and American Life Project, which reported 79 percent of Americans using the Internet at any location in April 2009.[19]

Second, intergroup inequality in access to the Internet has persisted as group-specific penetration rates have reached a plateau, contrary to the expectation that time would heal the "digital divide," as less prosperous and well educated groups caught up with early adopters. The most 2007 CPS data on home Internet use show a racial divide similar to the results from the models with specific externalities (a reasonable assumption given evidence on racial homophily in U.S. networks [McPherson, Smith-Lovin and Cook, 2001]). The low- and highest-homophily models each predicted about 62 percent

[19] For CPS results, see Figure 3. Pew data are from a the spreadsheet downloadable from http://www.pewinternet.org/Trend-Data.aspx (last accessed August 21, 2009). To be sure, equilibrium rates change in response to price-effect and network-effect parameters; but the general coincidence of predicted and observed rates is reassuring.

penetration for whites, and about 46 and 39 percent respectively, for Blacks, compared to 69 and 49 percent in late 2007 (Fig. 3). For the highest education group our models converge on a rate at equilibrium of 80 percent, with estimates of 39 and 32 percent for the lowest and highest homophily conditions for agents who did not complete high school. The lower figure squares well with the 33 percent rate for this group in the 2007 CPS, and approaches the rate for college graduates, which had risen to about 85 percent. (Comparable data for income are not available.) Overall, then, the model provides realistic estimates for subgroup rates.

This realism is actually somewhat surprising, given five ways in which our model deviates significantly from reality. First, income plays the dominant role in predicting adoption, through its direct effect and through its interaction with the percentage of network members who have adopted. The effect of income is stronger relative to that of education in our model than it is in the real world.

The second deviation from realism accounts for the first. In order to clarify the implications of particular mechanisms, we treat general externalities and specific externalities as distinct conditions. In fact, one way in which the Internet is different from the telephone (before cellular technology) is that adoption is governed by a mixture of specific and general externalities. To be sure, specific externalities – the desire to use the Internet to communicate with friends, family, and other personal contacts – were completely dominant in the early Internet (Abbate 1999; Castells 2001) and have remained primary (as instant messaging and social media have grown alongside e-mail) (Miller and Slater 2000; Kraut, Mukhopadhyay, Szczypula, Kiesler and Scherlis 1999; Hampton and Wellman 2003; Fallows, Rainie and Mudd 2004; Fox 2004).

Yet even by the late 1980s, a few users, for example, wealthy investors, were attracted to the Internet as a source of information and related utilities. With the emergence of browsers in the early 1990s and the commercialization of the Web in the mid-1990s, the amount of on-line information grew, expanding as the number of Web users increased. Under these conditions, two kinds of general externalities became important. First, some on-line activities conform to classic theories of network externalities: I am more likely to find the products I want on eBay, information on Wikipedia, stimulating political ideas in an on-line discussion, or useful tips on how to combat the virus that is slowing down my computer at a tech forum, the more other people use those sites. In these cases, contrary to the case for network-specific externalities, I may actually benefit more if the other users are strangers (people to whom I would not otherwise have access and who might therefore expose me to more new information) than if they are my friends (Granovetter 1973). Second, in addition to these classic general network effects, more Internet traffic attracts more revenue and larger investments which, in turn, increases the variety and attractiveness of on-line resources. Consistent with the expectation that general as well as network-specific externalities play a role in diffusion, when potential adopters are asked why they might want to go on-line, they usually mention information-gathering as well as communication (Selwyn, Gorard and Furlong 2005; Hargittai and Hinnant 2008).[20] Because demand for information is associated with education (Bonfad-

[20] We reran the model with an extra term added to the reservation price equation for cases with specific network externalities: $\cdot \lambda \cdot e_i \cdot N_{t-1}^{\alpha}$, where *e* represents education (a proxy for demand for information), N is the percentage of adopters in the population, and gamma and alpha are constants < 1. Qualitative results, including monotonic relationships between homophily and slope of, maximum, and equilibrium inequality, and the negative effect of homophily on equilibrium adoption for less privileged groups, were unchanged.

elli 2002; DiMaggio, Hargittai, Celeste and Shafer 2004), failing to model this relationship exaggerates the role of income relative that of education in adoption.

The model also departs from reality in depicting Internet adoption as a terminal state. Much research (summarized in DiMaggio and Celeste [2004]) indicates that many Americans who adopt the Internet (estimates range from 8 to 20 percent) subsequently discontinue its use. Consistent with Popielarz and McPherson's [1995] observation that "the closer a member is to the edge of a group's niche…the more likely s/he is to leave the group," analyses of a panel constructed from GSS data revealed that Internet users with the lowest probability of adopting in the first place are most likely to *dis*adopt (DiMaggio and Celeste 2004). A version of our model that included group-specific disadoption rates generated lower equilibrium adoption rates, produced higher levels of inequality, but otherwise retained the qualitative features reported earlier. (Results are available upon request.)

Fourth, our model deviates from reality in not taking account of the role of institutional decisions in bringing people online. About half of U.S. workers have Internet access at work (DiMaggio and Bonikowski 2008) and many workers first encounter the Internet in their workplace. Most universities provide Internet access to all of their students and require that they use it, creating hotbeds of intensive information technology use. It is possible that the "focused ties" (Feld 1981) that large organizations like corporations and universities induce among their members (Feld 1981) serve as bridges that moderate the impact of homophily on inequality. On the other hand, this will only be the case to the extent that students were *not* Internet users already; and that adults who first use digital communications technologies at work have sufficient autonomy to employ them for personal purposes. Absent reliable data on these issues, it is difficult to know how to incorporate organizational effects into a model or to predict how doing so would alter the results. Public policies aimed at increasing access are less difficult to model because they tend to act directly or indirectly upon the price of service, a parameter to which our qualitative results are robust.

Finally, our model departs from realism in treating network ties as exogenous. Real networks change constantly, of course, and it is possible (although there is no research on which one can rely) that people who fail to adopt communications technologies become marginal to their networks. We also ignore interdependence between technologies, both in the past (computer owners had a head start on home Internet service, having mastered relevant skills and purchased the necessary service [Rogers 2003: 350-51]) and prospectively, as the Internet migrates from personal computers to mobile devices of many kinds (Miyata, Boase, Wellman and Ikeda 2005).[21]

Conclusions to Technology Adoption Case

Like other research on and models of technology diffusion, our model demonstrates that network effects strongly enhance the slope and extent of technology adoption. But they also demonstrate a less appreciated effect of network externalities. When externalities are specific to the potential adopters' own networks – when, as in the case of most communication devices, ego cares less about how many people have adopted than about whether the members of her or his network have adopted – externalities interact with homophily to exacerbate the effects of social inequality on adoption.

[21] We also make no effort to model spatial heterogeneity in the supply and price of Internet service. Such heterogeneity induced initial regional and metropolitan-rural differences in trajectories of diffusion, but geographic differences have declined and their effect on socioeconomic inequality is not obvious.

In models with network-specific externalities, wherein each agent's choices are only influenced by the prior choices of those to whom he or she is directly connected, we find several robust results. First, the greater the degree of homophily bias, the steeper the slope of the diffusion curve. Second, the greater the degree of homophily bias, the lower is the equilibrium adoption rate. Third, equilibrium adoption rates are most negatively affected for the least privileged groups. Fourth, the greater the degree of homophily bias, the more quickly intergroup inequality increases, the higher the peak level of inequality, and the greater the difference between peak and equilibrium inequality. Fifth, the greater, the greater the degree of homophily bias, the greater is the degree of intergroup inequality at equilibrium, and the strength of this effect is a function of the status difference between subgroups. At the individual level, the net effects on adoption of status characteristics that do *not* affect choice directly (in these models race and education) become significant only at higher levels of homophily bias. When homophily bias is high, significant interactions between these characteristics and income suggest that homophily bias combines with network-specific externalities to further compound disadvantage.

Case 2: Internal Migration in Thailand

Our second case differs from the first in several important respects. First, whereas the Internet case explored changes over time in inequality among groups defined by race and socioeconomic status in a national population, our second case explores changes over time in inequality among villages in a region of Thailand. Second, whereas the first case employed a computational model, our second relies on empirical data collected over two decades. Finally, whereas the point of the first case was to document the impact of variation in externality strength and homophily on intergroup inequality, the goal of the second is to explore the tendency of homophily to generate inequality (in the sense of high-variance outcomes) where network effects are constant and socioeconomic differences at both the individual and village levels can be rigorously controlled.

Internal or international migration flows may begin for a variety of reasons: an attempt to increase individual or household income, a demand for low-wage workers in destination regions, or deteriorating economic conditions in sending communities. An intriguing idea in the migration literature posits that the conditions initiating migration flows may be quite different from those that perpetuate these flows over time. Although such factors as labor demand or wage differentials may continue to cause people to move, new conditions in the origin communities resulting from migration flows come to act as independent causes of migration themselves: Ties connecting individuals in origin to migrants in destination spread, acting as conduits of information or assistance for potential migrants (Massey 1990; Massey and Garcia-Espana 1987). Remittances from migrants change the income or land distribution within origin communities, propelling other individuals to migrate as well (Stark and Taylor 1989). An emergent 'culture of migration' creates expectations about the role of migration as a symbolic act of coming of age, or affirmation of identity (Levitt 1998). Through these transformations, migration flows become increasingly independent of the conditions that initiate them, leading to a process called 'cumulative causation of migration' (Massey 1990).

The idea of cumulative causation captures the self-feeding dynamics of migration flows, and explains why migrant streams between regions persist even when the initial conditions that provoked migration no longer prevail. Several studies from different contexts show evidence of cumulative migration flows (Dunlevy 1991; Curran and Rivero-Fuentes 2003; Massey and Espinosa 1997).

Recent research, however, points to heterogeneity in migration outcomes across time and space not expected within the current cumulative causation framework (Curran et al. 2005; Fussell and Massey 2004; Kanaiaupuni 2000). For example, recent evidence from Thailand, based on some of the data used in this study, observes cumulative causation underway, but also shows dramatic heterogeneity in migration patterns across communities. While migration to urban areas reaches mass levels in some rural communities, it lingers at low levels in others that are very similar in economic and demographic characteristics (Garip 2008). *The heterogeneity in migration patterns presents a puzzle that cannot be understood using the current economic, social, or cumulative explanations for migration*.

In this section, we apply a novel theoretical framework, combining elements of diffusion, network externalities, and social homophily, to explain the heterogeneity in migration patterns observed in these Thai communities. First, we build on cumulative causation theory to characterize the interconnectedness of migration decisions of socially-related individuals. We expect that information or assistance provided by prior migrants in a community will reduce the costs and risks and increase the benefits of migrating for potential migrants. We translate this idea into a model of migration as a diffusion process with network externalities, where the prior adoption of migration in a population alters the probability of migration for remaining non-adopters. We argue that potential migrants will be more likely to respond to information or assistance provided by prior migrants to whom they are socially connected. We identify household, village and regional ties as three broad categories of social relations that channel diffusion of migration.

Second, we consider how social homophily moderates the impact of network externalities, and creates differential migration diffusion patterns in communities. We cannot capture all aspects of social homophily in Nang Rong, as we do not have data on social ties within and between villages. McPherson and Smith-Lovin (1987) differentiate between *induced homophily,* which results from the degree of diversity in a population, and *choice homophily,* in which homogeneity is driven by individual's choice to interact with similar others. We have information on characteristics of all the individuals in the study villages, which allows us to produce measures of homogeneity that serve as proxies for induced homophily. We cannot measure choice homophily, however. To avoid confusion and clarify the scope conditions of our argument, we will use the term *homogeneity* to refer to population distributions within the villages. Our assumption is that, if choice homophily is similar among the villages in our population (and there is no reason to believe that it is not), then homogeneity will be strongly correlated with homophily at the village level.

We test four hypotheses about the relationship between network effects, externalities, and migration. Each hypothesis is subject to *ceteris paribus* conditions and modeled with extensive controls.

The first hypothesis is based on the expectation that, as noted earlier, migration in earlier time periods reduces the cost and raises the expected rewards of current migration. If this is the case, then:

Hypothesis 1: The greater the number of prior migrants, the higher the migration rate.

In testing this hypothesis we simply replicate the results of previous research demonstrating positive network effects on migration (Massey and Espinosa 1997; Massey and Garcia-Espana 1987).

The second hypothesis addresses the effect of social homogeneity (viewed as a proxy for network homophily) on migration rates. Previous research has suggested that social homogeneity reduces the diversity of information into a population, limiting the ability of members to learn from prior experience (Garip 2008). In the case of migration, homogeneity may limit the number of migration channels through which migrants are recruited, the diversity of positions that they occupy in the city, and the diversity of the urban social ties to which their experiences gives other villagers access. Thus:

Hypothesis 2: The more homogeneous the population, the lower the rate of migration.

Note that this hypothesis is also consistent with the result of our first case, which demonstrated that high levels of social homophily reduced the extent to which Internet use diffused.

The third hypothesis addresses the way in which social homogeneity moderates the strength of network effects. We anticipate that even if (as per hypothesis 2) homogeneity reduces the rate of migration by reducing the diversity of available information, homogeneous villages will circulate whatever information prior migrants provide more efficiently than will more heterogeneous places. As Putnam (2007) has argued, social heterogeneity may reduce social cohesiveness. This, in turn, is likely to reduce the velocity and scope of information flow within the population, limiting network effects to subsets of the larger population. By contrast, homogeneous populations may experience the benefit of prior migration more quickly and completely.[22] If this is the case:

Hypothesis 3: The more homogeneous the population, the stronger the network effects (i.e., the stronger the effects of prior on current migration).

Our fourth hypothesis applies to the village rather than the individual level of analysis. In the Internet case, we found that social homophily exacerbated intergroup inequality. By analogy, we anticipate that among Thai villages, social homogeneity will increase variance in outcomes by amplifying the effects of initial differences.

Hypothesis 4: At the village population level, the over-time increase in variance in migration rates will be higher among homogeneous villages than among heterogeneous villages.

We address these hypotheses in the analyses of the Thai case that follow.

Data and Methods

[22] Alternately one might argue that homogeneity would interact *negatively* with the number of prior migrants because each migrant brings back less useful information (as per Hypothesis 2), such that relatively heterogeneous village networks are likely to connect potential migrants to prior migrants from more diverse backgrounds who could potentially provide information about a diverse set of migrant destinations or work options, thus accelerating migration. We opt for Hypothesis 3 because we expect that the positive impact of heterogeneity is likely to be captured by its direct effect, rather than by its interaction with prior migration. Nonetheless, which mechanism is strongest is likely a function of attributes of specific cases, to that, where the difference in value of information gleaned from heterogeneous as opposed to homogeneous networks exceeds the positive effects of social cohesion on the circulation of information, a negative interaction between homogeneity and the number of prior migrants might be observed.

The data for the study come from the Nang Rong Survey of 22 migrant-sending villages between 1972 and 2000, when Thailand's economy shifted from agriculture to manufacturing, propelling the migration from rural areas to Bangkok and other urban destinations.[23] Nang Rong is a relatively poor district in an historically poor region of Thailand and an important source of migrants to urban centers, primarily Bangkok.

Earlier migration flows from this area consisted of mostly seasonal migrants, who moved to find alternative livelihoods for a few months during the droughts that preceded the monsoon rains (Phongpaichit 1990). This seasonal character of migration started to change in the mid-1980s with the growth in manufacturing exports, and an increased demand for labor in Bangkok and its provinces (Bello, Cunningham, and Poh 1998). Rural migrants in their teens and early twenties, mostly from the Northeastern region of the country (including Nang Rong), started to flock to factories, construction or service jobs in urban regions, not just seasonally, but for longer durations. Internal migration reached unprecedented rates, yet mostly retained its temporary character.

In 1984, the first year of the Nang Rong Survey, and roughly the beginning of the period of dramatic growth in rural-urban flows, the 22 study villages looked very similar. Ranging from 340 to 380 kilometers in distance to Bangkok, the villages shared common economic features: Agriculture, rice cultivation specifically, was the primary economic activity. 87% of individuals in the sample listed farm work as their main occupation with little variation across villages. Water shortages and irregular rainfall were reported in 16 out of 22 villages in 1984, with possible consequences for rice cultivation. Only one

[23] Nang Rong Survey is a collaborative effort between the University of North Carolina and Mahidol University, Thailand. More information is available at: http://www.cpc.unc.edu/projects/nangrong.

village had a nearby factory, providing an alternative to agriculture work, therefore, individuals in most villages had strong incentives to migrate seasonally. None of the villages yet had phone lines, and few had electricity (7 out of 22). Adding the similarities in demographic structure, number of households ranging from 70 to 154 (median 124) in 1984, we would expect similar rural-urban migration patterns across the 22 villages in the period that followed.

To observe these patterns over time, we employ the first three waves of data (collected in 1984, 1994, and 2000) for our analyses. The 1984 data collection was a census of all households and individuals residing in 22 villages within Nang Rong.[24] It included information on individual demographic characteristics, household assets and village institutions. The 1994 survey followed all 1984 respondents still living in the original village, as well as any new village residents. The questionnaire included all 1984 items along with a retrospective life history component collecting information about education, work and migration experiences. Most importantly, the 1994 survey included a follow-up module for migrants who were in destination at the time of the village survey. Related project manuscripts show that the success at finding migrants was relatively high for this kind of follow-up, and on average, 44% of the migrants were successfully interviewed in the six months following the 1994 village survey (Rindfuss et al. 2007). Prior work evaluates how representative the data are, and reports that missing information from migrants who could not be located is not likely to bias results on migration patterns (Garip and Curran 2009). The 2000 survey built on the previous data

[24] Surveys were conducted in 51 villages, but migrants were followed up in destinations for only 22 villages.

collection efforts by following all of the 1994 respondents and adding to the database any new residents and households in the original 22 villages.

The retrospective life histories of individuals span a period from 1972 to 2000, and allow us to model the diffusion of individual migration over time. The availability of information on multiple communities that demonstrate significant variability in these diffusion patterns provides a unique opportunity to discover the social mechanisms that create heterogeneous migration outcomes.

*Measures***.** We focus on first migration moves, and define migrant status using binary indicator that equals 1 if the index person has been out of the home district (Nang Rong) for the first time in his/her life for more than two months in a year.[25] A household is defined as a group of individuals who resided in the same house in 1984 and a village is defined by official boundaries in 1984. Changes in the household composition in 1994 and 2000 data collection waves are reflected in the definition of household in (and following) those years.

We operationalize homogeneity at the village level using two socioeconomic dimensions: education and occupation. These attributes are likely to shape individuals' migration opportunities, and the nature of information they can provide to their household or community members once they migrate.[26] Homogeneity is measured using

---

[25]A person who has been away for less than two months in a given year is not considered a migrant. This definition is from the Nang Rong life history survey, and captures labor migrants, who make trips of long duration to destination, rather than seasonal migrants.

[26] Education and occupation, when coded an ordinal scale (increasing in degree for the former, and in average earnings for the latter), are positively associated (polychoric correlation coefficient=0.31). Polychoric correlation uses the qualitative knowledge on the ordering of categories, and provides a more accurate measure of association in the case of ordinal variables compared with the Pearson's correlation. See Kolenikov and Angeles (2009) for details.

the complement of two diversity measures: the logarithm of variance for the ordinal attribute of education, and Shannon's entropy for the categorical attribute of occupation.[27] There are six major categories of occupation: student, farm work, factory, construction, service, and other. In raw form, these diversity measures capture heterogeneity, rather than homogeneity, within villages. For our purposes, both measures are normalized to [0,1] range, and homogeneity is computed as 1-diversity.[28]

To capture the effects of network externalities, we define households, villages and Nang Rong as a region as possible diffusion channels for migration. Then, for each individual, we count the number of prior migrants in his or her household (excluding the index individual) and village (excluding the index individual's household) up through the previous year. We also compute a global measure, which sums up prior migrants in all villages in Nang Rong (excluding the index individual's village). By comparing the effect of household and village migration rates on the one hand, and global migration rates on the other hand, we seek to assess the relative importance of specific and general network externalities for inducing migration flows.

Finally, to rule out alternative explanations, we include a number of exogenous factors that capture some of the social and economic incentives for migrating. For each

---

[27] Using Shannon's entropy index, occupation diversity is quantified as follows:

$$Diversity = \frac{-\sum_{i=1}^{n} p_i \times \log(p_i)}{\log(n)}$$

where n is the number of possible occupations and p is the proportion of individuals in occupation category i. Minimum diversity occurs when all individuals are concentrated in one occupation and the index equals zero. Maximum diversity occurs when each occupation category contains the same proportion of individuals, yielding an index of 1.

[28] Results are robust to alternative measures of diversity. Using an index of qualitative variation (Lieberson, 1969) instead of Shannon's entropy for the categorical occupation measure does not affect the results. Similarly, employing a Gini coefficient instead of the log variance for the continuous education variable leaves the results unchanged.

individual, we include indicators of age, sex, education and marital status. At the household level, we measure total land owned by household members as a proxy for household's wealth and to represent a condition that influences the migration propensity of each member. Also included as a household demographic indicator is the dependency ratio (proportion of those older than 64 or younger than 15 to those who are neither). Similarly at the village level, we include indicators of schools, newspaper reading centers, factories, and rice mills, available electricity and village population to roughly represent levels of economic opportunities and development that can cause migration propensities of village members to be correlated. Table A1 in the Appendix provides descriptive statistics for the individual, household, and village level variables.

Analytic Approach: *Individual-level Analysis.* We use event-history methods to model the individual-level diffusion of migration, as these methods allow us to take into account the timing of migration and censoring in data (i.e., not every individual migrates before the end of the study period) (Marsden and Podolny 1990; Strang 1991). An individual's adoption of migration is modeled as a process occurring over time, where the number of prior adoptions in household, village, or Nang Rong are time-varying factors affecting the hazard of adoption. We use the proportional hazards model,

$$(15)\ \log h_i(t) = \alpha + \beta x_i(t) + \gamma z_i$$

in which the dependent variable is $h_i(t)$, the hazard of first migration at time t for individual i, α is a constant, $x_i(t)$ is a vector of time-varying covariates characterizing individual i, and z_i is a vector of time-constant covariates for individual i . β is a vector of parameters giving effects of time-varying covariates on the hazard rate, and γ is a vector of parameters giving the effects of time-constant variables on the hazard rate.

In applying this model to the diffusion of migration, the key component is the vector $x_i(t)$, which includes the number of adopters in an individual's household and village (excluding the household), as well as the region of Nang Rong (excluding the village). Using each indicator, we can test our first hypothesis suggesting effects of prior on current migration. By comparing the magnitude and significance of the corresponding β parameters for these indicators, we can also evaluate the relative importance of different foci for social relations on migration (Feld 1981; Strang 1991). For instance, if the household tie is the sole relation structuring the diffusion of migration, then adoption rates for an individual will be a function of the number of prior adopters in the household. If village ties are important influences on diffusion, an individual's adoption of migration will also be a function of the number of villagers who have already migrated.

Another key variable of interest is the level of homogeneity in the village (in terms of education or occupation), which varies over time, and is also part of $x_i(t)$. Our second and third hypotheses respectively predict lower migration rates and higher network effects with higher homogeneity in the village. To test these ideas, we add measures of occupation and education homogeneity in the village, as well as interaction terms between these measures and the number of prior migrants in the village.[29] We also include several time-varying (age, marital status, education, household land, dependency ratio, village population, factory, school, newspaper reading center in village, electrific-

[29] We focus on village-level indicators because we are interested in understanding the variation in migration patterns *between* villages. Interactions at the household level, between household-level homogeneity and prior migrants in the household, could be introduced to model the variation between households in a village, which is out of the scope of this analysis.

ation of village) and time-constant (sex) control variables (as part of $x_i(t)$ and z_i respectively) to our model.

Analytic Approach: Village-level Analysis. To evaluate our fourth hypothesis suggesting higher heterogeneity of migration patterns among more homogenous villages, we use descriptive analysis and graphically compare the variance in migration patterns among villages categorized as high or low in education and occupation homogeneity.

Migration as a Diffusion Process with Specific Network Externalities

We introduced Thai internal migration primarily as a case to illustrate the uses of our theoretical framework, which integrates elements of network externalities and social homophily. We expect this approach to help explain the heterogeneity in migration flows observed in the 22 study villages. We seek to achieve both goals by (1) modeling individuals' first migration as a diffusion process, where prior adoption rates among social contacts influence future adoption, and where homophily in social ties may be induced by population homogeneity; and (2) drawing the village-level implications of individuals' migration behavior under these conditions over time.

We begin by observing the cumulative number of migration events between 1972 and 2000 in Nang Rong displayed in Figure 10. The cumulative number of adopters (i.e. migrants) approaches an S-shaped curve over time (Fig. 10), while the frequency of the new adopters per year approximates a normal, bell-shaped curve (Fig. 11). Early in the diffusion process, relatively few individuals adopt in each time period. Gradually, the rate of adoption speeds up until a majority of the individuals in the region have migrated.

Figure 10: Diffusion of Migration in Nang Rong

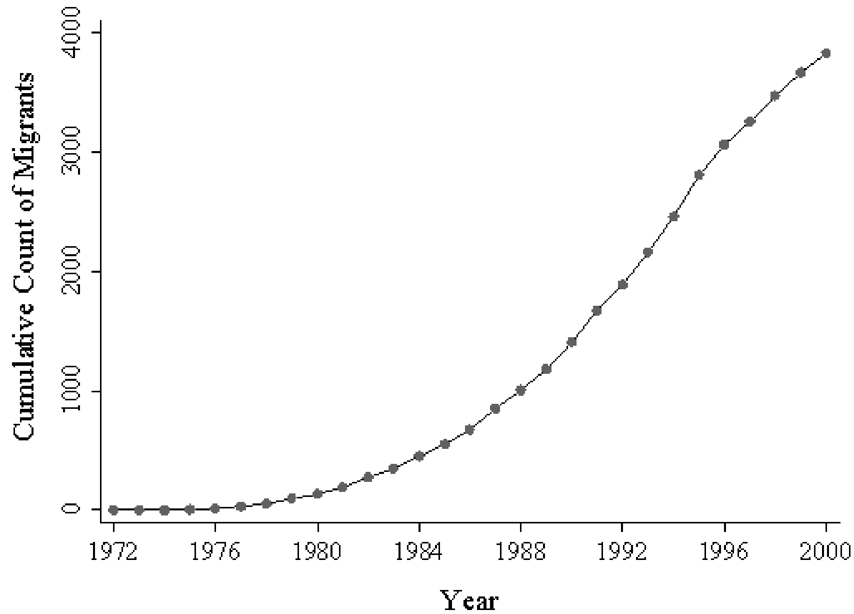


Figure 11 The Number of New Migrants Each Year

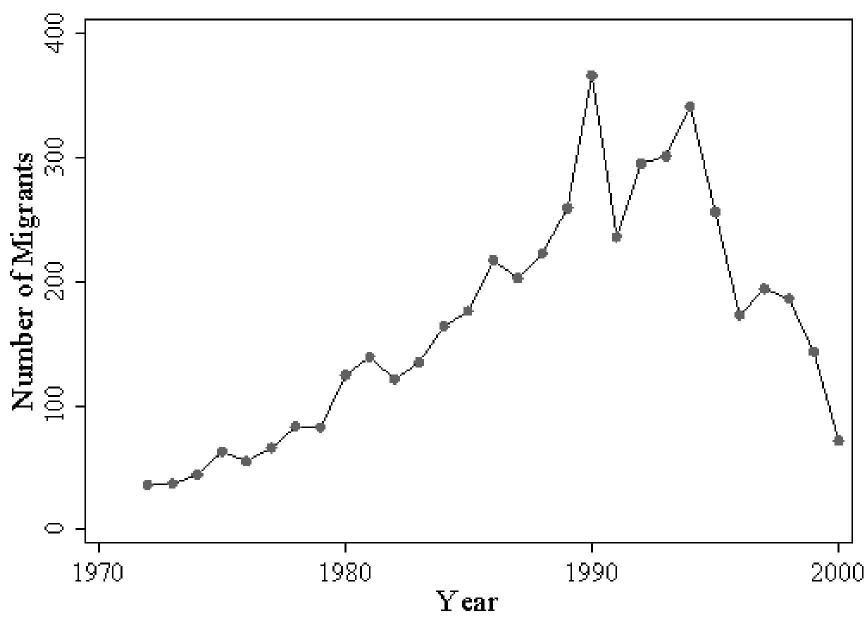
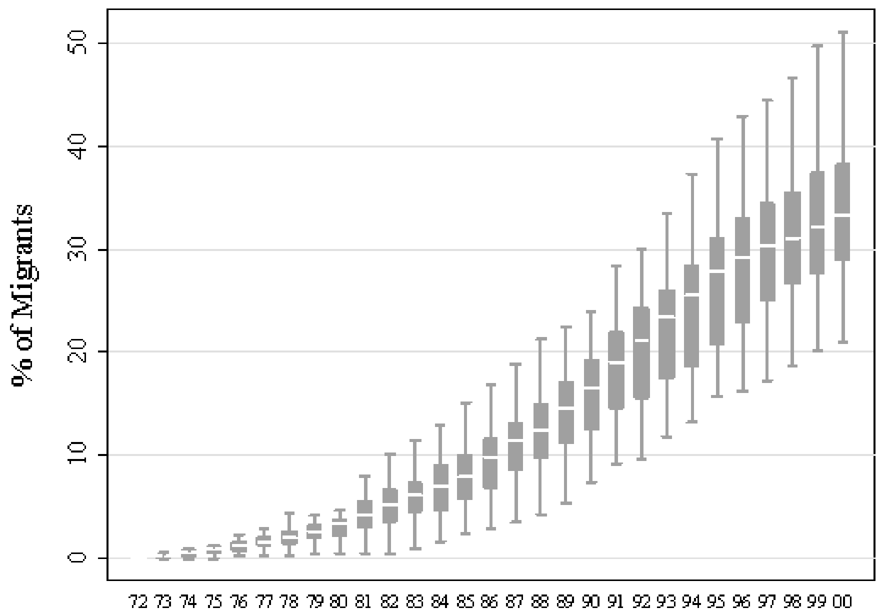


Figure 12 Dispersion of Cumulative % of Migrants across Villages over Time



The uniform path displayed in Figure 10 conceals the variation in migration paths across villages. This variation becomes apparent in the box plot in Figure 12, which demonstrates the distribution of cumulative migrants as a percent of village population across villages for each year between 1972 and 2000. Not only is there variation in the cumulative percentage of migrants across villages in each year, but this variation keeps growing over time. (Similar patterns are observed for annual, rather than cumulative, percentage of migrants.) The observed pattern actually implies a persistent and increasing between-village inequality in migration rates. As a result of this inequality, prior work suggests, some villages in Nang Rong reach very high levels of migration and become integrated into the urban economy, while others linger at low levels of migration and remain isolated (Garip 2008).

If diffusion of migration is a population-level process operating within Nang Rong, we expect similar adoption rates across villages, especially given how comparable these villages are in terms of economic and demographic structure, as noted above. The significant variation in the migration rates of the 22 villages suggests that there are specific diffusion channels within villages, which we consider in our analysis.

We investigate three social relations as possible channels of diffusion to test our hypothesis linking past to current migration. The first is membership in the same household. Because migration is a risky undertaking, individuals are more likely to rely on family ties to obtain trustworthy information about migration opportunities (Massey and Espinosa 1997). A second plausible channel of diffusion is the village level. There are a number of reasons why migration should diffuse through village networks more quickly than between them. First, individuals are more likely to know migrants in their own village than outside their village. Second, reciprocity obligations are likely to operate more within than outside the village, and individuals are more likely to receive information or help from their fellow villagers with the expectation of reciprocating the favor in the future (Portes and Landolt 2000). A final channel of diffusion is Nang Rong as a region. While the first two channels of diffusion represent specific networks, where the identities of alters are likely to be known and significant to the individual, the third channel reflects a general network, where alters are likely to be anonymous. Specific network externalities accrue through information or help household or village members provide to potential migrants, which reduces the costs and risks of migrating. General network externalities, on the other hand, may work through the institutions that develop in the region over time to support migration flows. For instance, high migration flows from the Nang Rong region could prompt employers in urban areas to send recruiters, fostering more migration out of each village in the long run. This is a common practice one of the authors observed during fieldwork in the region in 2005.

Table 1. Proportional Hazards Model of Individual Migration in 22 Nang Rong Villages (1973-2000)

		Diffusion Channel		
Variable	(0) Baseline	(1) Household	(2) Village	(3) Nang Rong
Age	0.941 **	0.940 **	0.935 **	0.932 **
	(0.011)	(0.011)	(0.012)	(0.012)
Sex (Male=1)	1.132 **	1.143 **	1.134 **	1.133 **
	(0.043)	(0.043)	(0.043)	(0.043)
Years of education	1.098 **	1.097 **	1.091 **	1.085 **
	(0.009)	(0.009)	(0.009)	(0.009)
Married	0.619 **	0.623 **	0.618 **	0.618 **
	(0.038)	(0.038)	(0.038)	(0.038)
Household land (rai)	0.997 **	0.997 **	0.997 **	0.997 **
	(0.001)	(0.001)	(0.001)	(0.001)
Household dependency ratio	1.242 **	1.262 **	1.232 **	1.221 **
	(0.035)	(0.036)	(0.035)	(0.035)
Village population	1.000 **	1.000 **	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)
Factory w/in 5km of village?	0.976	0.986	1.011	0.999
	(0.049)	(0.050)	(0.051)	(0.050)
Number of rice mills	1.006	1.007	1.008	1.007
	(0.012)	(0.012)	(0.012)	(0.012)
School in village?	0.992	1.004	0.977	0.982
	(0.043)	(0.044)	(0.043)	(0.043)
Newspaper reading center in village?	1.202 **	1.183 **	1.123 **	1.129 **
	(0.049)	(0.048)	(0.047)	(0.047)
Years since village electrified	1.011 **	1.005	0.986 **	0.977 **
	(0.005)	(0.005)	(0.006)	(0.007)
No of prior migrants in household		1.105 **	1.083 **	1.077 **
		(0.015)	(0.015)	(0.015)
No of prior migrants in village (excl. hh)			1.002 **	1.001 **
			(0.000)	(0.001)
No of prior migrants in Nang Rong (excl. village)				1.000 **
				(0.000)
N (person-years at risk)	50,198	50,198	50,198	50,198
Likelihood ratio χ^2 vs. prior model		50.15 **	35.31 **	8.52 **

*p<0.05; **p<0.01. Std errors given in parentheses. Results presented in hazard ratios.

Table 1 displays results for four individual-level models of migration diffusion between 1973 and 2000. The baseline model includes a number of exogenous variables that may influence an individual's hazard of migrating. The estimates show that the rate of migration decreases with age, increases with education, is higher for men than women, and for unmarried than married individuals. Rate of first migration decreases with the land owned by an individual's household, and increases with the dependency ratio (proportion of those older than 64 or younger than 15 to those who are neither) in the household. Among village characteristics, village population, presence of a newspaper reading center (where village members get together and potentially share information about migration experiences) and the availability of electricity (which proxies the level of economic development) all increase the individual rates of migration.

Each of the added covariates in Models 1, 2 and 3 measures the number of prior migrants within the individual's particular social relation. Model 1 presents the results for diffusion through the household channel. Model 2 and 3 additionally introduce prior migrants in the village and Nang Rong (excluding the individual's village) as alternative channels. The lower panel of Table 2 provides likelihood ratio tests of adding covariates, which in each case improve the model fit significantly. If we focus on Model 3, which includes the three covariates representing alternative diffusion channels, we see that the rate of migration significantly increases with the number of migration events in the household, village, and Nang Rong. The rate of migration increases most steeply as a function of the number of migrants in the household, followed by the number of migrants in the village, and then only slightly by the number of migrants in Nang Rong.[30] This result provides strong support for our first hypothesis, and suggests the importance of specific, rather than general, network externalities in the case of rural-urban migration,

[30] Standardizing the variables to mean-deviation form to account for their different ranges does not alter the results.

where the effect of others' behavior on an individual's migration increases with their social proximity to the individual.

*Social Homogeneity, Network Externalities and Migration***.** In Table 2, we test our second hypothesis, which links higher village homogeneity to lower migration rates. We introduce two social homogeneity indicators that measure the similarity of individuals in a village by education and occupation, respectively. These indicators are highly correlated at the village level (Pearson's correlation = 0.74), and therefore not included in the same model.

Model 1 includes a village-level indicator of education homogeneity standardized to [0,1] range, so that villages with the lowest and highest homogeneity obtain the values of 0 and 1, respectively. As per our hypothesis, we expect high education homogeneity to limit the diversity of information available to potential migrants, and to lower the migration rates in a village. The estimate supports this expectation, and shows that individuals living in the village with the highest level of homogeneity (=1) have a 61 percent (0.39-1=-0.61) lower rate of first migration than individuals living in the village with the lowest homogeneity (=0). Note that the model also controls for the mean level of education in the village, which does not have a significant effect on migration. The effect of occupation homogeneity (also standardized to [0,1] range), estimated in Model 3, is even stronger as individuals in the village with the highest occupation homogeneity have a 69 percent (0.31-1=-0.69) lower rate of migrating compared with their counterparts in the villages with the lowest homogeneity.

Table 2. Proportional Hazards Model of Individual Migration in 22 Nang Rong Villages (1973-2000[a])

	Education Homogeneity		Occupation Homogeneity	
Variable	(1)	(2)	(3)	(4)
No of prior migrants in household	1.076 **	1.061 **	1.063 **	1.048 **
	(0.015)	(0.015)	(0.015)	(0.015)
No of prior migrants in village (excl. hh)	1.001 *	0.995 **	1.002 **	0.993 **
	(0.001)	(0.001)	(0.001)	(0.001)
No of prior migrants in Nang Rong (excl. village)	1.000 **	1.000	1.000 **	1.000
	(0.000)	(0.000)	(0.000)	(0.000)
Mean education in village	0.976	1.045		
	(0.063)	(0.070)		
% working in a farm in village			0.384 *	0.305 **
			(0.149)	(0.123)
% working in a factory in village			0.644	3.449
			(0.468)	(2.551)
% working in construction in village			2.039	9.691
			(2.785)	(13.474)
% working in service in village			0.000 **	0.855
			(0.001)	(1.711)
Homogeneity in year [0,1]	0.391 **	0.030 **	0.309 *	0.126 **
	(0.165)	(0.016)	(0.177)	(0.076)
Homogeneity * No of prior migrants in village		1.037 **		1.025 **
		(0.004)		(0.003)
N (person-years at risk)	50,198	50,198	50,198	50,198
Likelihood ratio χ^2 vs. model w/o interaction		102.79 **		96.18 **

*p<0.05; **p<0.01. Std errors given in parentheses. Results presented in hazard ratios.

[a] Controls for age, sex, years of education, marital status, houshold land and dependency ratio; population, factories, rice mills, schools, newspaper reading center and electrification in the village are included.

Our third hypothesis suggests stronger network effects in highly homogeneous villages due to increased cohesiveness and velocity of information flows between past and current migrants. This idea is tested by interacting the number of prior village migrants with education homogeneity in Model 2 and with occupation homogeneity in Model 4. Likelihood ratio tests displayed in the lower panel show that the fit is improved significantly in both cases compared with models without the interaction term. The positive coefficient estimates for the interaction terms, in both Models 2 and 4, provide strong support for the hypothesis. In the village with the lowest education homogeneity, each prior migrant in the village *reduces* the rate of migration by 0.5 percent (0.995-1=--0.005. By contrast, in the village with the highest education homogeneity, each prior migrant in the village *increases* the rate of migration by 3 percent (0.995*1.037-1=0.03).

A similar pattern in observed for occupation homogeneity in Model 4. In the lowest occupation homogeneity village, the effect of network externalities is reduced, and each prior migrant in the village decreases the rate of migration by 0.7 percent (0.993-1=-0.007). In the highest homogeneity village, by contrast, each prior migrant increases the rate of migration by 2 percent (0.993*1.025-1=0.018). These results suggest that whereas homogeneity (in education or occupation) has a negative direct effect on the rate of individuals' first migration (consistent with Hypothesis 2), it has a positive indirect effect through network externalities in the village (as posited by Hypothesis 3).

Explaining the Heterogeneity in Migration Patterns across Villages

So far we attempted to demonstrate how migration diffuses through household, village, and global regional channels, and how the impact of village channels is moderated by social homogeneity. We now turn to Hypothesis 4: Can we employ the ideas of network externalities and social homophily to explain the puzzling heterogeneity in migration patterns across villages?

The individual-level models presented above suggested that social homogeneity in the village decreases the rates of migration, but increases the impact of number of prior migrants on the rates of migration. By amplifying the effects of prior migration, homogeneity (and, by implication, homophily) increases inequality among villages over time, boosting the effects of prior migration for villages with many migrants and penalizing villages with relatively few. By contrast, where heterogeneity is high, migration will be driven primarily by individual-level factors and household experience, generating less divergent outcomes between villages. In other words, because social homogeneity amplifies the effect of network externalities, then, we should observe a wider dispersion of migration patterns among villages classified as high homogeneity than in villages with low homogeneity.

We begin with a descriptive analysis of how migration patterns across villages vary by levels of social homogeneity. We group villages as high versus low homogeneity (with 11 villages in each category) along two dimensions of education and occupation, and observe the distribution of cumulative number of migrants across villages over time. Figure 13 displays the box plot of cumulative number of migrants across high and low education homogeneity villages over four time periods, 1985, 1990, 1995 and 2000. Figure 14 repeats the same analysis for occupation homogeneity. In both cases, we see that the distribution of cumulative migration is similar across high and low homogeneity villages in 1985 and 1990, the early initiation period of migration flows. As migration flows mature and gain prevalence from 1990 on, we see a wider spread of cumulative migration across high homogeneity communities leading to high levels of between-village inequality. By contrast, the distribution of cumulative migration has a much narrower spread in low homogeneity villages implying low levels of between-village inequality. These results provide descriptive evidence to our fourth hypothesis that social homogeneity, through its impact on network externalities, exacerbates inequality between villages in migration patterns over time.

Figure 13: Dispersion of Cumulative Migration Across Villages by Education Homogeneity

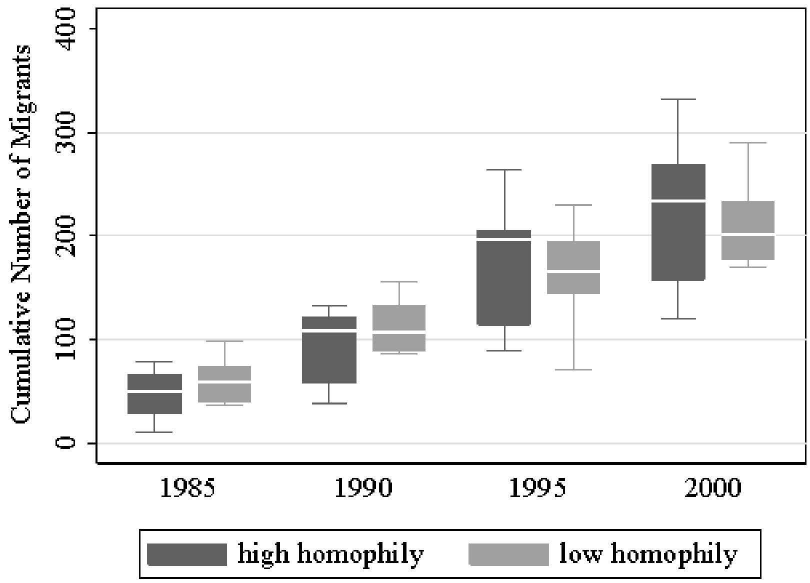
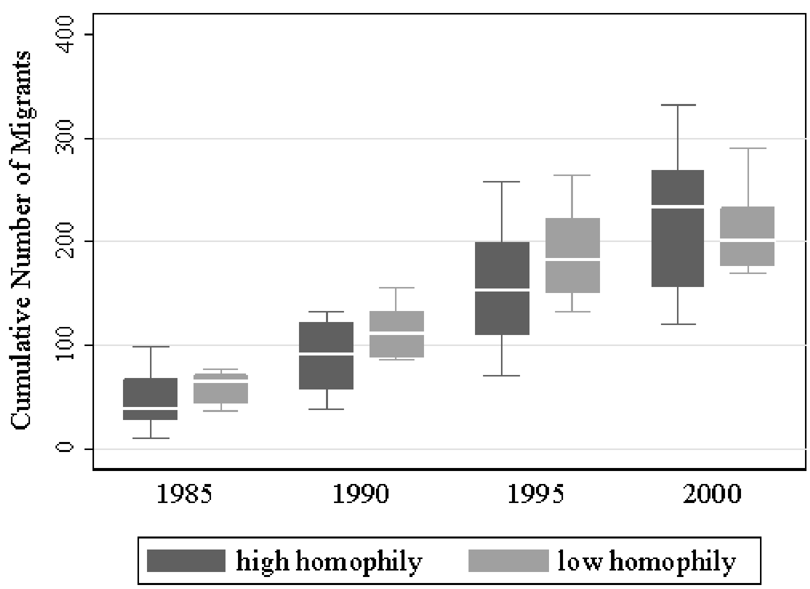


Figure 14: Dispersion of Cumulative Migration Across Villages by Occupation Homogeneity



Conclusion

Certain kinds of social structures generate more inequality than others. Sociologists have long appreciated the tendency towards cumulative advantage (what Robert K. Merton referred to as the “Matthew Principle”); and the tendency of inequality to reproduce itself intergenerationally (Bourdieu 1977). We have argued that inequality also comes out of a combination of particular forms of interdependence, those that make courses of action more profitable or productive for actors whose friends and associates also pursue them; and particular kinds of networks, those characterized by status homophily. Inequality, we contend, emerges out of the interaction of these two features of social structure: positive interdependence (network externalities) and homophily.

We have presented two studies. The first was a computational model of changing intergroup inequality in access to the Internet over the course of that technology’s diffusion. We varied the model to explore the impact of types of externalities and degrees of status homophily on the rate at which different groups went online and on the degree of inequality that resulted. The results demonstrated that, other things equal, diffusion processes with identity-specific externalities and homophily produce more rapid adoption and greater inequality than processes marked by general externalities or by specific externalities in the absence of homophily; and the stronger the bias towards homophily, the greater the degree of inequality. Substantively, these analyses provide insight into the tenacity of the digital divide.

Whereas the first study focused on explicating the mechanism and exploring its logic, the second used the framework to address an empirical puzzle --- the surprising heterogeneity of rural-urban migration patterns in ostensibly similar Thai villages.

Using retrospective and contemporary data collected over a period of nearly thirty years, we tracked the rise (and recent decline) of migration rates in twenty-two neighboring villages. Consistent with the notion that inequality emerges out of diffusion processes characterized by identity-specific network externalities and social homophily – and, specifically, that potential migrants face less risk and better prospects if they can benefit from the counsel and connections of friends and kin who have gone before them – we found that (a) migration rates increased as a function of prior migration in one's household, village, and region, with the strongest effects coming from the most local ties; (b) educational and occupational homogeneity, which we argue are ordinarily associated with higher levels of structurally determined homophily, are associated with lower migration rates at the individual and village levels; (c) despite its negative direct effect, homophily amplifies network effects on migration; and (d) because of this tendency for homophily to amplify initial inequalities (a result observed as well in the Internet-diffusion model), villages characterized by homophily display systematically higher variance in migration rates over time than those that are more heterogeneous.

Many questions remain about the scope of the mechanism we have identified. Not all practices have positive externalities --- most of the goods and services we consume are rival, with one person's enjoyment reducing rather than enhancing another's; and many behaviors are competitive (my college degree is worth less if everyone else has one). Moreover, not all goods and services with network externalities have *identity-specific* externalities: We all want enough people to like Japanese food so that our favorite Japanese restaurant can keep its doors open, or enough people to contribute to our favorite charity that it can continue to pursue causes we believe in, or enough people use eBay that we can find what we want on it rapidly and at a reasonable price – but we don't particularly care who those people are. One challenge in developing the ideas in this paper further is to figure out how to estimate change over time in the contribution of practices characterized by interdependent choices of members of networks characterized by status homophily to overall inequality, and to understand when and where their effects are likely to be felt most strongly.

A related question concerns those situations in which the mechanism we describe may actually *undermine* inequality. At the onset of the Internet, as people observed the advantage of youth in mastering the new technology, some suspected that the Internet would open the doors of opportunity to young people from all walks of life. That never happened. But there probably *are* status-conferring or otherwise beneficial activities in which subaltern groups possess an initial advantage that is amplified by processes of homophily and network-specific externalities. The advantages of African-American youth with respect to hip-hop and basketball, of young French Canadian men with respect to ice hockey, or of Italian-Americans in popular music in the 1950s and 1960s were hardly sufficient to pose significant challenges to entrenched privilege. Yet they demonstrate that the mechanism we describe *can*, under some circumstances, work to the advantage of the non-elites.

At the same time, the less privileged may also suffer disproportionately from the amplification of negative effects of adopting practices that are more harmful if others in one's network practice them. The risks of smoking cigarettes, for example, are greater if your network alters reinforce your behavior, discourage quitting, and teach you to value the practice. The cost of taking drugs may be greater if you are surrounded by other drug users, who may introduce riskier behaviors or bring the novice user into contact with the criminal-justice system. Similarly, the risk of unprotected sex is much greater if one is in a social milieu in which no one uses condoms than if condom use is widespread (Tavory and Swidler 2009). In so far as networks are characterized by status homophily, such cases illustrate what may be a powerful source of cumulative disadvantage. Note that if externalities were entirely negative – if their only effect were to raise the cost or reduce the benefit of a behavior – then the negative externalities would *discourage* adoption of risky behaviors. What these cases have in common is that homophilic networks may offer social rewards (affiliation, identity, status, participation in shared rituals) that are highly salient in the short run, whereas the risks are longer term and easier to discount. Thus an understanding of this mechanism may contribute to a theory of counternormative subcultures.

There is still much we do not understand about the ways that networks produce cumulative advantage by enhancing the value of new practices to their adopters. Our models suggest that homophily produces more or less linear increases in inequality in the presence of positive network-specific externalities. Yet the process may be dependent upon the number and distribution of bridges --- actors who, because they possess unusual combinations of identities, link networks of different kinds and can introduce practices to actors who might otherwise not encounter it – in ways that are poorly understood (by analogy, for example, to small-world networks [Watts 2003]). Other details of the mechanism – for example, the functional form of the relationship between number of adopters in one's network and the probability that one will adopt – are also likely to be consequential in less than obvious ways. If this paper convinces the reader that these questions are worth pursuing, it will have accomplished its purpose.

References

Abbate, Janet. 1999. *Inventing the Internet*. Cambridge: MIT Press.

Abrahamson, Eric and Lori Rosenkopf. 1993. "Institutional and Competitive Bandwagons: Using Mathematical Modeling as a Tool to Explore Innovation Diffusion." *Academy of Management Review* 18: 487-517.

Abrahamson, Eric and Lori Rosenkopf. 1997. "Social Network Effects on the Extent of Innovation Diffusion: A Computer Simulation." *Organization Science* 8: 289-309.

Arthur, W. Brian. 1989. "Competing Technologies, Increasing Returns, and Lock-In by Historical Events." *The Economic Journal* 99:116-31.

Baron, James N. and William T. Bielby. 1980. "Bringing the Firms Back in: Stratification, Segmentation, and the Organization." *American Sociological Review* 45: 737-765.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades." *Journal of Political Economy* 100: 992-1026.

Bello, W., S. Cunningham, and L.K. Poh. 1998. *A Siamese Tragedy: Development and Disintegration in Modern Thailand*. London: Zed Books, Ltd.

Blau, Peter M. *Inequality and Heterogeneity*. 1977. New York: Free Press.

Bonfadelli, Heinz. 2002. "The Internet and Knowledge Gaps: A Theoretical and Empirical Investigation." *European Journal of Communication* 17: 65-84.

Booth, Alan, John Edwards and David Johnson. 1991. "Social Integration and Divorce." *Social Forces* 70: 207-24.

Boudon, Raymond. 1973. *Education, Opportunity and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.

Breen, Richard and John H. Goldthorpe. 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society* 9: 275-305.

Brown, Richard S. 2005. "Changes in Advanced Placement Test Taking in California High Schools 1998-2003." University of California All Campus Consortium on Research for Diversity, Working Paper RR-002-0105. eScholarship Repository. http://repositories.cdlib.org/ucaccord/papers/RR-002-0105/ Last accessed, August 10, 2009.

Bruch, Elizabeth E. and Robert Mare. "Neighborhood Choice and Neighborhood Change." *American Journal of Sociology* 112: 667-709.

Buhai, Sebastian and Marco van der Leij. 2008. "A Social Network Analysis of Occupational Segregation." Manuscript, Tinbergen Institute, Erasmus University, Rotterdam.

Buhler, Christoph and Ewa Fratczak. 2007. "Learning From Others and Receiving Support: The Impact of Personal Networks on Fertility Intentions in Poland." *European Societies* 9: 359-82.

Burt, Ron. 1987. "Social Contagion and Innovation: Cohesion versus Structural Equivalence." *American Journal of Sociology* 92: 1287-1335.

Burt, Ron. 2000. "The Network Structure of Social Capital." *Research in Organizational Behavior*, ed. Robert Sutton and Barry Staw. Greenwich: JAI Press.

Cairncross, Frances. 1997. *The Death of Distance: How the Communications Revolution Will Change Our Lives*. Boston: Harvard Business School Press.

Castells, Manuel. 2001. *Internet Galaxy: Reflections on the Internet, Business and Society*. New York: Oxford University Press.

Church, Jeffrey and Ian King. 1983. "Bilingualism and Network Externalities." *Canadian Journal of Economics* 26: 337-45.

Coleman, James S., Elihu Katz and Herbert Menzel. 1957. "The Diffusion of an Innovation Among Physicians." *Sociometry* 20: 253-70.

Curran, Sara, Filiz Garip, Chang Chung, and Kanchana Tangchonlatip. 2005. *Social Forces* 84: 225-55.

Curran, Sara and E. Rivero-Fuentes. 2003. "Engendering Migrant Networks: The Case of Mexican Migration." *Demography* 40(2): 289-307.

DiMaggio, Paul and Bart Bonikowski. 2008. "Make Money Surfing the Web: The Impact of Internet Use on the Earnings of U.S. Workers." *American Sociological Review* 73: 227-52.

DiMaggio, Paul and Coral Celeste. 2004. "Technological Careers: Adoption, Deepening, and Dropping Out in a Panel of Internet Users." Russell Sage Foundation Working Paper Series. Available on-line at http://www.russellsage.org/publications/workingpapers/. Last accessed March 24, 2008.

DiMaggio, Paul and Joseph Cohen. 2004. "Information Inequality and Network Externalities: A Comparative Study of the Diffusion of Television and the Internet." In

The Economic Sociology of Capitalism, edited by Victor Nee and Richard Swedberg. Princeton: Princeton University Press.

DiMaggio, Paul, Eszter Hargittai, Coral Celeste, and Steven Shafer. 2004. "From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality." In *Social Inequality*, edited by Kathryn Neckerman. New York: Russell Sage Foundation.

Dunlevy, J. A. 1991. "On the settlement patterns of recent Caribbean and Latin immigrants to the United States." *Growth Change* 22(1): 54-67.

Eastin, Matthew S. and Rober LaRose. 2000. "Internet Self-Efficacy and the Psychology of the Digital Divide." *Journal of Computer-Mediated Communications* 6. Retrieved September 6, 2006

Economides, Nicholas and Charles Himmelberg. "Critical Mass and Nework Evolution in Telecommunications." Pp. 47-64 in *Toward a Competitive Telecommunications Industry: Selected Papers from the 1994 Telecommunications Public Research Conference,* ed. Gerald W. Brock. Mahwah, NJ: Lawrence Erlbaum Associates.

Fallows, Deborah, Lee Rainie and Graham Mudd. 2004. "The Popularity and Importance of Search Engines." Washington: Pew Internet and American Life Project. http://www.pewinternet.org/pdfs/PIP_Data_Memo_Searchengines.pdf Last accessed March 25, 2008.

Featherman, David and Robert Hauser. 1978. *Opportunity and Change*. New York: Academic Press.

Feld, Scott. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86: 1015-1035.

Fischer. Claude. S. 1992. *America Calling: A Social History of the Telephone to 1940.* Berkeley: University of California Press.

Fox, Susannah. 2004. *Older Americans and the Internet.* Washington: Pew Internet and American life Project. http://www.pewinternet.org/pdfs/PIP_Seniors_Online_2004.pdf Last accessed, March 25, 2008.

Fussell, E. and Douglas Massey. 2004. "The Limits to Cumulative Causation: International Migration from Mexican Urban Areas." *Demography* 41(1):151-171.

Garip, Filiz. 2008. "Social Capital and Migration: How Do Similar Resources Lead to Divergent Outcomes?" *Demography* 45(3): 591-617.

Garip, Filiz and Sara Curran. 2009. "Increasing Migration, Diverging Communities: Changing Character of Migrant Streams in Rural Thailand." Forthcoming in *Population Research and Policy Review*.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-1380.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83: 1420-43.

Granovetter, Mark and Roland Soong. 1988. "Threshold Models of Diversity: Chinese Restaurants, Residential Segregation, and the Spiral of Silence." *Sociological Methodology* 18: 69-104.

Gustaffson, Bjorn and Mats Johansson. 1999. "In Search of Smoking Guns: What Makes Income Inequality Vary over Time in Different Countries." *American Sociological Review* 64: 585-605.

Hampton, Keith and Barry Wellman. 2003. "Neighboring in Netville: How the Internet Supports Community and Social Capital in a Wired Suburb." *City and Community* 2: 277-311.

Hargittai, Eszter. 2003. "Informed Web Surfing: The Social Context of User Sophistication." Pp. 257–74 in *Society Online: The Internet in Context,* edited by Philip Howard and S. Jones. Beverly Hills, CA: Sage Publications.

Hargittai, Eszter and Amanda Hinnant. 2008. "Digital Inequality: Differences in Young Adults' Use of the Internet." *Communication Research* 2008: 602-21.

Hedstrom, Peter. 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. New York: Cambridge University Press.

Horrigan, John B. and Aaron Smith. 2007. "Home Broadband Adoption." Washington, D.C.: Pew Internet & American Life Project. http://www.pewinternet.org/pdfs/PIP_Broadband%202007.pdf Last accessed March 26, 2008.

Jaeger, Mads Meier. 2007. "Economic and Social Returns to Educational Choices: Extending the Utility Function." *Rationality and Society* 19: 451-83.

Kalev, Alexandra, Frank Dobbin and Erin Kelly. 2006. "Best Practices or Best Guesses?" Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies." *American Sociological Review* 71: 589-617.

Kandel, Denise. 1978. "Homophily, Selection and Socialization in Adolescent Friendships." *American Journal of Sociology* 84: 427-36.

Kao, Grace and Jennifer S. Thompson. 2003. "Racial and Ethnic Stratification in Educational Achievement and Attainment." *Annual Review of Sociology* 29: 417-42.

Karabel, Jerome. 2005. *The Chosen: The Hidden History of Exclusion and Admission at Harvard, Yale and Princeton.* Boston: Houghton Mifflin.

Katz, Michael and Carl Shapiro. 1985. “Network Externalities, Competition and Compatibility.” *American Economic Review* 75: 424-40.

Klopfenstein, Kristin. 2004. “Advanced Placement: Do Minorities Have Equal Opportunity.” *Economics of Education Review* 23: 115-31.

Kolenikov, Stanislav and Gustavo Angeles. 2009 “Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer?” *Review of Income and Wealth* 55(1): 128-165.

Kraut, Robert, Tridas Mukhopadhyay, Janusz Szczypula, Sara Kiesler and Bill Scherlis. 1999. “Information and Communication : Alternative Uses of the Internet in Households.” *Information Systems Research* 10: 287-303.

Lazarsfeld, Paul and Robert K. Merton. 1954. “Friendship as Social Process: A Substantive and Methodological Analysis.” Pp. 18-66 in *Freedom and Control in Modern Society*, ed. Monroe Berger, Theodore Abel, and Charles Page. New York: Van Nostrand.

Leibenstein, Harvey. 1950. “Bandwagon, Snob and Veblen Effects in the Theory of Consumption.” *Quarterly Journal of Economics* 64: 183-207.

Leigh, Andrew and Robert D. Atkinson. 2001. *Clear Thinking on the Digital Divide.* PPI Policy Report (June 26). Washington: Progressive Policy Institute. http://www.ndol.org/documents/digital_divide.pdf . (Last accessed March 23, 2008.)

Levitt, Peggy. 1998. “Social Remittances: Migration Driven Local-Level Forms of Cultural Diffusion.” *International Migration Review* 32(4): 926-948.

Lieberson, S. (1969). “Measuring Population Diversity.” *American Sociological Review* 34(6): 850-862.

Lin, Nan. 1999. “Social Networks and Status Attainment.” *Annual Review of Sociology* 25: 467-87.

Loges, William E. and Joo-Young Young. 2001. “Exploring the Digital Divide: Internet Connectedness and Age.” *Communication Research* 28: 536-62.

Macy, Michael W. and Robert Willer. 2002. “From Factors to Actors: Computational Sociology and Agent-Based Modeling.” *Annual Review of Sociology* 28: 143-66.

Marsden, Peter and Joel Podolny. 1990. “Dynamic Analysis of Network Diffusion Processes.” In J. Weesie and H. Flap (Eds.) *Social Networks Through Time*. Utrecht: ISOR.

Marx, Karl and Friedrich Engels. 1967 [1888]. *The Communist Manifesto.* New York: Penguin Classics.

Massey, Douglas. 1990. "Social Structure, Household Strategies, and the Cumulative Causation of Migration." *Population Index* 56(1): 3-26.

Massey, Douglas and Felipe Garcia-Espana. 1987. "The Social Process of International Migration." *Science* 237: 733-38.

Massey, Douglas, L. Goldring and Jorge Durand. 1994. "Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities." *American Journal of Sociology* 99:1492-533.

Massey, Douglas and R. Zenteno. 1999. "The Dynamics of Mass Migration." *Proceedings of the National Academy of Sciences* 96(9): 5328-35.

Massey, Douglas and K. E. Espinosa. 1997. "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102(4):939-99.

McPherson, J. Miller and Lynn Smith-Lovin. 1987. "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups." *American Sociological Review* 52(3): 370-379.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2006. "Social Isolation in America: Changes in Core Discussion Networks over Two Decades." *American Sociological Review* 71: 353-375.

McPherson, J. Miller., Lynn Smith-Lovin, and J.M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415-44.

Miller, Daniel and Don Slater. 2000. *The Internet: An Ethnographic Approach*. New York: Berg.

Miyata, Kakuko, Jefrey Boase, Barry Wellman, and Ken'ichi Ikeda. 2005. "The Mobile-izing Japanese: Connecting to the Internet by PC and Webphone in Yamanashi." Pp. 143-64 in *Personal, Portable, Pedestrian: Mobile Phones in Japanese Life*, ed. Mizuko Ito, Daisuke Okabe, and Misa Matsuda. Cambridge, Massachusetts: MIT Press.

Morgenstern, Oskar. 1948. "Demand Theory Reconsidered." *Quarterly Journal of Economics* 62: 165-201.

Mueller, Milton, and Jorge Reina Schement. 1996 "Universal Service from the Bottom Up: A Study of Telephone Penetration in Camden, New Jersey." *The Information Society* 12: 273-292.

OECD. 2003. *OECD Communications Outlook: 2003 Edition.* OECD: Science and Information Technology Series, Volume 2003, Number 4.
http://puck.sourceoecd.org/vl=5968845/cl=15/nw=1/rpsv/~6681/v2003n4/s1/p1l Last accessed August 17, 2009.

OECD. 2005. *OECD Communications Outlook: 2005 Edition.* OECD: Science and Information Technology Series, Volume 2005, Number 11.
http://titania.sourceoecd.org/vl=1021756/cl=19/nw=1/rpsv/~6681/v2005n11/s1/p1l Last accessed, August 17, 2009.

OECD. 2007. *OECD Communications Outlook: 2007 Edition.* OECD: Information and Communications Technologies Series, Volume 2007, no. 4.
http://titania.sourceoecd.org/vl=1021756/cl=19/nw=1/rpsv/~6671/v2007n4/s1/p1l Last accessed, August 17, 2009.

Offe, Claus and Helmut Wiesenthal. 1980. "Two Logics of Collective Action." *Political Power and Social Theory* 3: 67-115.

Pew Internet and American Life Project. 2008. "Percentage of U.S. Adults On-Line." http://www.pewinternet.org/trends/Internet_Adoption_3.18.08.pdf Last accessed March 26, 2008.

Phongpaichit, Pasuk. 1980. "The Open Economy and Its Friends: The "Development" of Thailand." *Pacific Affairs* 53(3): 440-460.

Popielarz, Pamela and J. Miller McPherson. 1995. "On the Edge or In Between: Niche Position, Niche Overlap, and the Duration of Voluntary Association Membership." *American Journal of Sociology* 101: 698-720.

Portes, Alejandro. 1998. "Social Capital" Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 24: 1-24.

Portes, A. and P, Landolt. 2000. "Social Capital: Promise and Pitfalls of its Role in Development." *Journal of Latin American Studies* 32: 529-547.

Putnam, Robert. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture." *Scandinavian Political Studies* 30:137-74.

Rindfuss, R., T. Kaneda, A. Chattopadhyay, and C. Sethaput. 2007. "Panel Studies and Migration." *Social Science Research* 36: 374–403.

Quillian, Lincoln. 2006. "Can Social Capital Explain Persistent Racial Poverty Gaps?" National Poverty Center Working Paper Series #06-12, Northwestern University, Evanston, Illinois.

Quillian, Lincoln and Mary E. Campbell. 2003. "Beyond Black and White: The Present and Future of Multiracial Friendship Segregation." *American Sociological Review* 68: 540-66.

Rogers, Everett M. 2003. *Diffusion of Innovations,* 5th edition. New York: Free Press.

Rosenkopf, Lori and Eric Abrahamson. 1999. "Modeling Reputation and Information Influences in Threshold Models of Bandwagon Innovation Diffusion." *Computational and Mathematical Organization Theory*. 5: 361-84.

Rossman, Gabriel, Ming Ming Chiu, and Joeri Mol. 2006. "Modeling Diffusion of Many Innovations Via Multilevel Diffusion Curves: Payola in Pop Music Radio." California Center for Population Research, On-Line Working Paper Series, CCPR-051-06, UCLA.

Sampson, Robert J., Jeffrey D. Morenoff and Thomas Gannon-Rowley. 2002. "Assessing `Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28: 443-78.

Schelling, Thomas. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1: 143-86.

Schement, Jorge. 1995. "Beyond Universal Service: Characteristics of Americans without Telephones, 1980-1993." *Telecommunications Policy* 19: 477-85.

Selwyn, Neil, Stephen Gorard and John Furlong. 2005. "Whose Internet is it Anyway?: Exploring Adults' (Non) Use of the Internet in Everyday Life." *European Journal of Communication* 20: 5-26.

Shrum, Wesley, Neil Cheek and S.M. Hunter. 1988. Friendship in school: gender and racial homophily. *Sociology of Education* 61:227–39

Shy, Oz. 2001. *The Economics of Network Industries.* 2001. New York: Cambridge University Press.

Shy, Oz. 2007. "Dynamic Models of Religious Conformity and Conversion: Theory and Calibration." *European Economic Review* 51: 1127-53.

Simon, Herbert. 1957. "The Compensation of Executives." *Sociometry* 20: 32-36.

Skvoretz, John. 1990. "Biased Net Theory: Approximations, Simulations, and Observations." *Social Networks* 12: 217–38.

Stark, O. and E. Taylor. 1989. "Relative Deprivation and International Migration." *Demography* 26: 1-14.

Strang, D. 1991. "Adding Social Structure to Diffusion Models." *Sociological Methods and Research* 19(3): 324-353.

Tarde, Gabriel. 1903. *The Laws of Imitation.* New York: Henry Holt.

Tavory, Ido and Ann Swidler. 2009. "Condom Semantics: Meaning and Condom Use in Rural Malawi." *American Sociological Review* 74: 171-89.

Tilly, Charles. 1999. *Durable Inequality*. Berkeley: University of California Press.

Tilly, Charles. 2006. *Identities, Boundaries, and Social Ties.* Boulder, Colorado: Paradigm Publishers.

Turkle, Sherry. 1995. *Life on the Screen*. New York: Simon and Schuster.

Turner, Fred. 2006. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism.* Chicago: University of Chicago Press.

Uhlig, Harald. 2006. "Regional Labor Markets, Network Externalities and Migration: The Case of German Reunification." SFB 649 Discussion Paper 2006-004. Berlin: Humboldt University.

Van den Bulte, Christophe and Gary Lilien. 2001. "Medical Innovation Revisited: Social Contagion versus Marketing Effort." *American Journal of Sociology* 106: 1409-35.

Van den Bulte, Christophe and Stefan Stremersch. 2004. "Social Contagion and Income Heterogeneity in New Product Diffusion." *Marketing Science* 23: 530-44.

Watts, Duncan J. 2003. *Six Degrees: The Science of a Connected Age*. New York: W.W. Norton.

Watts, Duncan J., Peter Dodds, and Mark Newman. 2002. "Identity and Search in Social Networks." *Science* 296: 1302-05.

Watts, Duncan. 1999. "Networks, Dynamics and the Small-World Phenomenon." *American Journal of Sociology* 105:

Weber, Max. 1922 [1968]. "Class, Status, and Party." In *Economy and Society*. New York: Bedminster Press.

Western. Bruce. 2006. *Punishment and Inequality in America.* New York: Russell Sage Foundation.

Western, Bruce, Deirdre Bloome, and Christine Percheski. 2008. "Inequality Among American Families with Children, 1975 to 2005." *American Sociological Review* 73: 903-20.

White, Harrison C. 1970. *Chains of Opportunity: System Models of Mobility in Organizations*. Cambridge: Harvard University Press.

World Bank. 2008. *World Development Report 2008.* http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTWDRS/EXTWDR2008/0,,menuPK:2795178~pagePK:64167702~piPK:64167676~theSitePK:2795143,00.html Last accessed March 25, 2008..