NBER WORKING PAPER SERIES

SOCIAL NETWORKS AND THE AGGREGATION OF INDIVIDUAL DECISIONS

D. Lee Heavner
Lance Lochner

Working Paper 8979
http://www.nber.org/papers/w8979

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2002

Social Networks and the Aggregation of Individual Decisions
D. Lee Heavner and Lance Lochner
NBER Working Paper No. 8979
June 2002
JEL No. D8, H0, J0

ABSTRACT

This paper analyzes individual decisions to participate in an activity and the aggregation of those decisions when individuals gather information about the outcomes and choices of (a few) others in their social network. In this environment, aggregate participation rates are generally inefficient. Increasing the size of social networks does not necessarily increase efficiency and can lead to less efficient long-run outcomes. Both subsidies for participation and penalties for non-participation can increase participation rates, though not necessarily by the same amount. Punishing non-participation has much greater effects on participation rates than rewarding participation when current rates are very low. A program that provides youth with mentors who have participated themselves can increase participation rates, especially when those rates are low. Finally, communities plagued by the flight of successful participants will experience lower short- and long-run participation rates.

D. Lee Heavner
Economic Analysis LLC
2049 Century Park East, Suite 2310
Los Angeles, CA 90067

Lance Lochner
Department of Economics
University of Rochester
Rochester, NY 14627,
Hoover Institution,
and NBER
lochner@hoover.stanford.edu

Social Networks and the Aggregation of Individual Decisions*

D. Lee Heavner
Economic Analysis LLC

Lance Lochner
University of Rochester
Hoover Institution
NBER

April 10, 2002

Abstract

This paper analyzes individual decisions to participate in an activity and the aggregation of those decisions when individuals gather information about the outcomes and choices of (a few) others in their social network. In this environment, aggregate participation rates are generally inefficient. Increasing the size of social networks does not necessarily increase efficiency and can lead to less efficient long-run outcomes. Both subsidies for participation and penalties for non-participation can increase participation rates, though not necessarily by the same amount. Punishing non-participation has much greater effects on participation rates than rewarding participation when current rates are very low. A program that provides youth with mentors who have participated themselves can increase participation rates, especially when those rates are low. Finally, communities plagued by the flight of successful participants will experience lower short- and long-run participation rates.

1 Introduction

Economic models typically assume that decision makers are fully informed, rationally process all information, and generally make the 'right' economic decision. Yet, for many decisions like finishing high school, attending college, joining a gang, taking drugs, or engaging in premarital sex, youth possess limited information about the potential outcomes of their choices. Furthermore, much of that information seems to come from older friends, neighbors, acquaintances, and parents who have already made this decision and can offer their advice. In short, youth look to role models and their social networks for advice.

To the extent that social networks are drawn from the neighborhoods in which youth live, the information and choices of youth will depend critically on the choices made by and outcomes of others in their neighborhood. This does not bode well for youth growing up in impoverished ghetto neighborhoods, as one such youth explains

*Much of this work was completed while Lee Heavner was at the A.B. Freeman School of Business, Tulane University. We thank participants at the 2000 SED Conference in Costa Rica and the University of Rochester Applied lunch for their comments.

> Our neighborhoods will only show us so much...See, I'm twenty-one years old and already an OG [original gangster – someone who has been in a gang for a number of years]...The OG's is getting a little bit younger, and I helped contribute to that because the younger ones were looking up to me just as I was looking up to older people. All I saw for a long time was people trying to get over and hustle, just hustle, hustle. I never seen no doctor get up at three in the morning, wash off his car, take his briefcase and his stethoscope, and go straight to the hospital. I didn't see no lawyers, no rocket scientists. All I seen was just somebody trying to get over. (Marshall and Wheeler, 1996)

Attempting to better understand these issues, this paper examines the evolution of community outcomes and the role of government policy when information is transmitted imperfectly through social networks. While standard economic models emphasize how heterogeneous tastes and abilities can lead to different choices across individuals, our model focuses on the role of heterogeneity in information received by decision makers. Many of our model's implications differ markedly from the predictions of models with fully informed decision makers.

A number of recent theoretical papers have studied the aggregation of decisions when individuals are not fully informed about their choices. The information cascade literature has examined decision-making when individuals sequentially choose whether or not to take some action based on their own private signals and the history of actions taken by their predecessors (e.g. Banerjee, 1992, and Bikchandani, Hirshleifer, and Welch, 1992).[1] In related work, Ellison and Fudenberg (1993, 1995) and Banerjee and Fudenberg (1999) consider the aggregation of decisions and information when a large number of decision makers receive signals about the payoffs to different choices and know the choices made by individuals during the previous period (but they do not observe the entire history of choices), which they refer to as word-of-mouth learning. This literature has primarily focused on understanding when economies characterized by limited information transmission converge to an equilibrium of decision making conformity in which decision makers ignore their own private signals.

Following the theoretical interest in social networks, a number of recent empirical papers have empirically examined the role of social information transmission in a variety of contexts. Most have concluded that learning from others is an important determinant of individual decisions, although little is known about the actual decision rules used by agents.[2] Contrary to the predictions of

[1] See Bikchandani, Hirshleifer, and Welch (1998) for an excellent discussion of this literature.

[2] Empirical findings suggest that social learning may play an important role in the decision to use contraception (Munshi and Myaux, 2000), criminal activity (Glaeser, Sacerdote, and Scheinkman, 1996), welfare participation

the information cascade and word-of-mouth learning literatures; however, it is rare to observe any (interesting) choice for which all individuals, even within homogenous groups, make the same decision. In order to analyze the effects of policy intervention and network size on aggregate outcomes, we, therefore, develop a model in which information is imperfectly transmitted through interactions with others *and* in which identical individuals sometimes make different choices even in the long-run steady state. The policies we examine are either ill-suited for study in models with full information or produce very different effects when information is transmitted through social networks.

We model the decision to participate in an activity when decision makers learn about the participation choices and receive signals about the payoffs associated with those choices from a small sample of individuals from the previous generation. Our framework differs from the information cascade literature and Banerjee and Fudenberg (1999) in that we assume that each generation uses the same decision rule for participation. This assumption is appropriate when individuals from different generations have similar prior beliefs about the state of the current regime or when individuals are limited in their abilities to compute complex Bayesian decision rules and, instead, use simple rules of thumb (as assumed in Ellison and Fudenberg, 1993, 1995). Our decision process is also general enough to account for individual tastes for conformity or the fact that "...a person's patterns and norms of behavior tend to be shaped by those with which he or she has the most frequent or sustained contact and interaction." (Wilson, 1986, p. 61) Because we impose very few assumptions on the form of the decision rule, our results apply to a wide class of social learning models.

We find that the economy will only converge to an equilibrium in which all individuals make the same choice if the 'must-see' to choose assumption (i.e. individuals never choose an option that they do not observe in their sample) of Ellison and Fudenberg (1993, 1995) is imposed. Because full conformity is an unrealistic prediction for most interesting social and economic problems, we relax this assumption in our analysis. Our main results focus on understanding how various policies effect both short- and long-run aggregate outcomes.

Our model generates many interesting predictions that contrast with the predictions of traditional economic models with fully informed decision makers. The imperfect transmission of information through social networks generally causes some individuals not to participate even when

(Bertrand, Luttmer, and Mullainathan, 2000), decisions about employment retirement plans (Duflo and Saez, 2000) and health insurance plans (Sorensen, 2001), stock market participation (Hong, Kubik, and Stein, 2001), and technological adoption in agriculture (Foster and Rosenzweig, 1995, and Conley and Udry, 2001, 2002). See Manski (1993) for an analysis of identification and estimation of actual social learning decision rules.

participation is *ex ante* optimal. This implies that there may be a positive role for government intervention. For example, subsidizing the best choice or punishing an inferior one can induce more individuals to make the right decision.

Contrary to conventional wisdom based on models with fully informed decision makers, aggregate choices need not respond symmetrically to changes in the payoffs of different alternatives when information is transmitted through social networks. For example, an increase in the earnings of high school graduates need not have the same impact on graduation rates as a similar decrease in earnings among drop outs – observing the net change in the graduate - drop out earnings differential is not sufficient to determine the aggregate response. The key to this asymmetry lies in the information (or, more appropriately, lack of information) people obtain from their observation sample. When aggregate graduation rates are low, most decision makers observe samples with very few graduates. As a result, only limited information about an increase in the earnings of graduates will reach them. On the other hand, information about a reduction in the earnings of drop outs will be conveyed clearly. The same reasoning shows that communities with high graduation rates will respond more to improvements in the earnings of graduates and less to reductions in the earnings of drop outs. Therefore, policies that improve the outcomes of graduates may successfully reduce drop out rates in middle and upper class communities while failing to do so in disadvantaged communities.

The problems associated with imperfect information transmission are not necessarily improved by blindly increasing the size of social networks (the number of people agents receive signals from). But, selectively augmenting social networks with individuals who have made good decisions, as a mentor program might do, offers promise. Empirically, mentor programs have been introduced to disadvantaged communities where high school drop out and crime rates are high. Tierney and Grossman (1995) find that the well-known Big Brothers/Big Sisters program reduces drug/alcohol use and violence while increasing school attendance and average grade levels.[3] We model mentor programs by adding a randomly drawn participant to each decision maker's observation sample. In communities with low aggregate participation rates, this can substantially improve information about the benefits of participation and induce many more to choose participation. The effects are substantially weaker in areas with already high participation rates.

Wilson (1987) has identified the outflow of successful blacks from ghetto neighborhoods as

[3] Johnson (1998) reports that the probability of a successful mentoring relationship is positively related to similarities between the mentor and youth, which suggests that youth rely more on information from others with similar attributes.

an important cause for the continued problems faced by these communities. We study how this phenomenon affects short- and long-term participation rates and find that both are reduced. In contrast, communities in which participants are particularly active in the social networks of youth will experience higher short- and long-term participation rates. In contrast to the previously defined mentor programs, a policy that encourages active involvement of participants in their communities will improve the outcomes most in neighborhoods with intermediate aggregate participation rates.

The next section describes a general model of social networks and identifies some important properties of the economy. Section 3 applies the model to analyze the decision to graduate from high school using a decision rule based on that of Ellison and Fudenberg (1993). We explore the affects of network size, mentors, an outflow of participants, and changes in the returns to different choices on short- and long-term aggregate outcomes. We offer some concluding remarks in the final section.

2 A General Model of Social Networks

We consider an environment in which individuals must choose whether or not to participate in an activity, focusing on situations in which: (i) accurate aggregate statistics either do not exist or individual heterogeneity makes nationwide statistics uninformative for individual decision makers; and (ii) individuals cannot acquire information from their own previous experiences.[4] It seems best to consider each local community as a separate economy, since in this case, individuals are likely to be similar to and rely on the information provided by other individuals within that community.[5] While national statistics about education, crime, teen pregnancy, etc. are publicly available, community-level statistics on these types of activities are more difficult to come by.

We recognize that decision makers infer information from observed payoffs as well as from observed participation rates. Observed outcomes of individuals from an earlier generation provide direct information about one's own expected outcomes, and observed participation rates give the fraction of previous decision makers that expected participation to be optimal. Intuitively, observed participation rates contain information about historical returns from participating, since individuals in the observed sample based their participation decisions on earlier observed payoffs and participation rates. The relative popularity of a choice in one's own sample serves as a proxy for that option's historical performance.

[4] In a context with repeated decisions, Lochner (2001) shows that individuals learn about the probability of arrest from their own experiences with crime and whether or not they are arrested.

[5] For example, national measures of the earnings of high school and college graduates may be only weakly related to the potential earnings opportunities for youth growing up in inner city ghettos.

To focus on the effects of limited information transmission, we study individuals that are homogeneous in tastes and abilities. We also abstract from the decision to acquire information by assuming that each decision maker randomly samples n independent members of the previous generation. (We emphasize observation samples that are small but greater than one.) A decision maker learns the participation choices and the payoffs associated with each individual he samples. We follow Banerjee and Fudenberg (1999) and Ellison and Fudenberg (1993, 1995) in assuming that decision makers cannot access the information used by previous decision makers.[6]

Let x denote the number of observed sample members who participated. Participating individual i conveys the payoff signal $\pi_p^i = \theta_p + \varepsilon_p^i$, and non-participating individual j conveys the payoff signal $\pi_{np}^j = \theta_{np} + \varepsilon_{np}^j$ where the ε_a are independent and identically distributed in the population as $N(0, \sigma_a^2)$ for $a \in \{p, np\}$. The ε_a terms represent a combination of randomness in returns and noise in the payoff signals.[7] Let $\theta \equiv \theta_p - \theta_{np}$ represent the expected gain from participating, which is common across all individuals. As a normalization, assume that $\theta > 0$, so that participation is always the *ex ante* optimal choice. In a world with full information, everyone would choose to participate and the economy would always be efficient (i.e. everyone would make the *ex ante* optimal choice).

Consider the following individual decision rule for participation:

$$\text{participate if and only if} \begin{cases} \bar{\pi}_p \geq \gamma^p(n) & \text{when } x = n \\ \bar{\pi}_p \geq \gamma(\bar{\pi}_{np}, x, n) & \text{when } x \in [1, n-1] \\ \bar{\pi}_{np} < \gamma^{np}(n) & \text{when } x = 0, \end{cases}$$

where $\gamma^p(n)$, $\gamma^{np}(n)$, and $\gamma(\bar{\pi}_{np}, x, n)$ are exogenously given threshold functions. We assume that γ is weakly increasing in $\bar{\pi}_{np}$ and would generally expect that it is weakly decreasing in x. We make no assumptions about the effects of network size (n) on the decision rule. Individuals who observe only participants will choose to participate themselves as long as the average outcome signal for participants, $\bar{\pi}_p$, is high enough, while those who only observe non-participants will choose to participate if the average outcome signal for observed non-participants, $\bar{\pi}_{np}$, is low enough. Individuals observing a sample consisting of both participants and non-participants will choose to participate when the average participant signal is sufficiently high, where higher signaled payoffs from non-participants and a lower observed participation rate increase the minimum average participant

[6]This assumption is most reasonable when (i) decision makers have little incentive to record such information after making their decision and (ii) information acquisition is too costly (relative to its rewards) for an organization to profit from recording and communicating such information.

[7]The source of variation is irrelevant for the individual decision process and the dynamics of the aggregate system as long as it does not affect individual decision rules. Variation may also be due to unknown *ex ante* heterogeneity, which should not affect decisions as long as it is not recognized by decision makers.

payoff signal that will induce participation.

This decision rule may reflect a simple rule-of-thumb decision process due to bounded rationality as in Ellison and Fudenberg (1993, 1995). Or, it may represent an underlying Bayesian decision process where individuals from all generations have the same set of prior beliefs over $(\theta_a, \varepsilon_a)$ for $a \in \{p, np\}$ and over the number of generations that have passed since the current (θ, ε) regime began.[8] Finally, it may simply represent individual tastes for conformity or the social transmission of norms and values across people.

Given the decision rule described above and the independence of outcome signals across individuals, the probability that an individual observing x participants will choose to participate themselves is given by μ_x where

$$\mu_x = \begin{cases} \Phi\left(\frac{\theta_p - \gamma^p(n)}{\sigma_p/\sqrt{n}}\right) & \text{when } x = n \\ \int\limits_{-\infty}^{\infty} \Phi\left(\frac{\theta_p - \gamma\left(\theta_{np} + z\frac{\sigma_{np}}{\sqrt{n-x}}, x, n\right)}{\sigma_p/\sqrt{x}}\right)\phi(z)dz & \text{when } x \in [1, n-1] \\ \Phi\left(\frac{\gamma^{np}(n) - \theta_{np}}{\sigma_{np}/\sqrt{n}}\right) & \text{when } x = 0, \end{cases}$$

and $\Phi(\cdot)$ and $\phi(\cdot)$ represent the standard normal cdf and pdf, respectively.

Aggregate Participation Rates

Assuming that individual samples are randomly drawn from a large population, we can determine the evolution of the aggregate participation rate using the law of large numbers. If y is the current population participation rate, then the probability an individual observes j participants in his sample is

$$Pr(x = j|y) = \binom{n}{j} y^j (1-y)^{n-j}.$$

The participation rate next period, y', is then given by

$$y' = \sum_{j=0}^{n} \binom{n}{j} y^j (1-y)^{n-j} \mu_j.$$

[8] For example, suppose individuals know the joint distribution of $(\varepsilon_p, \varepsilon_{np})$ and that this distribution is time invariant. If regimes characterized by (θ_p, θ_{np}) can change at any time with some probability q, each regime is independent of the entire history, individuals do not know when a regime change occurs, and all individuals have the same prior beliefs about q and the distribution from which (θ_p, θ_{np}) is drawn, then individuals from each generation will face the same decision problem and will follow the same Bayesian decision rule. It is important that individuals are unaware of the date when regimes begin; otherwise Bayesian decision rules will differ across generations – later generations will place more importance on observed participation rates (e.g. Banerjee and Fudenberg, 1999). While we do not explicitly analyze a stochastic economy, one can think of the steady states in our model applying within any given regime.

The economy will be in *steady state* if $\Delta(y) \equiv y' - y = 0$. The following proposition establishes that neither full participation nor zero participation will be steady states when $\gamma^p(n) > -\infty$ and $\gamma^{np}(n) > -\infty$.

Proposition 1. *(i) If $\gamma^p(n) > -\infty$, then $y = 1$ is not a steady state. (ii) If $\gamma^{np}(n) > -\infty$, then $y = 0$ is not a steady state. (iii) If $\gamma^p(n) > -\infty$ and $\gamma^{np}(n) > -\infty$, then there must exist at least one interior steady state, y^*, satisfying $y^* \in (0, 1)$.*

Proof: (i) $\Delta(0) = \mu_0 > 0$ when $\gamma^{np}(n) > -\infty$, and (ii) $\Delta(1) = \mu_n - 1 < 0$ when $\gamma^p(n) > -\infty$. (iii) By the Intermediate Value Theorem, there must be some $y \in (0, 1)$ for which $\Delta(y) = 0$. □

Thus, the economy will be inefficient (i.e. some individuals will not participate even though participation is *ex ante* optimal) unless individuals use a 'must see' to choose decision rule in which case they never choose an option they do not observe in their sample (i.e. $\gamma^p(n) = \gamma^{np}(n) = -\infty$). If individuals do use a 'must see' decision rule, then both zero and full participation are possible steady states; although they are unlikely to be stable unless the sample participation rate plays a strong role in the decision process.[9]

When individuals may choose an option they do not observe, there will be at least one interior steady state with partial participation, and there may be as many as n steady states (no more than $\frac{n}{2} + 1$ can be stable). Consequently, there will be few stable steady states when networks are small.

Since $\theta > 0$, a higher participation rate represents a more efficient economy. Proposition 2 shows that by increasing θ, either by subsidizing participants or by taxing non-participants, the government can increase aggregate participation rates and improve efficiency of the economy (ignoring any deadweight losses associated with such policies). Of course, these policies will also increase *ex post* inequality between participants and non-participants.

Proposition 2. *For any given level of participation, y, (i) changes in participation rates are weakly increasing in θ_p $\left(\frac{\partial \Delta(y)}{\partial \theta_p} \geq 0\right)$ and (ii) changes in participation rates are weakly decreasing in θ_{np} $\left(\frac{\partial \Delta(y)}{\partial \theta_{np}} \leq 0\right)$. (iii) For almost every y, changes in participation rates are not equivalent for the same increase in θ_p and decrease in θ_{np} $\left(\frac{\partial \Delta(y)}{\partial \theta_p} \neq -\frac{\partial \Delta(y)}{\partial \theta_{np}}\right)$. (iv) An increase (decrease) in θ need not increase (decrease) changes in the participation rate.*

Proof: See Appendix A.

[9] In our context, a steady state is stable if the system returns to the same steady state after small perturbations. The zero participation steady state will be stable if and only if $\mu_1 < 1/n$ and full participation will be a stable steady state if and only if $\mu_{n-1} > (n-1)/n$. An interior steady state will only be stable if $\frac{\partial \Delta(y)}{\partial y}|_{y=y^*} < 0$.

As one would expect, increases in the payoffs to participants and reductions in the payoffs to non-participants both weakly increase aggregate participation rates. However, in contrast to standard models with fully informed decision makers, increases in participation payoffs and reductions in non-participation payoffs of the same magnitude will have differential effects on aggregate participation rates. Thus, it matters if the return to participation increases because participants have become better off or if it is because non-participants have become worse off. Put another way, net changes in payoffs are not enough to determine whether aggregate participation rates increase or decrease.

The key to this asymmetric response lies in the information decision makers obtain from their social networks and not necessarily in any asymmetry of individual decision rules. The differential effects of changes in θ_p and θ_{np} are most clear at very high and low participation rates. For example, when the participation rate is very high, an increase in the payoff to participation will be immediately recognized by most decision makers, since they will primarily observe participants in their network sample. However, a reduction in non-participant payoffs will go largely unnoticed, since few individuals will observe any non-participants in their sample. Consequently, an equal increase in participant and non-participant payoffs (leaving θ unchanged) will increase changes in participation rates when current participation rates are high and will reduce those changes when current participation rates are low. A steady state participation rate that is high will increase, while a low steady state participation rate will decline in response to increases in payoffs to both participation and non-participation.

3 An Example

In this section, we study the affect of changes in the payoffs to different choices, mentors, network size, and an outflow of participants on aggregate participation rates using a simple rule-of-thumb decision rule for participation based on Ellison and Fudenberg (1993). Specifically, assume that $\gamma(\bar{\pi}_{np}, x, n) = \bar{\pi}_{np} + m(1 - 2\frac{x}{n})$, where m is a non-negative constant determining how much weight decision makers place on the information contained in the observed participation rate. Individuals observing mixed samples of participants and non-participants, therefore, choose to participate if the difference in average observed payoffs ($\bar{\pi}_p - \bar{\pi}_{np}$) exceeds the cutoff $m(1 - 2\frac{x}{n})$. We assume that γ^{np} and γ^{p} are equal, finite constants, and independent of n.[10]

This decision rule provides a simple way to incorporate the information contained in observed returns and observed participation rates (it may also incorporate tastes for conformity). The

[10]See Appendix B for a more detailed discussion.

parameter m determines the importance of the sample participation rate in determining individual decisions when decision makers observe a mixed sample of participants and non-participants. For m very large, individuals will simply follow the choice made by the majority of their sample regardless of the payoff signals they receive. For $m = 0$, individuals only compare the payoff signals they observe for participants and non-participants, ignoring the sample participation rate altogether. Even though the decision rule treats $\bar{\pi}_p$ and $\bar{\pi}_{np}$ symmetrically, we will observe asymmetric responses to changes in θ_p and θ_{np} as described earlier in Proposition 2.

Any number of patterns for $\Delta(y)$ can be generated by this model depending on the parameterization. While the broad issues we discuss are not sensitive to the parameters we choose, it is instructive to present a specific example. We apply the model to the high school graduation decision, noting that evidence presented in Betts (1996) and Dominitz and Manski (1996) suggests that youth are not fully informed about the earnings associated with different education levels. Betts (1996) also demonstrates that the errors in student earnings expectations are positively correlated with parental earnings levels, suggesting that parents are in their children's social networks. We choose parameter values for our model that produce a pattern consistent with the 1970-80 changes in high school graduation rates (among young white men) across metropolitan areas in the U.S. and the average earnings of high school graduates and drop outs in 1970. As Figure 1 shows, high school graduation rates increased much more from 1970 to 1980 among young men living in metropolitan areas with lower initial graduation rates than those living in areas with high graduation rates. While we recognize that many models might explain this pattern, we focus exclusively on the role of social networks in what follows.[11]

Linking the model to our example, let participation represent the decision to finish high school and non-participation represent the decision to drop out. Then, we set $\theta_p = 8.595$ and $\theta_{np} = 8.352$ based on average log earnings among young graduates and drop outs. We also assume a network size of $n = 6$ along with $m = \sigma^p = \sigma^{np} = 1$ for our base case. (See Appendix B for further details on the base case parameterization.) As seen in Figure 2, our base case parameter values produce a pattern for $\Delta(y)$ like that of Figure 1 (note that Figure 1 only shows $\Delta(y)$ for a limited range of y values for which data are available). The only steady state participation rate is approximately $y^* = 0.85$, and it is stable. Changes in the participation rate, $\Delta(y)$, are initially increasing for very

[11]We note that two obvious explanations are not supported by the data. First, measurement error in graduation rates across metropolitan areas cannot fully explain the negative correlation between changes in graduation rates and initial rates. Second, the pattern in Figure 1 is not generated by differential changes in the rate of return to high school completion across metropolitan areas. See Appendix B for a more detailed discussion of these empirical issues and the Census data used to generate Figure 1.

Figure 1: 1970-80 Change in High School Graduation Rates by 1970 High School Graduation Rates

(White Males Ages 20-25 by Metropolitan Area)

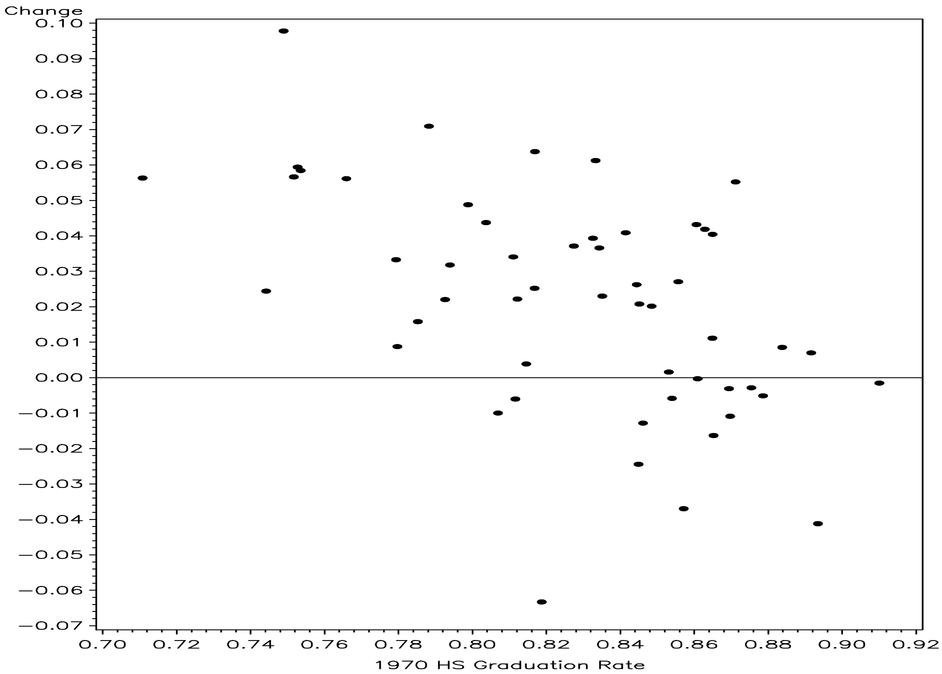
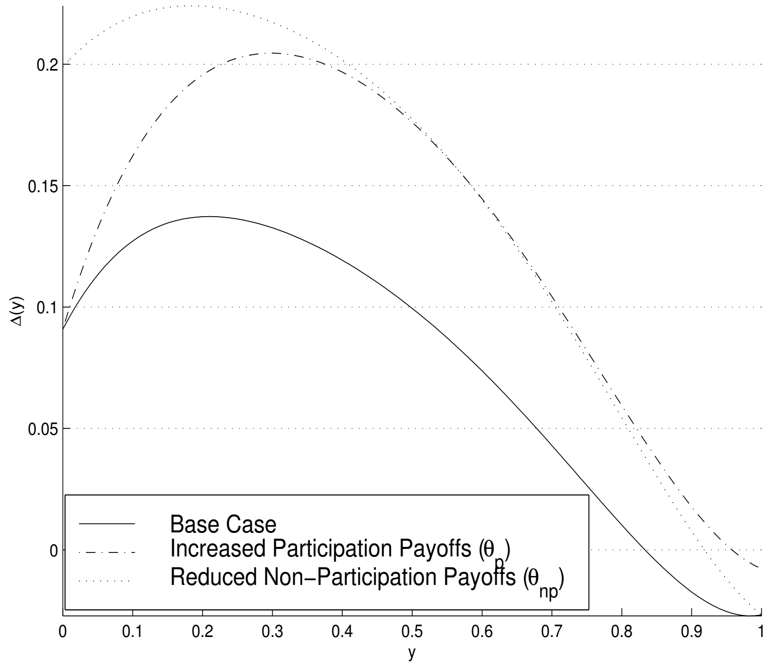


Figure 2: The Effects of Changes in Payoffs to Participation and Non-Participation



low y, then declining thereafter. Thus, convergence to the steady state from a very low participation rate will be slow at first, increasing until $y = 0.2$, then slowing down again until y settles at y^*.

To demonstrate the asymmetric response of $\Delta(y)$ and y^* to changes in θ_p and θ_{np}, Figure 2 also shows the effects of increasing the net payoff to graduation by 0.2 either by raising θ_p or lowering θ_{np}. While both changes increase $\Delta(y)$, the magnitude of the response depends on which payoff has been changed and the current participation rate. An increase in θ_p has very small effects when participation rates are low, while the largest effects of a decline in θ_{np} occur for $y < 0.5$. For $y \in (0.3, 0.9)$, a change in either payoff produces a similar increase in $\Delta(y)$. Because the initial steady state is very high, an increase in the payoff to participation increases the steady state participation rate more than a decline in the payoff to non-participants.

Empirically, the wages of high school graduates and drop outs have changed considerably since the 1960s (e.g. see Katz and Murphy, 1992). From 1963-71, wages increased substantially across all education groups with very similar increases for both high school graduates and drop outs (around 17%). Despite a negligible change in relative wages over that period, the model predicts that schooling should have declined (or increased at a slower rate) in communities with low graduation rates while it should have increased in communities with very high graduation rates. In contrast, the 1980s were characterized by falling wages among both graduates and drop outs (with slightly larger declines among drop outs), which should have produced gains in schooling for the most disadvantaged communities while producing little change in graduation rates for the most educated communities. More generally, declines in wages for both drop outs and graduates should lead to faster convergence in graduation rates across communities, while increases in their wages should slow convergence.

A policy that increases the earnings of graduates or punishes drop outs will not only increase the steady state graduation rate, but it will also increase the rate at which graduation rates move toward that steady state for most values of y.[12] Interestingly, among communities with very low graduation rates, punishing drop outs should generate greater increases in schooling early on than a subsidy to graduates, so the quickest way to raise the graduation rates in such a community may be to raise compulsory schooling ages or the severity of fines associated with violating them.

The Size of Networks

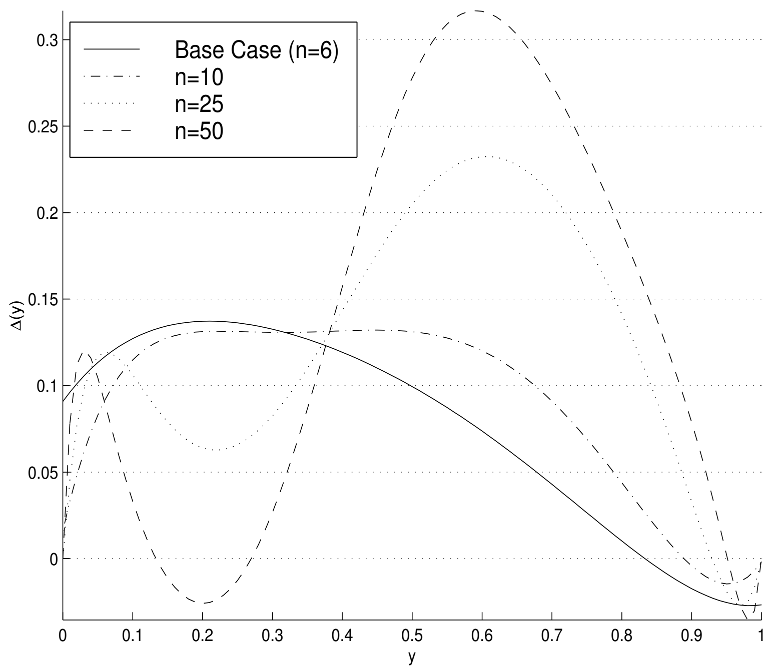
Can we ensure full participation by simply increasing the size of networks, n? As Figure 3 shows, over a large range of n, y^* is, indeed, increasing in n. However, as n grows, so does the possibility that other (lower) steady states emerge.[13] For $n = 50$, two stable steady states exist, one with a participation rate of 0.12 and a second with a participation rate of 0.95 (a third, unstable, steady state also exists at a participation rate of 0.28). In this example, an economy that begins at a low initial participation rate may settle at a low participation steady state when networks are large.[14] Thus, larger networks do not necessarily guarantee that the economy will be more efficient. Intuitively, when networks are very large and participation rates are very small, the 'herding' effect of observing mostly non-participants, which will be the case for most individuals when y is low, can overwhelm the effect of observing better average payoff measures. On the other hand, when participation rates are high, the 'herding' effects work in favor of participation. (Similar tendencies emerge when m is large.)

[12]Note that these results implicitly assume that individuals do not adjust their decision rules in response to the policy. This assumption is appropriate when policies are unannounced or have indirect effects on the payoffs to different choices that are not easily recognized by decision makers. Examples might include the establishment of empowerment zones in disadvantaged communities, increased enforcement of truancy laws, changes in school quality, etc.

[13]For very large n, a full information model is likely to be more appropriate for studying individual decisions.

[14]With other parameterizations, multiple steady states may also exist for smaller network sizes.

Figure 3: The Effects of Network Size

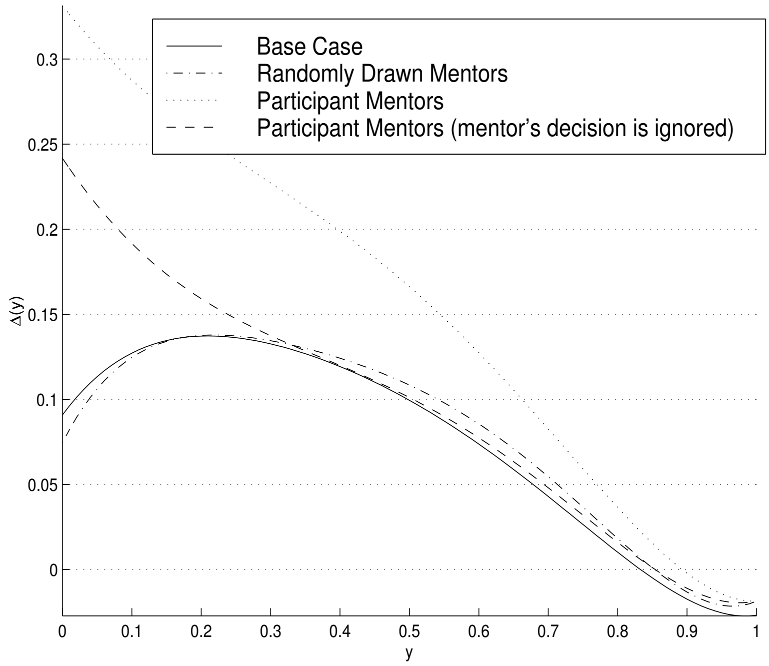


Mentors

We next explore the potential for mentor programs to encourage youth to finish high school. Figure 4 shows $\Delta(y)$ for the base case with $n = 6$ and a case where individuals observe an additional person randomly selected from the population. Both $\Delta(y)$ and y^* are quite similar for the two cases, suggesting that a program which randomly selects mentors for youth should have little effect on the economy.

We next analyze a program that provides youth with a mentor who is randomly drawn from the population of participants, so all youth observe at least one participant in their network. We explore two cases. Consider first the case in which individuals continue to use the same decision rule as before. In this case (the dotted line), $\Delta(y)$ increases for all values of y, especially for lower values. The steady state participation rate rises to about 0.9. Here, a mentor program not only increases the number of participant payoff signals observed for decision makers, which improves each individual's estimate of the payoff to participation, but it also increases the fraction of participants each decision maker observes, which lowers the bar for participation. The former effect is particularly important for those who would otherwise observe few participants, which explains why the effects are so large when aggregate participation rates are low. In a sense, the latter effect assumes that youth are unaware (or ignore) that mentors are drawn exclusively from the participant sample. Instead, youth

Figure 4: The Effects of Mentor Programs



may disregard the information provided by their mentor's participation decision and only use his additional payoff signal. In other words, individuals may adjust their decision rule so that they only consider non-mentors in computing their sample participation rate.[15] In this case (the dashed line), $\Delta(y)$ still increases substantially when aggregate participation rates are low, but there is little effect when participation rates are above 20%.

The way mentors are chosen and the response of decision makers to that selection mechanism is important. A program that randomly assigns a mentor from the population at large is equivalent to an increase in the size of social networks and offers little additional information about the payoffs to high school graduation when aggregate graduation rates are low, since few of the mentors will actually be participants. However, a program that draws mentors from the pool of graduates can have large effects in areas with very low graduation rates, since it provides information about the payoffs to a choice most decision makers are likely to be uninformed about.[16] Mentor programs offer a powerful way to improve outcomes in areas where graduation rates are quite low.

[15] Let $\bar{\pi}_p^m$ represent the average payoff among all observed participants including the mentor, and x represent the number of observed participants (excluding the mentor). Then, an individual observing a sample with $x \in [0, n-1]$ participants chooses to participate if and only if $\bar{\pi}_p^m - \bar{\pi}_{np} \geq m(1 - 2\frac{x}{n})$ (individuals with $x = 0$ observe a single participant signal from their mentor).

[16] Note that if prior beliefs are such that youth believe the payoffs to graduation are quite high (much higher than they actually are), then a program that provides high school graduates as mentors may actually reduce the likelihood that youth choose to graduate. With unbiased prior beliefs (or a symmetric decision rule), this is unlikely.

Active and Inactive Participants

Wilson (1987) forcefully argues that the flight of middle- and working-class blacks from inner city neighborhoods has negatively impacted those communities by reducing the number of positive role models for youth. This can be studied within our framework by assuming that some fraction of participants become inactive or detached from the community. Specifically, assume that only a fraction ξ of participants remain in the population from which youth draw their network samples.[17] As shown in Figure 5, which assumes that 10% of participants become inactive ($\xi = 0.9$), both short- and long-term participation rates are negatively affected by the flight of participants. The largest impacts on short-term participation changes occur for intermediate values of y. If the participants experiencing the highest payoffs are most likely to become inactive or leave the community, then the negative impacts of their inactivity would be even greater. Disadvantaged communities plagued by 'brain drain' or an outflow of successful adults will tend to improve at a slower rate and settle at a worse long-run steady state than otherwise similar communities. On the other hand, communities in which successful adults are more active will face better short- and long-run prospects as shown in Figure 5 (represented by $\xi = 1.1$). In contrast to the mentor programs already discussed, however, policies that generally encourage successful adults to become more active in their communities will have their greatest impacts when participation rates are neither very high nor very low.

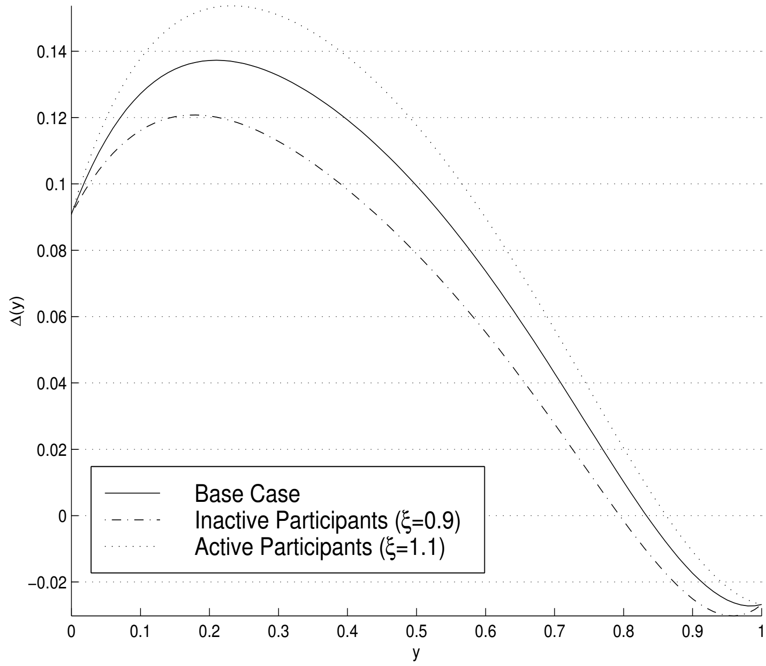
4 Conclusions

The way in which individuals gather information and make decisions has important implications for policies designed to influence decisions to finish high school, attend college, engage in crime, etc. However, most existing economic research ignores the process of information acquisition by assuming that individuals are fully informed about the potential outcomes of their choices. This research cannot explain how mentor programs achieve their effects or why ghetto neighborhoods plagued by the flight of successful adults fail to improve. To address these shortcomings, we study individual decisions when information is acquired from social networks. A number of our findings can be useful for guiding policy in a variety of social contexts.

Our model of social networks suggests that steady state participation rates will generally be inefficient. When participation in an activity is optimal, subsidies to participants or penalties on non-participants can improve the efficiency of the economy. It is important to note, however, that

[17] In this case, $y' = \sum_{j=0}^{n} \binom{n}{j} z^j (1-z)^{n-j} \mu_j$ where $z = \xi y/(1 + (1-\xi)y)$. Whether those no longer in the network pool actually leave the community or simply become detached is irrelevant here.

Figure 5: Active and Inactive Participants



the choice between these two policies matters, especially when participation rates are either very low or very high. A subsidy for participants will have little effect on the economy when current participation rates are low, while a penalty on non-participants will have its greatest impact in these cases. The relative effectiveness of these policies is reversed when participation rates are high. This asymmetry has important practical implications. For example, in areas where high school graduation rates are low, a policy that punishes high school drop may increase graduation rates more than a policy that rewards graduates. On the other hand, such a policy may have little effect in communities with high graduation rates. Strict enforcement of compulsory schooling laws should, therefore, boost graduation rates in disadvantaged communities while having little effect in middle and upper class areas. Similarly, a policy that imposes stronger punishments on gang members is likely to reduce gang membership more than programs that improve legitimate opportunities in gang-infested neighborhoods. However, in areas with few gang problems, encouraging legitimate opportunities can be more effective at slowing or preventing the growth of gangs. Finally, programs that indirectly aid teenage mothers are unlikely to increase teen pregnancy in most communities where few teenagers choose to have children, but they may substantially increase teen motherhood in impoverished communities where teen pregnancy is already a problem.

We also find that efficiency is not necessarily improved when networks are expanded. However, efforts to supplement networks with mentors can be effective at raising participation rates if mentors are drawn from the pool of participants rather than the population at large. When aggregate participation rates are low, the additional information about the payoffs to participation provided by mentors who participated themselves can substantially increase participation rates. A mentor program will be most effective in communities where youth are generally drawing their information from a previous generation that was making suboptimal choices. So, providing youth in disadvantaged communities with mentors that finished high school, attended college, and avoided drugs, crime, and gangs is likely to improve aggregate outcomes in all of these dimensions. On the other hand, disadvantaged communities plagued by the flight of successful adults will improve at a slower rate and settle at worse long-run outcomes than otherwise similar communities.

References

Banerjee, A. V. (1992), 'A simple model of herd behavior', *The Quarterly Journal of Economics* **107**, 797–817.

Banerjee, A. V. and Fudenberg, D. (1999), 'Word of mouth learning', mimeo, Massachusetts Institute of Technology.

Bertrand, M., Luttmer, E. and Mullainathan, S. (2000), 'Network effects and welfare cultures', *The Quarterly Journal of Economics* **115**(3), 1019–55.

Betts, J. (1996), 'What do students know about wages? evidence from a survey of undergraduates', **31**(1), 27–56.

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998), 'Learning from the behavior of others: Conformity, fads, and informational cascades', *The Journal of Economic Perspectives* **12**, 151–170.

Conley, T. and Udry, C. (2001), Social learning through networks: The adoption of new agricultural technologies in Ghana. Working Paper.

Conley, T. and Udry, C. (2002), Learning about a new technology: Pineapple in Ghana. Working Paper.

Dominitz, J. and Manski, C. (1996), 'Eliciting student expectations of the return to schooling', **31**(1), 1–26.

Duflo, E. and Saez, E. (2000), Participation and investment decisions in a retirement plan: The influence of colleagues' choices. Working Paper.

Ellison, G. and Fudenberg, D. (1993), 'Rules of thumb for social learning', *The Journal of Political Economy* **101**, 612–643.

Ellison, G. and Fudenberg, D. (1995), 'Word of mouth communication and social learning', *The Quarterly Journal of Economics* **110**, 93–125.

Foster, A. and Rosenzweig, M. (1995), 'Learning by doing and learning from others: Human capital and technical change in agriculture', *The Journal of Political Economy* **103**(6), 1176–1209.

Glaeser, E., Sacerdote, B. and Scheinkman, J. (1996), 'Crime and social interactions', *The Quarterly Journal of Economics* **111**(2), 507–48.

Hong, H., Kubik, J. and Stein, J. (2001), Social interaction and stock-market participation. NBER Working Paper No. 8358.

Johnson, A. (1998), *An Evaluation of the Long-Term Impacts of the Sponsor-A-Scholar Program on Student Performance*, Mathematica Policy Research, Inc., Princeton.

Katz, L. and Murphy, K. (1992), 'Changes in relative wages, 1963-1987: Supply and demand factors', *The Quarterly Journal of Economics* **107**(1), 35–78.

Lochner, L. (2001), A theoretical and empirical study of individual perceptions of the criminal justice system. RCER Working Paper No. 483.

Manski, C. (1993), 'Dynamic choice in social settings: Learning from the experience of others', *Journal of Econometrics* **58**, 121–36.

Marshall, J. and Wheeler, L. (1996), *Street Soldier: One Man's Struiggle to Save a Generation – One Life at a Time*, Dell Publishing, New York.

Munshi, K. and Myaux, J. (2000), Social norms, with an application of demographic transition in Bangladesh. Working Paper.

Sorensen, A. (2001), Social learning in the demand for employer-sponsored health insurance. Working Paper.

Tierney, J. and Grossman, J. (1995), *Making a Difference: An Impact Study of Big Brothers/Big Sisters*, Public/Private Ventures, Philadelphia.

Wilson, W. J. (1987), *The Truly Disadvantaged*, The University of Chicago Press, Chicago.

Appendix A

Proof of Proposition 2:

(i) and (ii) $\frac{\partial\Delta(y)}{\partial\theta_a} = \sum_{j=0}^{n}\binom{n}{j}y^j(1-y)^{n-j}\frac{\partial\mu_j}{\partial\theta_a}$ for $a \in \{p, np\}$. Straightforward calculus yields $\frac{\partial\mu_0}{\partial\theta_p} = \frac{\partial\mu_n}{\partial\theta_{np}} = 0$, $\frac{\partial\mu_n}{\partial\theta_p} = \frac{\sqrt{n}}{\sigma_p}\phi\left(\frac{\theta_p-\gamma^p(n)}{\sigma_p/\sqrt{n}}\right) \geq 0$, and $\frac{\partial\mu_0}{\partial\theta_{np}} = -\frac{\sqrt{n}}{\sigma_{np}}\phi\left(\frac{\theta_{np}-\gamma^{np}(n)}{\sigma_{np}/\sqrt{n}}\right) \leq 0$. For $x \in [1, n-1]$, we obtain

$$\frac{\partial\mu_x}{\partial\theta_p} = \frac{\sqrt{x}}{\sigma_p}\int_{-\infty}^{\infty}\phi\left(\frac{\theta_p-\gamma\left(\theta_{np}+z\frac{\sigma_{np}}{\sqrt{n-x}},x,n\right)}{\sigma_p/\sqrt{x}}\right)\phi(z)dz \geq 0$$

$$\frac{\partial\mu_x}{\partial\theta_{np}} = -\frac{\sqrt{x}}{\sigma_p}\int_{-\infty}^{\infty}\phi\left(\frac{\theta_p-\gamma\left(\theta_{np}+z\frac{\sigma_{np}}{\sqrt{n-x}},x,n\right)}{\sigma_p/\sqrt{x}}\right)\gamma_1\left(\theta_{np}+z\frac{\sigma_{np}}{\sqrt{n-x}},x,n\right)\phi(z)dz \leq 0.$$

Clearly, then $\frac{\partial\Delta(y)}{\partial\theta_p} \geq 0$ and $\frac{\partial\Delta(y)}{\partial\theta_{np}} \leq 0$. (iii) The previous equations imply that the magnitude of each effect will differ for nearly all y values. (iv) Given the above definitions, it is possible to increase both θ_p and θ_{np} such that θ increases and $\Delta(y)$ decreases for some values of y. For other values of y, a decrease in both θ_p and θ_{np} such that θ increases can cause $\Delta(y)$ to decline. A symmetric argument applies to reductions in θ. □

Appendix B

This appendix discusses the data used in generating Figure 1 and in determining the parameter values used in the base case model of Section 3.

Census Data

Data on earnings, age, and education for white men are taken from the 1970 and 1980 Decennial Censuses. In calculating high school graduation rates for Figure 1, we use white men ages 20-25. Graduation rates (and their changes) for fifty-seven Metropolitan Statistical Areas (MSAs) with 250 or more observations are shown in the figure, though the same pattern emerges when the entire sample is used.

We first note that this pattern is not generated by differential changes in the rate of return to high school completion across MSAs. A regression of 1970-80 changes in MSA high school graduation rates on 1970-80 changes in the rate of return to high school[18] and the 1970 graduation rate produces an estimated effect of -0.035 (standard error of 0.035) for changes in returns and -0.367 (standard error of 0.084) for the initial graduation rate.

While measurement error in graduation rates across MSAs will generate a negative relationship between 1970-80 changes in graduation rates and 1970 graduation rates, we now show that this can only explain a fraction of the negative correlation. In particular, we study the bias induced when using sample graduation rates rather than true graduation rates to estimate the relationship between $\Delta(y)$ and y. Let y_j denote the true graduation rate in community j and $\bar{y}_j = y_j + \eta_j$ the sample graduation rate taken from N random persons in that community; η_j represents iid measurement error. If d_j^i is an indicator variable that equals one if individual i in community j graduates and zero otherwise, then d_j^i will have a mean of y_j and variance of $y_j(1 - y_j)$. Since $\bar{y}_j = \sum_{i=1}^{N} d_j^i$, the variance of η_j is given by

$$Var(\eta_j) = \frac{y_j(1 - y_j)}{N} \leq \frac{1}{4N}. \tag{1}$$

Suppose we are interested in estimating the relationship

$$\Delta(y_j) = \alpha + \beta y_j + \omega_j,$$

[18]The returns to high school are estimated from MSA-level regressions of log wage earnings for white men ages 18-65 on indicators for education < 10 years, education $= 10$ or 11 years, education $= 12$ years, education $=$ 13-15 years, and education ≥ 16 years, as well as experience (age - education - 6) and experience-squared. The coefficients on the indicator for 12 years of school for each MSA are used as measures of the return to high school.

but we instead use OLS to estimate

$$\Delta(\bar{y}_j) = \alpha + \beta \bar{y}_j + \nu_j.$$

Assuming ω_j and ν_j are iid errors independent of each other as well as y_j, the OLS estimate of β ($\hat{\beta}$) converges to

$$\frac{Cov(\bar{y}_j, \Delta(\bar{y}_j))}{Var(\bar{y}_j)} = \beta \left(\frac{Var(y_j)}{Var(y_j) + Var(\eta_j)} \right). \tag{2}$$

Assuming β is negative, the OLS estimate provides an upper bound for the true β.

Given the upper bound on $Var(\eta_j)$ from equation (1), we know that $Var(y_j) \geq V(\bar{y}_j) - \frac{1}{4N}$ and, therefore,

$$plim\ \hat{\beta} \leq \left(\frac{4N \cdot Var(\bar{y}_j) - 1}{4N \cdot Var(\bar{y}_j)} \right) \beta. \tag{3}$$

Combining equations (2) and (3), we see that for $\beta < 0$ and $Var(\bar{y}_j) > \frac{1}{4N}$,

$$\left(\frac{4N \cdot Var(\bar{y}_j)}{4N \cdot Var(\bar{y}_j) - 1} \right) plim\ \hat{\beta} \leq \beta \leq plim\ \hat{\beta}.$$

Since we only use MSAs with $N \geq 250$ and $Var(\bar{y}) = .002$, any bias in estimating β or in the slope of $\Delta(y)$ in Figure 1 should be no greater than a factor of two. The empirical bias in the paper is likely to be substantially smaller for two reasons: (1) N is much larger than 250 for most MSAs in the sample and (2) $Var(\eta_j)$ is likely to be considerably smaller than 0.25, since graduation rates are much greater than 0.5.

Base Case Parameterization

In determining θ_p and θ_{np}, we use the 1970 Census, regressing log wage earnings on indicators for education < 10 years, education $= 10$ or 11 years, education $= 12$ years, education $=$ 13-15 years, and education ≥ 16 years, as well as experience (age - education - 6) and experience-squared. The values $\theta_p = 8.595$ and $\theta_{np} = 8.352$ are given by the coefficients on the indicators for 10 or 11 years of school and 12 years of school, respectively.

We then set $n = 6$, $m = \sigma_p^2 = \sigma_{np}^2 = 1$, and $\gamma^{np}(n) = \gamma^p(n) = 7.8068$ to produce a pattern for $\Delta(y)$ matching that of Figure 1.[19]

[19] Other parameterizations can also produce similar patterns over the limited range for y that is observed. The base values for $\gamma^{np}(n)$ and $\gamma^p(n)$ are derived from $m\left(1 - 2\left(\frac{n-1}{n}\right)\right) - \frac{\theta_p + \theta_{np}}{2}$. These thresholds are held fixed at their base values when network sizes are adjusted in our policy experiments.