ABSTRACT

SMITH, KARL WILLIAM. The Prior Trap. (Under the direction of Stephen Margolis).

Basic economic theory suggests that the decision to go to college should be based only on the expected costs and benefits of college. The income of the family the student comes from should have no effect. Yet, it does. The two common explanations for this discrepancy, inadequate primary school funding and liquidity constraints are at odds with the facts. I offer a third explanation, economically disadvantaged students attend college at lower rates because they have biased information. This analysis connects to the existing literature in at least three ways. It provides a rational basis for the neighborhood effect., extends work on human capital development indicating that educational paths are set at or before age 16 and helps provide an explanation for the both the increase in the return to education and the slowdown in college graduation growth among young men in the United States.

The Prior Trap

by

Karl William Smith

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina

2007

APPROVED BY:

Stephen Margolis Chair of Advisory Committee	Charles Knoeber
David Flath	Craig Newmark

DEDICATION

This work is dedicated to my future wife Larisa Yasinovskaya, who has been my partner, my confidant, my conscience and my best friend throughout this entire project. It was through her that I learned I could fake my way through a life unexamined. It was because of her that I set my targets outside of the expertise of my department. It was because of her that I had the courage to pursue this project without regard to its marketability. She has stuck by my side through the good and perhaps mostly bad times that it took to get this far.

Angel, we did it!

BIOGRAPHY

It is often said that good writers know what to write, great writers know what to steal. And, so I steal from the best.

Three passions, simple but overwhelmingly strong, have governed my life: the longing for love, the search for knowledge, and unbearable pity for the suffering of mankind. These passions, like great winds, have blown me hither and thither, in a wayward course, over a great ocean of anguish, reaching to the very verge of despair.

I have sought love, first, because it brings ecstasy - ecstasy so great that I would often have sacrificed all the rest of life for a few hours of this joy. I have sought it, next, because it relieves loneliness--that terrible loneliness in which one shivering consciousness looks over the rim of the world into the cold unfathomable lifeless abyss. I have sought it finally, because in the union of love I have seen, in a mystic miniature, the prefiguring vision of the heaven that saints and poets have imagined. This is what I sought, and though it might seem too good for human life, this is what--at last--I have found.

With equal passion I have sought knowledge. I have wished to understand the hearts of men. I have wished to know why the stars shine. And I have tried to apprehend the Pythagorean power by which number holds sway above the flux. A little of this, but not much, I have achieved.

Love and knowledge, so far as they were possible, led upward toward the heavens. But always pity brought me back to earth. Echoes of cries of pain reverberate in my heart. Children in famine, victims tortured by oppressors, helpless old people a burden to their sons, and the whole world of loneliness, poverty, and pain make a mockery of what human life should be. I long to alleviate this evil, but I cannot, and I too suffer.

This has been my life. I have found it worth living, and would gladly live it again if the chance were offered me

ACKNOWLEGEMENTS

I'd like to thank my mother, my sister, my father who is no longer with us, Ed Whitfield, and Leon Watson for supporting my annoying insatiable questioning from a young age. I thank Mike McElroy, John Lapp, John Seater, Kerry Smith and Patti Clayton for that same support at a slightly older age. I thank my committee for tolerating my work habits. A special thanks goes to Steve Margolis for always inviting me back to his office, no matter what I had done since my last visit. This was truly much more than anyone could reasonably ask for. I thank Pat Inman for increasing the probability that someone might actually be able to understand what I have written.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

The lower a student's income the less likely the student is to attend college. Basic economic theory suggests that the decision to go to college should be based on a comparison of the gains that a student expects to receive from going to college with the costs associated with preparation and attendance. The income of the family the student comes from should have no effect. Yet, it does.

There are two common explanations for the observed relationship between family income and college attendance, but both are at odds with the facts. The first is that students economically disadvantaged families do not have the money and cannot borrow the money to pay for a college education. The evidence suggests, however, that family income leading up to and during college is a much weaker predictor of college attendance than family income during primary school. The relationship between family income and future college attendance is established early in a student's life and so it is unlikely that liquidity constraints are the involved.

The second explanation is that students from less affluent communities attend less effective primary schools. The logic is that school funding is determined in part by local tax revenue and thus students in economically disadvantaged communities will go to schools with fewer resources. The problem here is that school funding doesn't seem to have a strong effect on student outcomes. Poorly funded schools in wealthy communities tend to do better than well funded schools in poor communities.

I offer a third explanation. Economically disadvantaged students go to college at lower levels because they have different expectations than students raised in more affluent communities. The community matters not because it determines school funding but because it determines student expectations. Formally, I argue that economically disadvantaged students attend college at lower rates because they have biased information. A key ingredient in college success is preparation. Students will only spend time and effort preparing for college if they believe that they have reasonable chance of succeeding. Students who grow up in communities where most adults have not completed or even attended college rationally conclude that their chances of completing college are low as well. Therefore, they invest little in college preparation. Because students have not prepared for college they do not get into college or perhaps more commonly they don't even bother applying.

My analysis extends findings in the existing literature in at least three ways. First, it provides a rational basis for the neighborhood effect. Students "prefer" to act like their peers and role models because peers and role models provide valuable information about the optimal way to behave. In short, in a world fraught with uncertainty, group socialization is individually optimal.

Second, this analysis extends work on human capital development which suggests that educational paths are set at or before age 16. Heckman and Carnerio find a crystallization effect during the early teenage years. Keane and Wolpin show that endowments at age 16 are highly determinate in a young men's career path.

Third, this paper helps provide an explanation for the both the increase in the return to education and the slowdown in college graduation growth among young men in the United States. Card and Lemiuex who show that the increase in the return to education could be explained two factors, a linear trend in the demand for educated workers and an exogenous slowdown in the growth rate of the supply of college graduates. The model I build endogenizes that supply side slowdown.

The rest of the paper proceeds as follows. Chapter 2 provides a non-technical overview of the student's information problem. Chapter 3 discusses some background trends in college going. Chapter 4 looks at some competing explanations for those trends and the empirical problems with those explanations. Chapter 5 provides a formal model of the students problem and develops tools for comparative statics of community level behavior. Chapter 6 outlines and provides the results of a numerical simulation of US white male college going rates since 1950 based on the analytics of the model developed in Chapter 5. Chapter 7 provides some concluding remarks.

THE INFORMATION PROBLEM THAT STUDENTS FACE

High school students must make numerous choices that could have long lasting consequences on their earning potential. I focus, however, on one choice: whether or not to devote time and energy preparing for college.

There are essentially three elements that go into making this choice.

1. The cost of studying
2. The benefits of successfully completing college
3. The probability that the student will finish college if he or she studies

Students who devote more time to studying may complete more years of college, on average, than those who study less and students who complete some college earn more on average than students who complete none. To make the analysis as simple as possible, however, I will present the argument in terms of binary choices and consequences. Either, the student prepares for college or does not. Either, the student finishes college or does not. While this reduces the complexity considerably the fundamental issue remains the same: a student must choose how much time and energy to devote to college preparation today, in the hopes that he or she will receive some gains in the future.

Rational students will choose to prepare for college if the expected benefits outweigh the cost. The costs are what the student gives up in order to prepare for college. In lieu of studying students could choose to either work, spend time with their friends and family or engage in other pursuits of their own choosing. Students who choose to devote time to studying cannot also be working and so must give up whatever wages they would have earned. Since all students 16 years of age or older have the opportunity to work at the minimum wage a student must at least give up 5.45 per hour in order to study. However, some students may value other activities at more than 5.45 an hour. If so those students are sacrificing even more.

The expected benefits of studying are determined in two steps. The first step is to calculate the benefits a student can expect if he successfully graduates college. The second step is to factor in the probability that the student will graduate from college if he or she prepares. The expected benefit after graduation multiplied times the probability that of graduation.

In the first, step the student has to add up the value of the extra earnings that a college graduate receives but subtract out the costs of tuition, books, and lost income during college. In the second step the student has to form an estimate of the probability of graduate. It is the student's difficulties in the second step which form the heart of this paper. Even, if the benefits from completing college are high the expected benefit can be low if the probability of success is low.

For example, a student who knows that he or she will receive at net benefit of $250,000 by graduating college but believes that there is a 10% chance of graduating if he or she studies faces an expected benefit of $25,000. If it costs more than $25,000 in lost part time work earnings then it will not be worth it to study. While both the benefit and the probability of graduating may be difficult to calculate, I argue that the student is likely to be much more uncertain about his chances of graduating than the rewards if he graduates.

A student's chances of graduating college depend at least in part on his scholastic ability. How scholastic ability is determined is outside the scope of this paper. It could be partially genetic or entirely environmental. However, I do assume that by age 16 a student possesses either scholastic advantages or disadvantages which will affect his probability of successfully completing college. Students with greater scholastic ability will have a greater chance of graduating college, but the exact relationship between scholastic ability and the probability of graduation may be complex and unknown.

I suggest that a priori students do not know the exact relationship between ability and the probability of graduation. They must back out this relationship by observing the world around them. The accuracy of the student's information can affect the choices he makes. In particular, I argue that residential segregation based on income systematically biases the information a student has and lead him to make decisions he would not have made if he had possessed all of the facts.

Imagine, once again that a student knows he will earn an extra $250,000 if he or she graduates from college. To prepare for college he must give up two years of part-time work or roughly $12,500. If the student believes that his chances of graduating college are 5% then his expected benefits will be at least (.05)*($250,000) = $12,500. In other words, a student who believes he has a 5% chance of graduating college if he prepares will find that the expected benefits are exactly equal to the cost. A student who believes that he has greater than a 5% chance will find that the expected benefits of preparing for college outweigh the costs. Conversely, a student who believes that he has less than a 5% chance of graduating college will find that the benefits are less than the cost. In this way, the student's estimate of his chances of completing college become crucial in determining whether or not he prepares.

A student's community provides information that will help him determine the probability that he will graduate from college if he prepares. He knows that his scholastic ability is a factor. However, he does know the exact relationship between scholastic ability and probability of success. The student, however, can observe the adults in his community. He knows whether or not they have graduated from college and he can likely deduce how their scholastic ability compares to his own. Using this information he can get an idea of how likely he is to graduate.

As long as the adults in the community are a random sample of the population as a whole this method is relatively effective. Each generation may not get the relationship exactly right but it will produce a better estimate of the relationship than the proceeding generation. After a few generations have passed the community will accurately interpret the relationship between scholastic ability and probability of graduation.

A problem is created, however, if the community is not a random draw from the population. If the community is created in such a way that college graduates have a disproportionate chance of either being included or excluded from the community then the information available to students will be biased. Students will systematically draw the wrong conclusions. In particular, if college graduates are unlikely to return to a community then students will underestimate their probability of graduating college. This will lead to under investments in college preparation and lower graduation rates. In fact, it is possible for the cycle to feed on itself until a community has no college graduates at all, and no students or only the students with the very highest aptitudes will prepare for college.

Since college graduates earn more on average than non-college graduates residential segregation on the basis of income will bias the distribution of graduates. Economically disadvantages communities will tend to have fewer college graduates and exhibit the effects described above. In fact, the greater the gap in earnings between college graduates and non-graduates the stronger the residential segregation will be and the more biased the information available to economically disadvantaged students will be. This can produce interesting supply side dynamics that I will return to later.

Student’s would be able to mitigate this effect if other sources of information available to them beside relative comparisons to others in their community. Absolute measures of scholastic ability, however, may be hard to come by. Grades are in part a relative measure of performance. Standardized test scores are available but such tests often measure accumulated knowledge rather that raw scholastic ability. Students may take the scholastic aptitude test, however, the relationship between the SAT and college graduation is debated even among experts.

Perhaps, more to the point accessing and analyzing this information may prove challenging to a high school student. Parents and teachers can offer advice about how to analyze this information and which sources to trust, however, talk is literally cheap and it is difficult to know which sources of information to trust.

It is a teacher’s job to help a student prepare for higher education and thus a student might expect a teacher to at least give lip service to the advantages of preparation even when the expected benefits are low. Parents might encourage a student to prepare for college in an effort to raise the family’s status, to fulfill a dream the parents have or simply because parents may place a low value on the sacrifices the student must make.

Perhaps most importantly, a student may discount what a parent says because he can observe what the parent does. In the end, actions speak louder than words. If someone is telling you that you should make a significant investment in education yet you observe that this person and this person's peers have not made that investment it would be reasonable to question the credibility of that information.

Moreover, many parents from economically distressed communities may not pressure their children to prepare for college because the parents *themselves* have low estimates of the return to education. If the only person encouraging you to prepare is the person who is paid to prepare you then there are strong reasons to consider that individual's information dubious at best. Who is the student to believe, his teacher or his lying eyes.

Unusually persuasive teachers may be able to overcome this trust gap and alter student expectations. I will return to this point later. Note, however, that this implies that it is a teacher's ability to motivate economically disadvantaged students, not instruct them that makes the difference.

The proposed link between residential segregation based on income and college preparation can cause some counter-intuitive results in the labor market. For example, if the gap between the income of college graduates and non-college graduates rises then there is a greater benefit to completing college. We would expect that this would lead more students to prepare for college. If, however, this rising gap leads to increased segregation between college and non-college graduates the number of students preparing for college could stagnate or even decline.

Card and Lemieux show that the rapid increase in the earnings gap between college and non-college workers beginning in the 1980s could be explained by steady growth in the demand for college educated workers since World War II but declining growth in supply since the mid 1970s. Increases in residential segregation based on income could serve as part of the explanation.

BACKGROUND TRENDS

Since 1975 the gap between the wages of college educated workers and non-college educated workers has grown substantially. However, the number of men who graduate from college has changed very little. The number of female college graduates has been growing; however, it is difficult to determine how much of that is due to social rather than economic forces. Economically disadvantaged and minority men seem have particularly muted response to the increase in the earnings gap.

Between 1974 and 2000 the earnings ratio between college and non-college educated American men grew from 1.5 to 2.0, an increase of 33%.

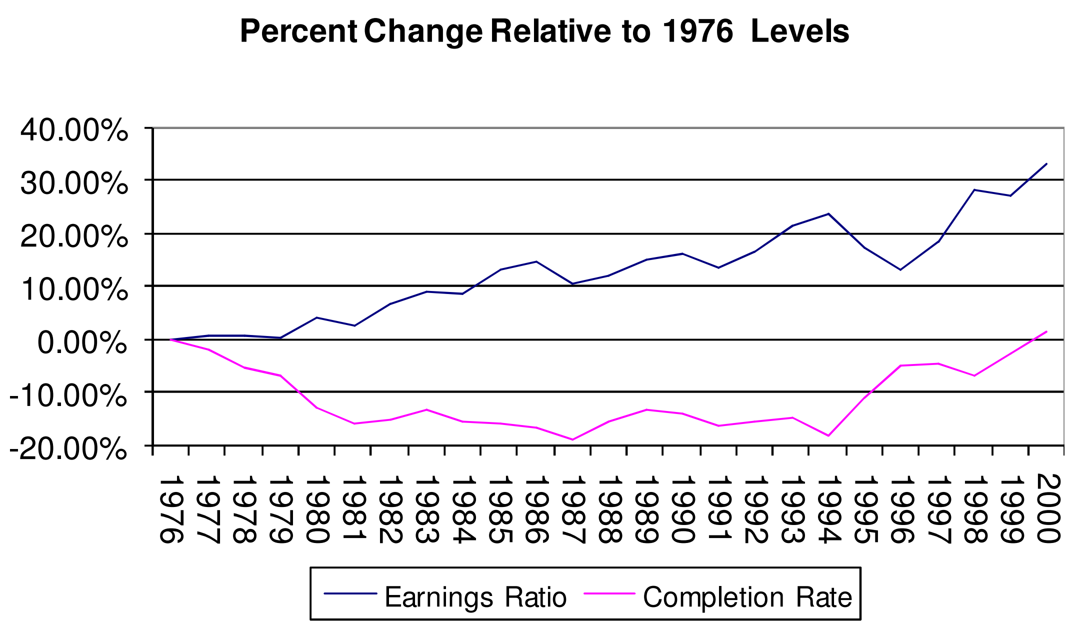

Figure 1 **Change in Earnings and Attendance**

The leading explanation for the increase in the earnings ratio is an increased demand for skilled workers driven by computer related technological change. For this paper, however, the more interesting question is why the supply response has been so small.

One might imagine that there was a marginal American man in 1976 for whom the increased earnings from a college education just failed to overwhelm the costs of tuition and four years of lost pay and work experience. In the 1990s that same man should have chosen to enter college but it seems he did not. The completion rate by the end of the century was nearly identical to 1976 levels.

As Card and Lemieux point out, changes in earnings and enrollment are actually negatively correlated over much of the period and the effect is even more dramatic when earnings are restricted to workers under 30. Young workers without a degree saw their real wages fall slightly while young college educated workers experienced larger wage increases than their older counter parts.

Card and Lemieux suggest that the slow growth of male college completion since the 1960s rather than computer related technological change is the driving force behind the increase in earnings inequality. While this is a strong possibility the fact I want to focus on here is that male college completion, was at a minimum, slow to respond to earnings growth over most of the period.

Of particular interest is that among poorer students college attendance is lower and seemingly slower to respond to changes in the earnings ratio. The chart below is used often by James Heckman and is taken from Heckman and Carnerio.

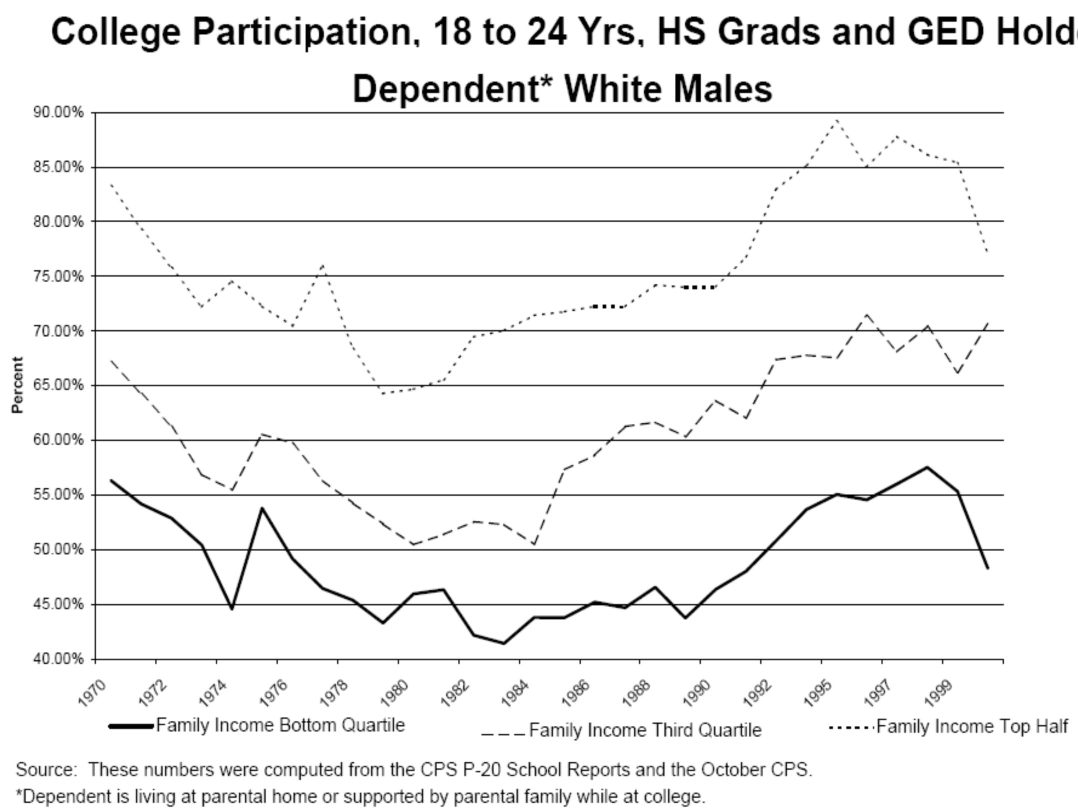

Figure 2 **College Participation**

The returns to college began to increase in the early 1980s and students from upper income families responded almost immediately. Students from the third quartile did not respond until the mid-80s and students from the lowest quartile did not respond until the 90s.

Racial groups show a similar pattern of levels and response timing. Blacks and Hispanics have both a lower level of college participation and are slower to respond to increases in the wage premium. The following chart, also taken from Heckman and Carnerio, illustrates that black and Hispanic college enrollment did not respond to the increase in the earnings premium until the late 80s and early 90s respectively.

Both charts, however, are likely to overstate the response of schooling to changes in the earnings ratio. The CPS enrollment data combines two and four year college enrollment and makes no adjustment for how long it takes students to finish college.

During the early 90s when minority and poor white college enrollment was growing the fastest, the percentage of students in rolled in two year colleges was at its height. A portion of the enrollment response is thus likely do to an increase in two year rather than four year college enrollments. In addition, Bound, Lovenheim and Turner report that for students in the 1972 National Longitudinal Survey, 51.5% of those who would go on to earn a baccalaureate degree from a public institution did so by their fourth year out of high school. In the 1988 National Education Longitudinal Survey that percentage had dropped to 34.1%.

Therefore, some of the increase in enrollment during the early 90s represents a greater fraction of two year college students and an increasing number of four year college students who took longer than four years to graduate.

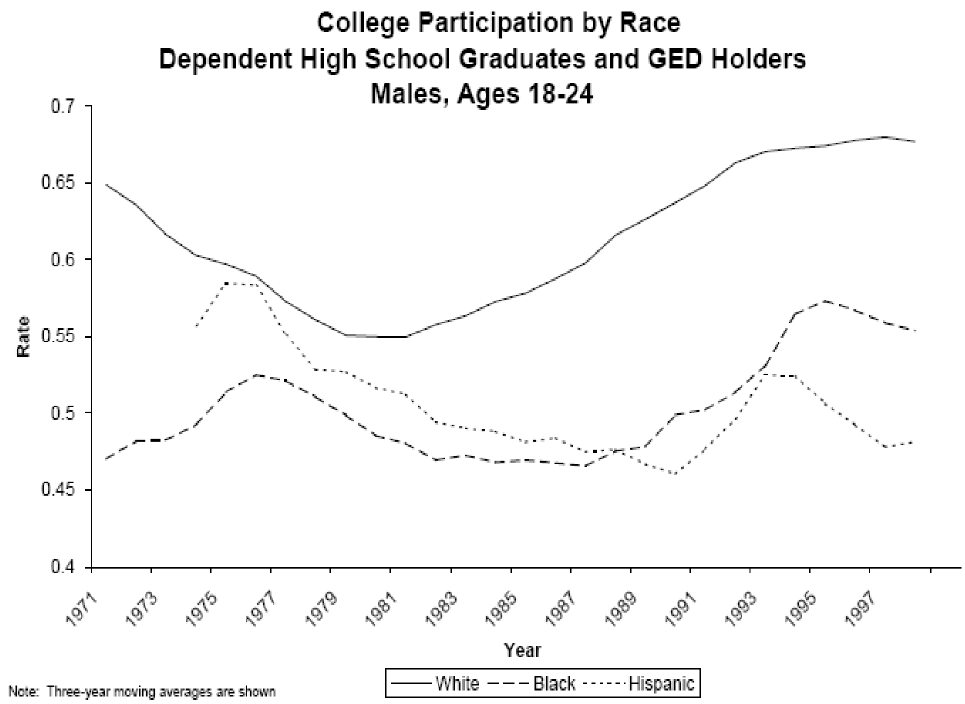

Figure 3 **College Participation by Race**

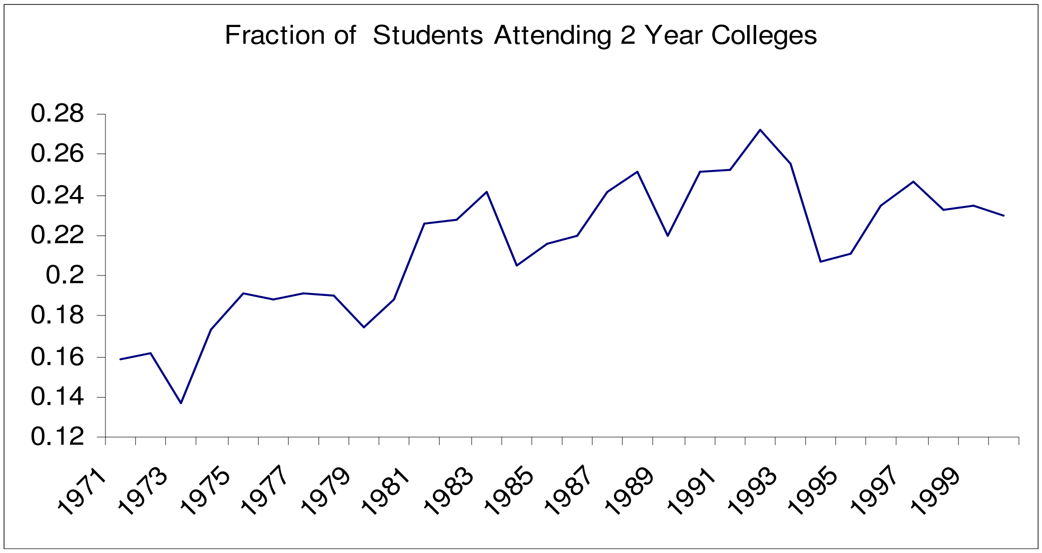

Figure 4 **Two Year College Attendance**

Poor students are more likely to live in the same communities as other poor students and minorities are more likely to live in the same communities as other minorities. The trends in college attendance, however, could simply be a result of income and race effects rather than community induced priors.

Mayer 2000, however, estimated the effect of state level income variance between census tracts and with-in census tracts on educational outcomes.[1] Geographic neighborhoods are not quite the same as communities. Communities exist within social space and have temporal lags with respect to geography. Individuals may interact frequently with friends or relatives who live across town and moving to a different geographic neighborhood does not necessarily produce immediate changes in the social community he or she

[1] For example, Illinois had a lot of variance in income between its census tracts. What effect does that have on the educational outcomes of a child growing up in Illinois.

interacts with. There is, nonetheless, a strong correlation between geographic neighborhoods and social communities. To the extent that an increase in earnings variance manifests itself as an increase in the variance of earnings between neighborhoods, communities' expectations should diverge. Conversely increases in the variance of earnings inside of neighborhoods should cause less, if any, divergence in community expectations.

Mayer's findings support this hypothesis. In her most complete specification, which includes family income and race, a one standard deviation increase in between tract income variance yields an estimated increase of .730 years of schooling for students in the upper half of the income distribution and a decrease of .499 years of schooling for students in the lower half of the income distribution for a total divergence on 1.229 years in schooling. On the other hand a one standard deviation increase in within tract variance results in an estimated increase of .418 years schooling for students in the upper half of the distribution and no response for students in the lower half of the distribution. Thus, increases in socioeconomic stratification between neighborhoods seem to encourage schooling among wealthier students and discourage schooling among poorer students.

There is no way to tell if the distribution of college graduates and the returns to education are what induce the differences in income variance in Mayer's study. Nonetheless, the response pattern of income groups to changes in between-tract inequality is similar to the response pattern of income and racial groups to changes in the earnings ratio and this raises suspicion they could have a common root.

COMPETING EXPLANATIONS FOR LOW COLLEGE GOING RATES

There are two common explanations for low college going rates among economically disadvantaged students. The first is that students from less affluent communities are not well prepared for college because of poorly funded primary and secondary schools. The second is that students simply lack the resources to pay for a college education

Inadequate funding for primary schooling is the most common explanation for the lower college going rates in economically distressed communities. If those students go to elementary and middle schools that are lacking resources then they may be unprepared for high school and therefore *correctly* estimate that their chances of getting into and succeeding in college are low.

Hanushek (JEL, 1986) surveys the literature and provides meta-analysis on K12 performance. Of particular interest is the fact that per pupil expenditures on current services expanded in real terms by 135% between 1960 and 1983. However, measures of pupil performance declined over that period.

Some of this effect may be due to rising costs. Increasing productivity in the non-education sectors tends to drive up labor costs in education even when the level of service remains constant. The level of service, however, was not constant over the period. Hanushek finds that teacher to student ratios actually fell by 26% from 1960 to 1980.

Nonetheless scores on Iowa's Tests of Educational Development peaked in the late 60s and declined precipitously into the 1980s. Nationally, the fraction of graduating seniors going on to college rose from 32.8% in 1960 to 46.1% in 1970 but has remained roughly constant since. The time series data seem to suggest that resources-per-pupil is not the dominant factor in determining educational performance.

Hanushek examines 147 studies of school performance based on instructional inputs. Sixty-five of those studies included per pupil expenditures. In only sixteen where expenditures significant and in three of those the effect was negative. In only 13 (20%), did expenditure have a significant positive effect.

To confound matters it is possible that in these few incidences the effect runs the other way. Suppose that students in less affluent communities do worse because they have worse expectations. The studies do not include an independent measure of income. Therefore, to the extent that economically disadvantaged communities have lower funding the expectations effect could be attributed to school funding.

An empirical regularity mentioned in Hanushek and supported with extensive panel data in Rivkin, Hanushek and Kain (Econometrica, 2005), is that while per pupil expenditure seems to have little effect on student performance, unobserved differences in teacher quality have large effects. This is consistent with the general notion that expectations are important.

Though not explicitly modeled here, it is conceivable that certain teachers are able to gain the trust of their students. Information that these teachers provide information will then be incorporated in a student's expectations. Teachers who are able to convince students that effort will be rewarded with a high probability will raise academic performance. Thus a teacher's ability to gain the trust of students may serve as the source for unobserved differences in ability.

The second explanation depends on the concept of liquidity constraints. Liquidity constraints imply that students are not able to borrow enough money to finance a college education. If young students could borrow all of the money that they needed then they could repay the loans with their increased earnings and still be better off. The evidence for liquidity constraints, however, is weak.

James. J. Heckman argues that cognitive skills become crystallized around the age 14. In particular, Heckman and Carnerio (2000), presents evidence that family income during the final years of high school have no effect on college going rates. Instead, family income during the primary school years seems to be a determining factor. If students were liquidity constrained then one would expect that losses in income leading up to high school graduation would lead students to at least defer college going. However, this does not seem to be the case. Either lower family income during primary school sets into motion some a chain of events that prevents college going or it serves as a proxy for some deeper effect.

One of the most impressive analyses of college going choices comes from Keane and Wolpin. The estimate a dynamic programming model of career choices based on data from the National Longitudinal Survey of Youth (NLSY). There model follows the choice of young men to go to school, work in white-collar occupation, work in blue-collar occupations, serve in the military or stay out of the labor force. Their model accounts for skill and knowledge accumulation and allows for differing preferences between the young men.

Keane and Wolpin find that 90 percent of the variance in career decision is determined by and young man's type laid down by age 16. That is, something happens to young men by age 16 which will largely guide their career paths for the rest of their lives. There are four possible types. Type one individuals tend to be the most educated averaging roughly 16 years of education and work almost exclusively in white collar occupations. Type two and three individuals work mostly in blue collar occupations though type threes spend some time in the military. Roughly half of type two and three individuals finish high school and a small fraction goes on to post-secondary education. Type four individuals are actually more likely to finish high school yet spend much of their post school time at home. These types are not determined by the researchers but are chosen by the computer after analyzing patterns in the data.

Most interestingly, types are heavily correlated with parental education and income. The computer simulation determines types by analyzing patterns in the young men's behavior. However, when the researchers compare those types parental education and income they find that highly educated and affluent parents are much more likely to have type one children. Lower income and less educated parents were much more likely to have type three children. The parent of type two individuals had similar education profiles as type the parents of type three individuals but tended to be slightly higher income. Interestingly, type four individuals were likely to have parents that were more educated but with lower incomes.

Keane and Wolpin interpret these types as skills the child was either born with or developed before age 16. However, I offer an alternative explanation. A young man's type may be influenced by the information he gathers from his community. The type then influences behavior not because of differences in skill but in expectations. In particular different type may have different estimations about the return to preparing for college.

Rather than setting the student's type, I set the student's estimation of college's "type." If college's is extremely difficult then preparation will not pay off for students who do not have the highest level of scholastic ability. However, if college is only moderately difficult then preparation can make the difference even for students with average levels of scholastic ability.

In the language of game theory, an unseen move by nature makes college either extremely or moderately difficult and the student has to form priors about the probability of each. These priors determine the structure of the game

Two students faced with the same objective constraints can end up playing two entirely different games if they have different priors. For one student college is a distant possibility and his high school years are better spent developing practical experience. Another student with the same native talent sees college prep as a no-brainer. A little hard work can ensure success in college and the returns to a degree outstrip anything that he could learn on the job[2].

Without accounting for the difference role models and peers have in shaping these two student's perspectives it will appear that they differ on some unobserved personal characteristic that is correlated with family experiences, neighborhood residence,

---

[2] Levitt makes an implicit point about priors and the crack epidemic in the late 1980s and early 1990s. Young gang members assumed that they would move up in the hierarchy just as older gang members had done before them. Given the incomes available to gang leaders this promised to be a good investment. Yet, the experiences of their role models were not an accurate guide for what would happen to them. The incentive structure had changed and the old gang members had no intention of turning over the reigns. This lead to unfortunate choices on the part many inner-city youth and a collapse in social wellbeing.

ethnicity, etc. In fact, this characteristic is their assumptions about the world gleaned from family experiences, neighborhood residence and ethnicity.

Heckman and Canerio argue that these assumptions develop early on and crystallize into non-cognitive skills that impact the student thoroughout his or her educational career. The model presented here deals with the priors of a young high school student deciding whether or not to exert effort during the final years of public school. A more general model, however, would also incorporate parental priors, which influence how children are raised and early childhood priors about the value of education and cognitive performance in general.

Children and parents whose vision of the world places a premium on cognitive performance will investment more time and energy in honing mental ability. Baby Mozart tapes and other activities designed to give an infant an edge in the inevitable college entrance exams are obvious examples but the effect can be more general. A community which views cognitive activity as a gateway to economic success will prize signs of mental agility in children and work to accentuate them.

Priors that educational success is unlikely may show up as some of the non-cognitive skill deficiencies pointed out by Heckman and Canerio. Heckman and Canerio point out that low performing students often suffer from poor motivation and behavior problems that can be alleviated by programs such as the Boy's and Girl's club. The express purpose of the Boy's and Girl's club is to provide positive role models and the reported effect is a change in what young people see as their opportunity set.

The central thesis here, however, is that the expectations young people have are not the arbitrary result of exogenous culture but priors developed from a rational interpretation of community behavior. Priors influence community behavior and community behavior shapes the priors of the next generation. Through this process community behavior supports an intergenerational transfer of information about the state of nature. The result is a system in which exogenous shocks can have an echo effect on community choices for generations. The persistence of these shocks depends on the extent to which individuals from different community interact and whether the resulting behavior insulates or exposes individuals to new information. [3]

Out of this interaction comes an economic return to diversity. In an isolated community an exogenous information shock can create priors that make members afraid to engage in some activity. Because members are reluctant to engage in the activity new information about the consequences of the activity cannot be obtained. Priors are updated very slowly and the community suffers from misinformation for an extended period of time.

[3] An isolated community that is exposed to a discrimination shock that prevents its members from attending college could have persistently low estimations of the return to early cognitive development. It is possible that this effect could have been present in some black communities in the United States. Early discrimination has an echo effect of depressing investment in human capital. I discuss later on how unfortunate integration strategies could actually exacerbate this effect by removing positive role models from the community.

If on the other hand the community is not isolated then it will trade members with communities that have different priors. The behavior of the new members will reflect priors that were not impacted by the negative shock. After observing this, members of the original community will community will be more likely to engage in the frightful activity, will learn from the experience, and will have their information updated accordingly.[4]

[4] A cute anecdote relating to this is the reluctance of Italians to eat tomatoes when they were first introduced to the Old World, in 1530. According to legend it was only after a food shortage in Naples that some peasants discovered that tomatoes, unlike other members of the nightshade family, were not poisonous. It took nearly 300 years for the anti-tomato prior to be overturned and the utility of Italian families doubtlessly suffered throughout the period.

ANALYTICAL MODEL

The focus of this paper, however, is on the decision to study hard in high school. This choice has serious consequences either way. If a student works hard he could be rewarded with acceptance to college and eventually a higher paying job. He must, however, give up part of his youth. He must forgo earnings today or, perhaps more importantly, time spent with friends, in exchange for the possibility of discounted future earnings. Many teenagers choose not to work at all during high school, so the wage may be a lower bound on the value of leisure time.

Loss of leisure applies both in and out of the classroom. Clearly, time spent doing homework is time that cannot be spent with friends. In the classroom, however, time spent paying attention to the lesson is time that cannot be spent goofing off, napping or engaging in side conversations. Attentive students are taking a big risk – daydreaming pays off with probability one.

Making matters worse, the student doesn't know for certain what his chances of getting and finishing college are. To some extent it depends on his scholastic aptitude. To some extent it depends on the difficulty of college.

He may have a sense for how academically talented he is relative to the people around him but to form a belief about his chances of success in college he needs to have some sense of how likely those around him are to get into and complete college themselves.

He is facing a game of incomplete information. He knows his own ability, but he doesn't know the underlying difficulty of college. About that the student has to form priors. Those priors will determine what the game looks like to him. This is the crucial step.

Two students with exactly the same ability facing the exact same objective constraints but with two different sets of priors will be playing two entirely different games. To one the idea of going on to college, even graduate school, may be a no-brainer, the obvious choice for someone with his ability. To the other it is a pipe dream. In the "real world" his analytical aptitude means that he will make an excellent head cashier.

Neighborhood observations help set each student's priors. He looks at the success rates in his community. If those success rates indicate that someone with his level of talent is likely to get into and graduate from college he will assume that the difficulty of college is relatively low – that college for a smart young man is no-brainer. If the success rates tilt the other way, He will assume that college is relatively hard and that his time is better spent on leisure or part time work than studying.

Analytical Model

In the analytical model a community has 2N residents. Half of the residents are adults and the other half are students. Residents live for two periods. In the first period they are students and must decide whether or not to exert effort in high school. In the second period they graduate from college or not and become adults.

Each agent falls into a family indexed $i \in \{1...N\}$. The agents are uniquely identified by family and the period in which they was born. At the beginning of each period, each student must decide whether or not he will study hard in high school. If he chooses to study hard, he pays a cost **C** and receives a probability of getting into and graduating from college determined by

(0.1) $$S_{i,t} = 1 - e^{\frac{-A_{i,t}}{D}}$$

S is the probability that the agent will graduate from college if she studies hard in high school. $\mathbf{A_{i,t}}$ is her aptitude, assumed to be set exogenously by nature or nurture by the end of primary education. **D** is the underlying difficulty of college. D, and hence S, is unknown to the agents. They must back it out from observations of aptitudes and college graduation rates.

If the agent successfully finishes college then she receives a benefit **B**. Therefore her expected benefit from studying hard is

$$(0.2)\ E\left[B_{i,t}\right] = B\left(1 - e^{\frac{-A_{i,t}}{D}}\right)$$

and so the agent will attempt college if[5]

$$(0.3)\ C < B\left(1 - e^{\frac{-A_{i,t}}{D}}\right)$$

which implies that he will attempt college so long as

$$(0.4)\ D < \frac{-A_{i,t}}{\ln\left(\frac{B-C}{B}\right)}$$

Which allows us to define

$$(0.5)\ A^* = -D\left(\ln\left(\frac{B-C}{B}\right)\right)$$

[5] At this point we are assuming risk neutrality, so that B and C can be seen conversely in monetary or utility terms.

Where **A*** is the threshold level of aptitude necessary to make studying hard worthwhile. Unfortunately, the agents do not know D. However, the aptitudes of the agents within the neighborhood are known as well as the success rate of those in the neighborhood.

Inverting the above equation for A* and plugging into the equation for S and averaging over the entire community yields

$$(0.6)\quad S(A^* \mid A_t, B, C) = \frac{\sum_{i=1}^{N}\left(1 - e^{\frac{A_{i,t}}{A^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)\delta(i)}{N}\left|\begin{array}{l}\delta(i) = 1 \Leftrightarrow A_{i,t} > A^* \\ \delta(i) = 0 \Leftrightarrow A_{i,t} \le A^*\end{array}\right.$$

Where S(A* | A,B,C) gives the success rate of the population that is consistent with A*, that is the fraction of the entire population that could be expected to succeed if the if the threshold were A*.

Lemma One: For A* > 1 and B>C>0 , S(A*|A,B,C) is monotonically decreasing in A*

Suppose not. Then there exists $A_1^* > A_2^*$ such that $S(A_1^*) > S(A_2^*)$. Let Q1 be the set of all i such that $A_{i,t} > A_1^*$ and let Q2 be the set of all i such that $A_1^* > A_{i,t} > A_2^*$. Then

$$(1.1)\quad S(A_1^*) = \frac{\sum^{Q1}\left(1 - e^{\frac{A_{i,t}}{A_1^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}$$

and

$$(1.2)\quad S(A_2^*) = \frac{\sum^{Q1}\left(1 - e^{\frac{A_{i,t}}{A_2^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right) + \sum^{Q2}\left(1 - e^{\frac{A_{i,t}}{A_2^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}$$

and $S(A_1^*) > S(A_2^*)$

Now, either Q2 is the empty set or it is not. If Q2 is the empty set then S(A*2) becomes

$$(1.3)\quad S(A*2) = \frac{\sum^{Q1}\left(1 - e^{\frac{A_{i,t}}{A*2}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}$$

Thus

$$(1.4)\quad S(A_2^*) - S(A_1^*) = \frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{A_2^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N} - \frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{A_1^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}$$

$$(1.5)\quad S(A_2^*) - S(A_1^*) = \frac{\sum^{Q1}\left(e^{\frac{A_{i,t}}{A^*1}\left(\ln\left(\frac{B-C}{B}\right)\right)} - e^{\frac{A_{i,t}}{A^*2}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}$$

$$(1.6)\quad S(A_2^*) - S(A_1^*) = \frac{\sum^{Q1}\left(\left(e^{A_{i,t}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)^{\frac{1}{A_1^*}} - \left(e^{A_{i,t}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)^{\frac{1}{A_2^*}}\right)}{N}$$

Now, the numerator contains a summation. Each term in that summation is the difference of two identical terms raised to a power. I will show that the base terms are fractional, the powers are positive and that the power of the first term is smaller than the power of the second.

$\frac{B-C}{B}$ is certainly fractional. Hence, its natural log is negative. By assumption $A_{i,t}$ is positive. Hence the base term is *e* raised to a negative power. The base term is therefore fractional. By assumption A_1*and A_2* are positive. The powers are the reciprocals of $1/A_1^*$ and $1/A_2^*$ respectively. However, by assumption $A_1^* > A_2^*$ hence $1/A_1^* < 1/A_2^*$.

Thus the difference of each pair of terms in the summation is positive. Hence, the summation is positive. The denominator is the sum of 1s and so is positive. Therefore the S(A*2) – S(A*1) is positive and hence S(A*2) > S(A*1).

Now suppose that Q2 is non-empty. Then

$$(1.7)\quad S(A*2)-S(A*1)=\frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{A*2}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)+\sum^{Q2}\left(1-e^{\frac{A_{i,t}}{A_2^*}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}-\frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{A*1}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}$$

However, (1.8) $$\frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{A^{*}2}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)+\sum^{Q2}\left(1-e^{\frac{A_{i,t}}{A_2^{*}}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}>\frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{A_2^{*}}\left(\ln\left(\frac{B-C}{B}\right)\right)}\right)}{N}\ \forall A_{i,t}>0$$

So by the above $S(A_2^*) > S(A_1^*)$. Thus S(A) is strictly decreasing in A.

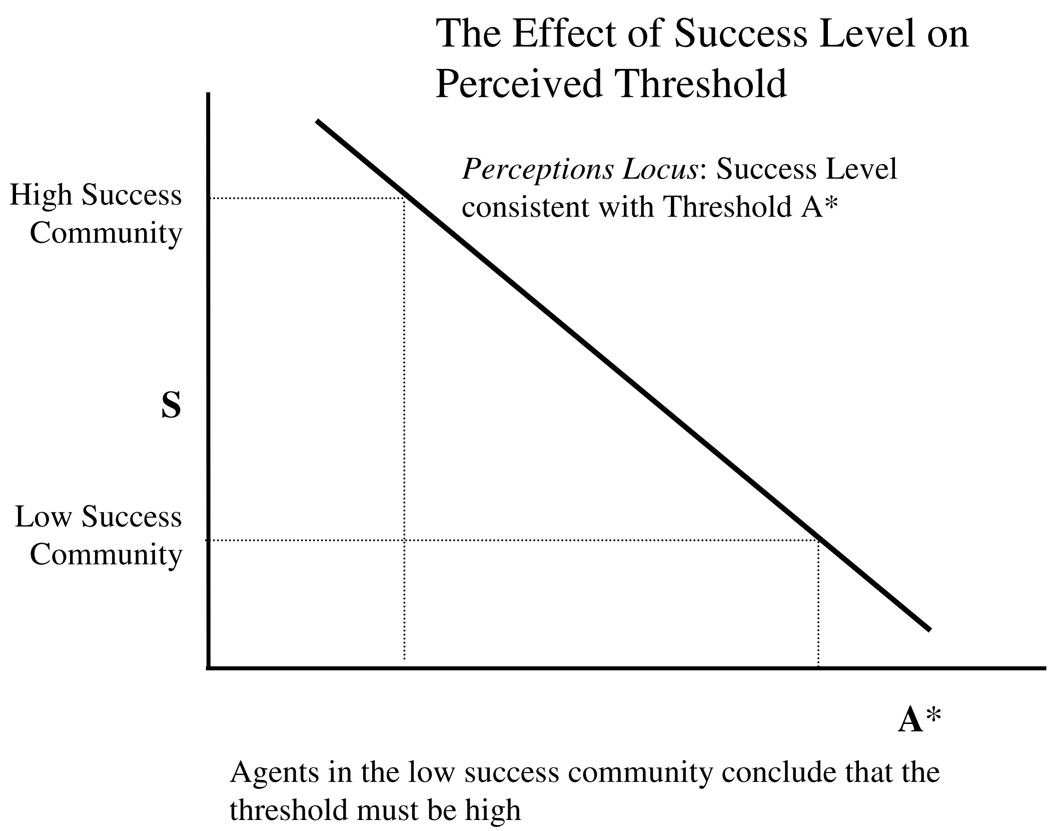

Figure 5 Perceptions Locus

Theorem: $\hat{A}^* = S^{-1}(S|A, B, C)$ ***is strictly decreasing in S***

Proof: By Lemma One $S^{-1}(S \mid A,B,C)$ exists and gives $\hat{A}*$, the communities estimate of A*, as a function of S. It follows immediately that $S^{-1}(S)$ is strictly decreasing in S.

This is the centerpiece of the argument. Low levels of community success are consistent with a high A* and therefore a high difficulty of college. A student backing out the difficulty of college from looking at the success level in his neighborhood will tend to conclude that college is more difficult if fewer people have completed it.

It's important to note that S^{-1} is conditional upon the vector A of aptitudes in the community. Two communities with the same success rate will in general not have exactly the same $\hat{A}*$. Holding S constant A` > A`` will have a higher $\hat{A}*$.

That is, if every agent in the first community is at least as talented as every agent in the second community and one or more agents in the first community is more talented then the first community will have a higher $\hat{A}*$. If talented people are succeeding at the same rate as less talented people one would guess they face higher obstacles. Thus, if a community has more talented members but the same success rate student are going to back out a higher threshold.

I call S^{-1} the Perceptions Locus. It maps out what agents perceive to be the state of nature given the community composition.

How Communities Become Educated

If a community has a very low level of success then investments in education will be retarded. A natural question might be how communities wind up with very low levels of education.

All communities were at some point uneducated. So, perhaps the more relevant question is how some communities became educated at all. This requires looking at generational developments and the Consequence Locus.

The number of students who try and the actual difficulty of college determine the success level of each generation. The generation process for the students' success level is

$$(2.1)\ S^* = \frac{\sum_{i=1}^{Q1} \varphi(i)}{N} \left| \begin{array}{c} i \in Q1 \Leftrightarrow A_{i,t} \geq A^* \\ \\ \varphi(i) = 1 \Leftrightarrow \textit{Agent Graduated College} \\ \varphi(i) = 0 \ \textit{Otherwise} \end{array} \right.$$

and so S* is stochastic. However E[S*] =

$$(2.2)\ \mathrm{E[S^*]} = S(A^* \mid D) = \frac{\sum_{i=1}^{NK}\left(1-e^{\frac{-A_{i,t}}{D}}\right)\delta(i)}{N} \left| \begin{array}{l} \delta(i)=1 \Leftrightarrow A_{i,t} > A^* \\ \delta(i)=0 \Leftrightarrow A_{i,t} \le A^* \end{array} \right.$$

Lemma Two: E[S*] is monotonically decreasing in A*

Proof: Let Q1 be the set of all i such that Ai,t > A_1* and let Q2 be the set of all i such that A_1* > Ai,t > A_2*. Now suppose A1* > A2* then. Then

$$(2.3)\ S^*(A_1^*) = \frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{D}}\right)}{N}$$

and

$$(2.4)\ S^*(A_2^*) = \frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{D}}\right)+\sum^{Q2}\left(1-e^{\frac{A_{i,t}}{D}}\right)}{N}$$

Now, once again either Q2 is the empty set or it is not. If Q2 is the empty set then S*(A1*) = S*(A2*) and we are done. If it is not then

$$(2.5)\quad S^*(A_2^*) = S^*(A_1^*) + \frac{\sum^{Q1}\left(1-e^{\frac{A_{i,t}}{D}}\right)}{N}$$

But the second term on the right is positive and so $S^*(A_2^*) > S^*(A_1^*)$ Thus the S*(A*), the Consequences Locus, is monotonically decreasing in A*

One more result is necessary before we can create a graph showing the relationship between Perceptions and intergenerational Consequences.

Lemma Three: $\hat{A}^* > -D\ln\left(\frac{B-C}{B}\right) \Leftrightarrow$ S(A*) > E[S*]

$$(2.6)\ \hat{A}^* > -D\ln\left(\frac{B-C}{B}\right)$$

$$(2.7)\ \frac{\hat{A}^*}{\ln\left(\frac{B-C}{B}\right)} < -D$$

$$(2.8)\ \frac{\ln\left(\frac{B-C}{B}\right)}{\hat{A}^*} > \frac{-1}{D}$$

Then

$$(2.9)\ \frac{A_{i,t}\ln\left(\frac{B-C}{B}\right)}{\hat{A}*} > \frac{-A_{i,t}}{D}\ \forall A_{i,t} > 0$$

Hence

$$(2.10)\ 1-e^{\frac{A_{i,t}\ln\left(\frac{B-C}{B}\right)}{\hat{A}*}} < 1-e^{\frac{-A_{i,t}}{D}}\ \forall A_{i,t} > 0$$

And if its true for all Ai,t then it is true for any sum of Ai,t Thus,

$$\hat{A}* > -D\ln\left(\frac{B-C}{B}\right) \Rightarrow \mathrm{S(A^*)} < \mathrm{E[S^*]}$$

And by Symmetry

$$\hat{A}* > -D\ln\left(\frac{B-C}{B}\right) \Leftarrow \mathrm{S(A^*)} < \mathrm{E[S^*]}$$

The following results also follow by symmetry

$$\hat{A}* < -D\ln\left(\frac{B-C}{B}\right) \Leftrightarrow \mathrm{S(A^*)} > \mathrm{E[S^*]}$$

$$\hat{A}* = -D\ln\left(\frac{B-C}{B}\right) \Leftrightarrow \mathrm{S(A^*)} = \mathrm{E[S^*]}$$

Therefore, we have immediately

Corollary One: S(A*) and E[S*] cross in one location. This produces the chart below

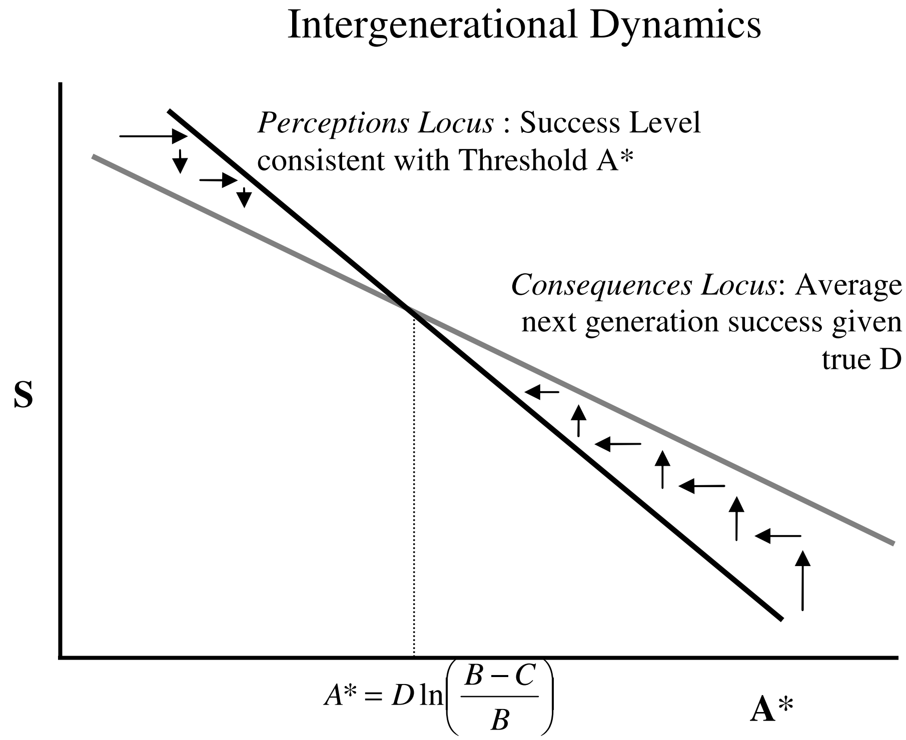

Figure 6 **Intergenerational Dynamics**

This leaves us with the very natural conclusion that if the members of a community guess that college is more difficult than it really is fewer will work hard in high school and thus fewer will go to higher education than would go under complete information. However, a higher fraction of those who do go will return successfully than the members would have predicted.

Conversely, if the community members think that college is easier than it is than it is, then too many will apply themselves in high school and too many will go on. However, fewer will finish successfully than they expected.

Over time the community perceptions will converge towards the true value. This is comforting both because it tends towards optimality and because it suggests that using ones community as a source of information is over time helpful.

A key insight here is that if there is variation in community action and there is no mechanism to bias the consequences a community will eventually converge on the truth. Over time someone will breakout and try something new, in this case going to college. That person will come back and influence the community. This in turn will encourage more people to try and community perceptions will gradually move towards the truth. Without parameterization, however, there is no way to tell how long this will take.

Bias, however, is a serious possibility in the community context. Community composition is endogenous and more importantly life-choices and hence perceptions influence community membership. If individuals select themselves into communities based on common beliefs then the resulting actions of a community will be systematically biased based on that belief and future generation will make incorrect inferences about the state of nature.

Ethnic prejudice may be example of beliefs about the state of nature that many readers are familiar with. Ethnic prejudice seems to be strongly influenced by the community in which one grew up. If a tight nit community believes that all people of a particular ethnicity share some trait then the children raised in the community are likely to believe that as well. This model of priors and endogenous community composition shows why these beliefs persist so easily.

Individuals sort into communities on the basis of ethnicity. This is particularly true when there are language barriers, as we might have imagined existed among immigrants to the US in the 19th century. Therefore, the sorting process reduces the experiences necessary to update prejudices. Even more to the point, individuals who, for some idiosyncratic reason, discard their prejudices are less likely to return to the community and affect the priors of the next generation.

Young people growing up in the community do not know the state of nature. They have no idea about the actual properties of ethnicities but they can deduce it from the behavior of the adults. Through the sorting process their expose to adults who do not share prejudice is limited and so the children rationally develop prejudice. Thus a system is created that reinforces previously negative priors about other ethnicities.

It is also immediately clear why negative false priors persist more strongly than positive ones. Positive false priors encourage community mixing where the priors are then updated. Negative false priors discourage it.

The sorting mechanism works similarly for beliefs about college. Adults who believe that college is too difficult are concentrated by sorting and so the environment in which the young students are raised is systematically biased. Thus, the young students rationally conclude that college is too difficult.

How do Some Communities Remain Uneducated?

The above analysis presumed that all college grads returned to the community from which the originated. However, suppose some fraction $1 - \alpha$ left the community. That is, after graduating from college they then joined a new community. For example, not all towns have colleges and graduates might tend to stay in the town the graduated in and become isolated from their original community.

In addition,n college graduates tend to have higher incomes than non grads and may move into a more affluent neighborhood. The motivations will be considered later but for now simply assume that each graduate moves to a new community with probability $1 - \alpha$.

If he moves to a new community then that graduates family line is replaced by a new non-graduate family line.[6]

$$(3.1)\ \mathrm{E}[\mathrm{S}^*|\ \mathrm{A}^*,\mathrm{D},\alpha] = S^*(A^*|D) = \frac{\sum_{i=1}^{NK}\left(1-e^{\frac{-A_{i,t}}{D}}\right)\delta(i)\alpha}{N}\left|\begin{matrix}\delta(i)=1 \Leftrightarrow A_{i,t} > A^* \\ \delta(i)=0 \Leftrightarrow A_{i,t} \le A^*\end{matrix}\right.$$

α passes through to yield:

$$(3.2)\ \mathrm{E}[\mathrm{S}^*\ |\ \mathrm{A}^*,\ \mathrm{D},\ \alpha] = \alpha S^*(A^*|D) = \alpha\mathrm{E}[\mathrm{S}^*|\mathrm{A}^*,\mathrm{D}]$$

[6] The assumption that the new family line is a non-graduate is arbitrary but simplifies the analytics immensely. We could alternatively assume that a graduate leaves the community with probability γ and is replaced by an agent who is a college graduate with probability η. Then $1 - \alpha = \gamma - \gamma\eta$.

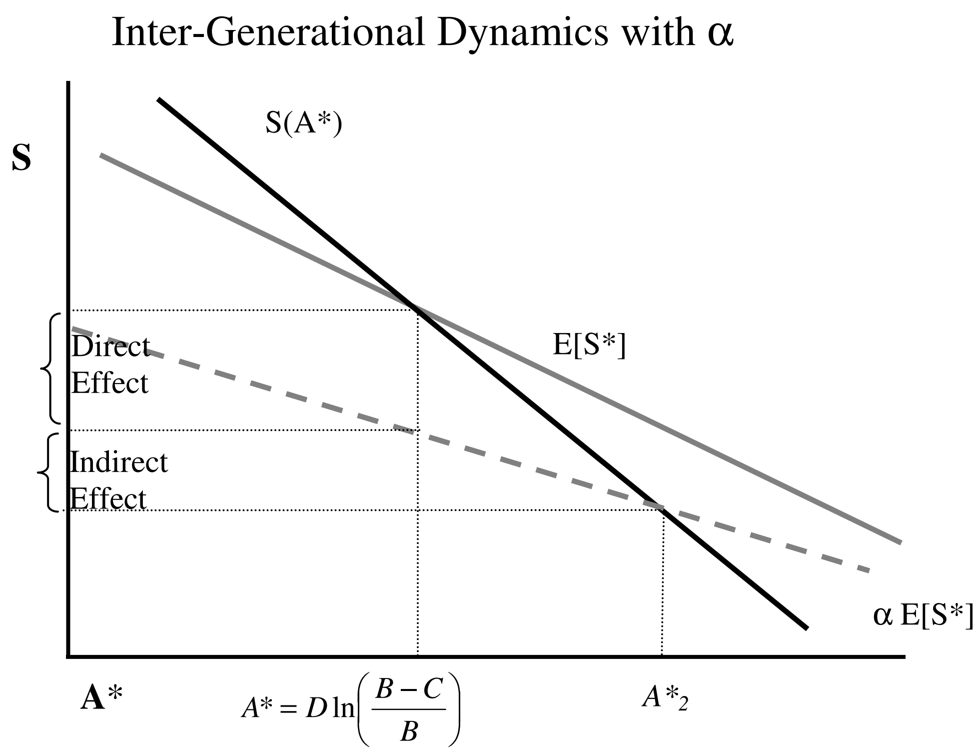

Figure 7 **Dynamics with Alpha**

Simply put the fraction of expected graduate is reduced by the factor α.

The success rate in the community, however, is reduced by more than α. Not only do fewer graduates return to the community, but the lower apparent success rate discourages the next generation from trying. Rather than settling down to the true value of A* the community settles on a higher value consistent with lower success.

In this simple model community membership is binary. You are either in or out. There is not possibility of being distantly connected or probabilistically connected. Therefore, when individuals leave they no longer effect the estimation of the new generation. This may not be that strong of an assumption.

We only have to assume that those who are partially disconnected from the community are worse sources of information than those who are fully engaged. In that case if some fraction β of graduates are partially disconnected, then there is an information loss of equivalent of some α totally disconnected members.

There are two effects of α. There is the direct loss of graduates because some graduates are ending up in another community. There is also the indirect loss of graduates because a lower apparent success level discourages young students from working hard.

Interestingly, the size of the indirect effect is inversely proportional to the slope of S(A*). If a wider range of A*s are consistent with a given range of success then the indirect effect is increased. In equilibrium, the community must underestimate the success of those who try by a factor of α. A shallow slope means that the tendency for agents to underestimate the probability of success is low.

Therefore, a very high threshold and low success rate is necessary to pull underestimation down by a factor of α. The better the agents are at surmising the true relationship between success and difficulty the more damaging alterations in the success rate are.

Shallow vs. Step Perceptions

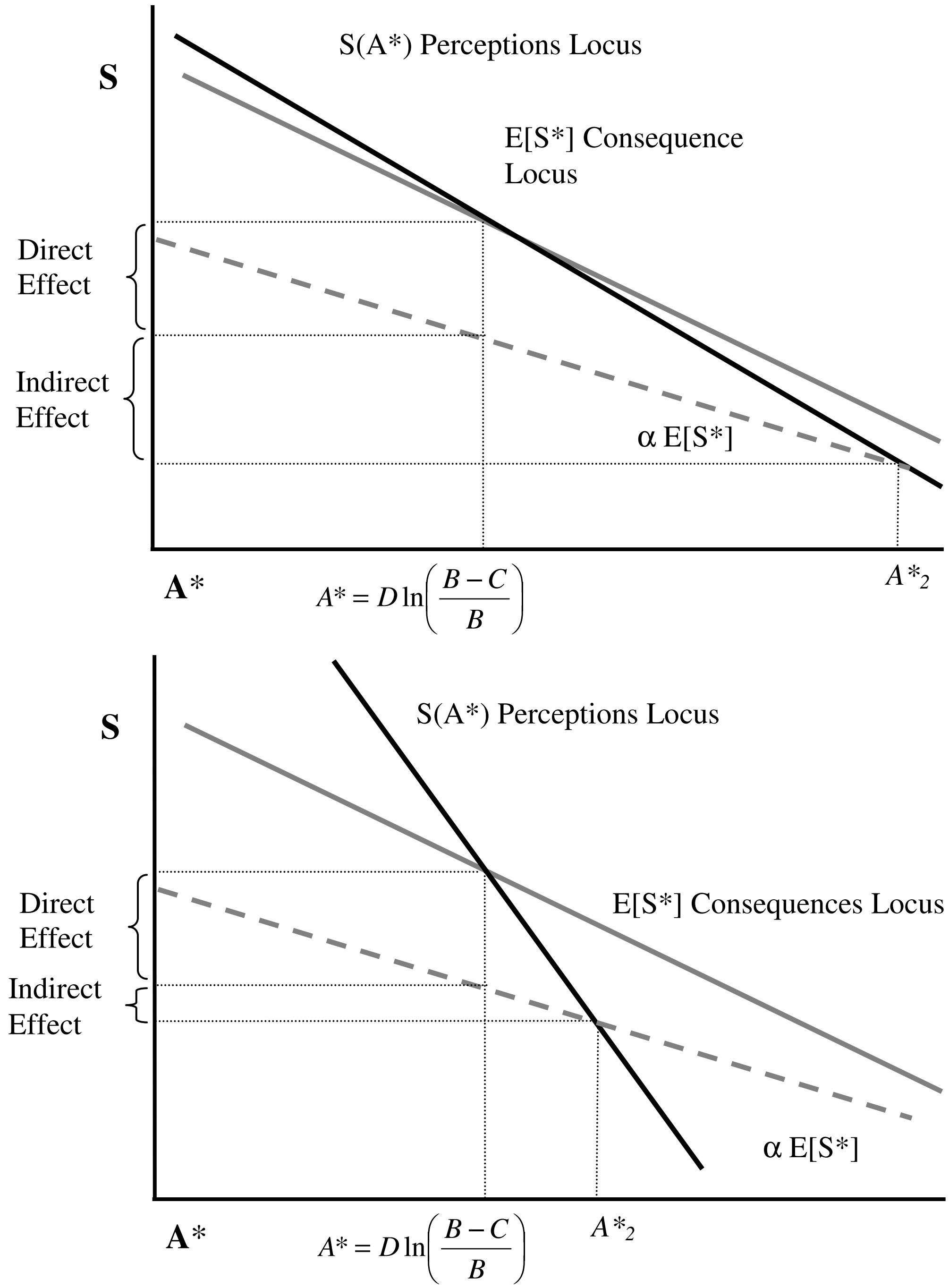

Figure 8 **Shallow vs. Step Perceptions**

Why Would Graduating Agents Not Return to the Community?

There are really two questions. Why would graduating agents not return and why don't graduated agents from other communities move in. If the community is defined by geography, that is all agents in a particular community tend to live in the same place, the answer could simply be locational amenities.

Some places have more amenities than others. Given a choice people would prefer to live in the nicer places. If there is a free market for housing then prices will allocate living spaces. In general nicer places will be more expensive and richer people will be more likely to live in nice places.

Therefore, even if all communities started out the same, those with more income would tend to migrate towards communities with more locational amenities and those with lesser income would move towards the communities with fewer. If college graduates have higher incomes then they would tend to move out of the community.

Wilson argued that things did not really become bad in the inner city until racial segregation began to end. Suppose that during segregation blacks were relegated to lower amenity areas, so that, successful blacks had no choice but to move back to the community. The community may have been disadvantaged but it was not suffering from a prior trap.

On the contrary, the fact that all blacks were forced into the same community reduced the probability of a prior trap. However, once restrictions were lifted, blacks who were more successful could move to higher amenity areas. The community's success rate would begin to fall and with it the effort level.

Another implication is that small changes in exogenous school quality could cause large changes in performance. If a school is even slightly better than its local competitors then parents and students interested in educational quality will bid higher to move into its district. If parents who are interested in educational quality tend to be educated themselves then this increases the educational level in the district and hence changes the perceptions of students in the district.

The Truly Disadvantaged

It is further possible that parents realize the empirical relationship between community success and the probability of their own child's success. It is also possible that all else being held equal parents would prefer their children to graduate from college. Then α itself could be a function of the communities' success rate. That is individuals would be less likely to return to a neighborhood with a low success rate than to a neighborhood with a high success rate.

So suppose that

$$\alpha_i = \alpha(S) \text{ and } {d\alpha_i}/{dS} < 0$$

Then $\alpha(S)E[S]$ crosses S(A*) in one location, two locations or no locations.

Consider in particular the extreme case of a community where $\alpha(0) = 0$. That is, no graduates wish to return to this community if there were no other graduates. The implication is that this is a community with very low amenities, perhaps isolated and otherwise unpleasant rural areas or the most rundown of urban communities.

Now, consider first the case in which $\alpha(S)E[S]$ crosses S(A*) in one location. Because $\alpha(S)E[S]$ crosses S(A*) in only one location the equilibrium for this community must be zero. That is, no matter what position the community starts in, over time it will move to a point where the number of college graduates is zero.

At this point the fate of students in this community becomes sensitive to assumptions about what happens in an information vacuum. That is, without any college graduating role models at all, what are your priors about college? One, perhaps hopeful assumption, is that with no graduates A_t is simply equal to the maximum At from the previous generation.

That is, since no in the previous generation is a college grad one can expect to graduate if his or her aptitude is higher than anyone from the previous generation. This will happen from time to time and so even very disadvantaged communities will occasionally produce grads and those graduates will be superstars. However, none of them will return and the community will remain disadvantaged.

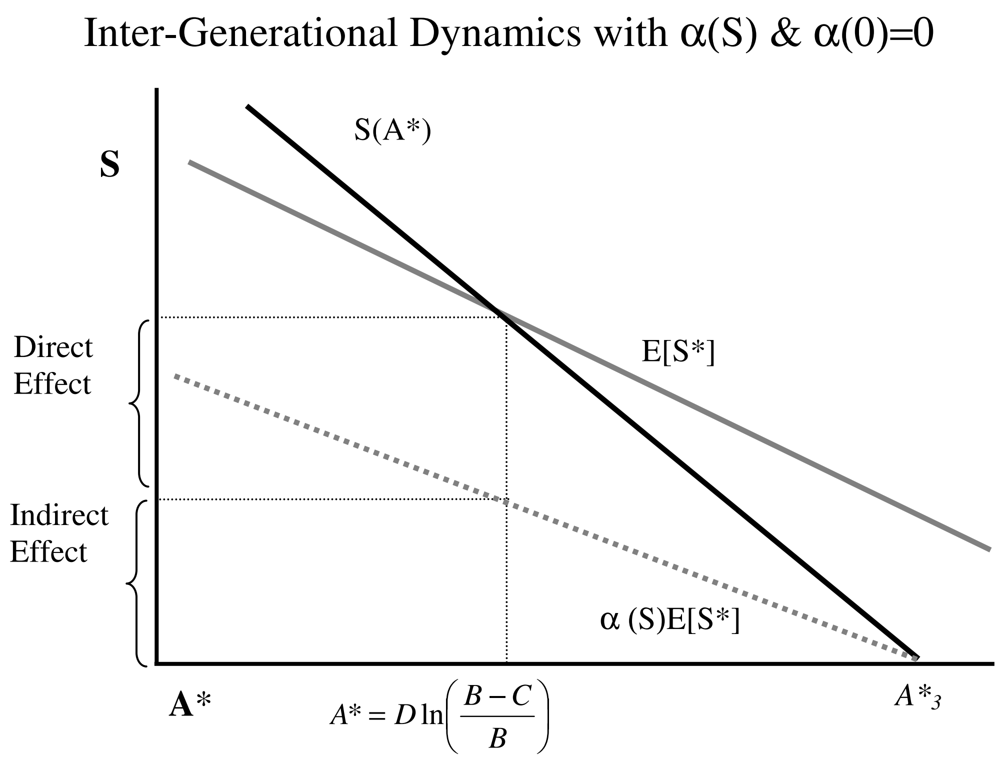

Figure 9 **Alpha at Zero**

Several Puzzles Solved

These dynamics produce a Horatio Alger type story. A kid from the roughest streets beats the odds to graduate top of his class at Harvard. Observers may say that if only all the students in his community had his determination and talent they would all succeed. Those observers would be right. Yet, it is equally true that the determination of other students in that community is discouraged by a rational interpretation of the information immediately available to them. Asking them to try anyway is asking them to be irrationally hopeful and is likely to fall on deaf ears.

Since a lack of successful community members can discourage students from trying successful community members are a positive externality. Parents are, therefore behaving rationally when they try to move into better districts and exclude communities with low success rates from joining their district. It also means there is something to the tribalism that often puzzles economists. If you care about your children's future then it is important not just that you succeed but that your peers also succeed. They will help form the role models that influence your children. Furthermore, successful peers who do not return to the community are harmful and it makes sense that people would dislike them.

This model can also help explain the super premiums that some neighborhoods command without an appeal to simple status. A high premium implies fewer uneducated members which implies a higher premium. Over high ranges of success $\alpha(S)$ is negative and more graduates bring other graduates back with them to a neighborhood causing the success rate and the premium to spiral forward until the premium is so high that no unsuccessful members can afford to live there.

For some neighborhoods it is possible that there is a range over which $\alpha(S)$ is negative and a range at which $\alpha(S)$ is positive. This community can then exhibit two equilibria, one a source and the other a sink. The figure below shows such an example.

In this case the community will tend towards the sink equilibria unless some shock throws the success rate above the source in which case the community will race off to 100% success. This might represent the real life story of an urban community that experiences an influx of educated singles and childless couples who are unconcerned about the ambient success rate. This is enough to make the community attractive to successful couples with children and the community takes off towards 100%.

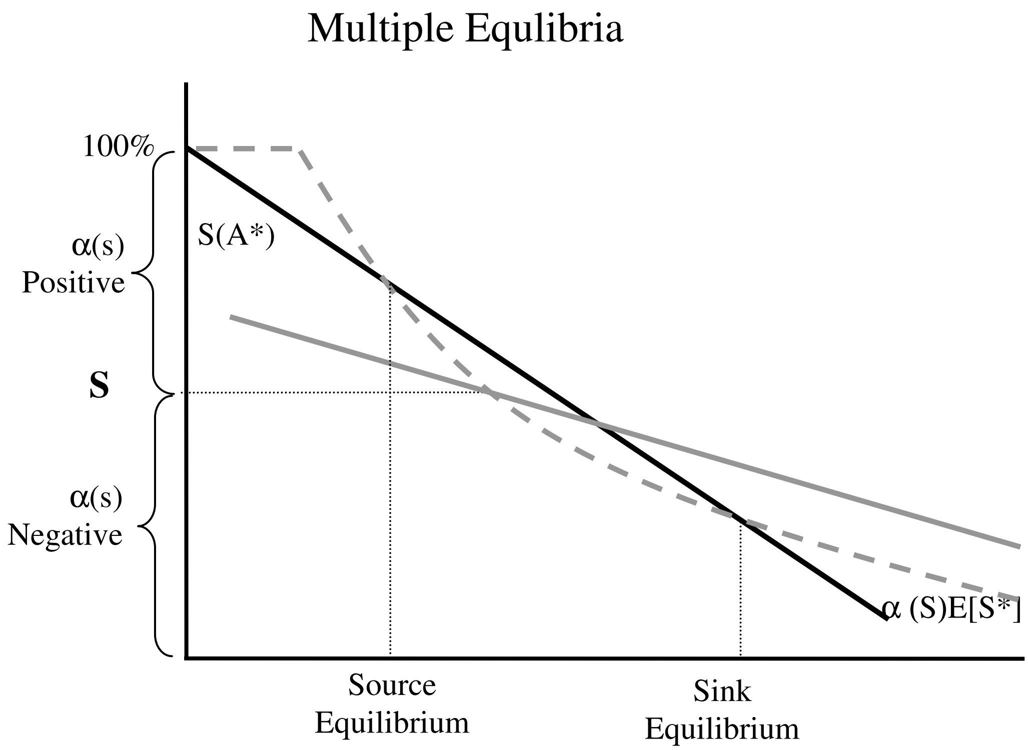

Figure 10 Multiple Equilibria

The Perception and Consequences curves below were simulated using a cost of 100, a benefit of 220 and a difficulty of 1000. A community of 2000 members was randomly drawn from a normal distribution with mean aptitude of 1000 and a standard deviation of 110. Aptitude mean and deviation were chosen to simulate SAT scores. The benefit/cost ratio was chosen to match the estimates from the BLS on wage differences between college and non-college workers and the minimum wage forgone at 20 hours a week during high school. With these numbers the A* is 606. That is, if a student can make at least 606 on his SAT then he should attempt college prep. This is more than three standard deviations below the mean and so it includes just about everyone. However, it does not mean that just about everyone will graduate.

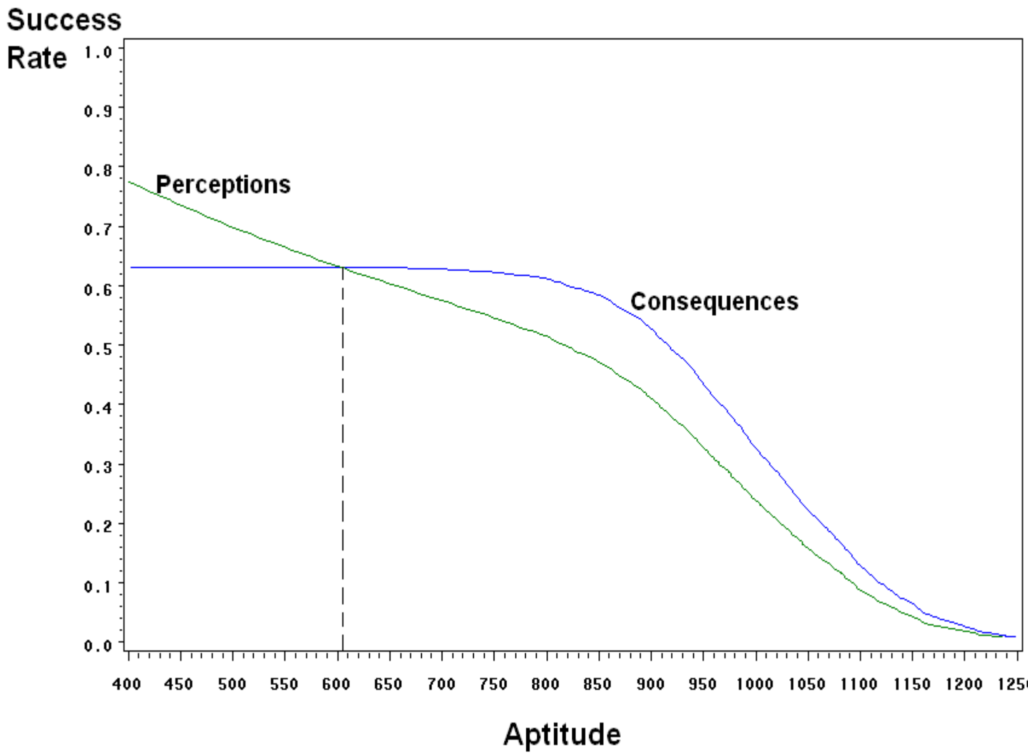

Figure 11 Simulated Curves

Even if everyone tries, not everyone will succeed. In this case only about 63% of the population will successfully graduates even if they all work hard in high school.

Lower Levels and a Slower Response

The above analysis demonstrates that a community with a lower alpha will tend to have a lower college going rate. Furthermore, since a lower alpha corresponds to less desirable communities there will be a tendency for communities with a low alpha to be poor. Yet, where does the slow response come in?

Increases in the wage premium encourage students in the best communities to go to college but may discourage students in the worst communities. However, suppose that on net the effect is positive. If there is an osmosis of families between communities then eventually the higher grad population in the best communities will seep into the second-best communities. Students in these communities will then be more likely to study and the grad population will rise. If it rises enough some families will seep into the third-best community by osmosis and the process will continue slowly until all communities have an elevated response.

How this process evolves, however, is highly dependent on the elasticities of the Perception and Consequences curves, the alpha's and the extent of osmosis or seepage between communities.

NUMERICAL SIMULATION

When analyzing the relationships among multiple communities the combination of discrete stochastic choices and dependence on past events makes an analytic solution difficult. Because agents move between communities increases or decreases in college going in one community can have an impact on the effective alpha of other communities. Suppose for example, that there were two communities A and B and that in general a college graduate was more likely to select community A than community B. This would suggest an alpha greater than one for community A and lower than one for community B. However, if the number of graduates from each community diverges this relationship may not hold. Community A may produce so many graduates that on average they spill out of the community into community B. Likewise community B may produce so few that even small amounts of spillover could induce a positive alpha.

Therefore, it can be helpful to create numerical models to examine relationships between communities. The models can allow us to see how these complex interactions might take place. Equally, important numerical models allow us to check the predictions of the analytical model against the real world data. This will allow us to determine whether or not the patterns in college going behavior among white males in the US could be explained by a prior trap.

The analytics of the prior trap, however, speak only to the supply of education workers. The return to successfully completing college is determined outside of the model. To address that, I borrow the demand side model of Card and Lemieux. I use their structure and elasticites wherever possible. There are a few additional parameters that must be independently calibrated, including the difficulty of school and the personal discount rate of the students.

The results match several interesting empirical facts:

1) Lower levels of college attendance among less affluent communities

2) A more muted response to the increase in the college premium among lower income communities

3) A leveling off of college going rate in the later 20th century

4) A subsequent rise in the wage premium and inequality

5) A higher expected return to exogenous shocks on college going in less affluent communities.

In addition the results indicate a significant and growing premium to being raised in a community with many college educated adults. This could explain an increase in housing prices in neighborhoods with high level of college educated adults.

Model Structure

The supply of college educated workers is determined by the mechanism of the prior trap. I assume that there are F families in N communities for a total of i = (1, . . . , N*F) family lines. The lines are followed over K periods. In each period a family consists of an adult and a student. At the end of each period the old adult dies, the student becomes the adult and a new student is born.

When born each student has a scholastic aptitude which is drawn from the normal distribution with mean 1000 and a standard deviation of 110. This is the scale on which SAT scores are standardized and should allow the reader to easily determine how extraordinary a score may be. It's important to remember, however, that in this model aptitudes are determined exogenously and there is nothing that an agent can do to alter or improve his aptitude.

At the beginning of each period the student must determine whether or not he is going to study in high school. If he chooses not to go to college then he receives

$$\sum_{i=1}^{49} \frac{1}{(1+r)^i}$$

the discounted present value of the non-college wage from age 16 until age 65. The non-college wage is normalized to one.

If he chooses to study then he pays a cost equal to

$$\frac{1}{2} \sum_{i=1}^{6} \frac{1}{(1+r)^i}$$

Which represented half of the wage stream from age 16 until 22. I assume that the student can work no more than part time and still be devoted to his studies.

If the student succeeds then he receives the discounted present value of the college wage equal to:

$$\sum_{i=8}^{49} \frac{p}{(1+r)^i}$$

Where p is the wage premium determined in the labor market.

Following the model presented in Chapter 5 a student's probability of success is determined by the function:

$$s = 1 - e^{-\frac{A_{ik}}{D}}$$

Where A_{ik} is the scholastic aptitude of the student from the i-th line in the k-th period. As in the analytical model the student cannot observe D. Also, as in the analytical model there exists a threshold aptitude such that any student above that aptitude will study and below will not. The students must back out that threshold from observing the adults in their community.

Once, they do so those who have chosen to study either succeed and receive the benefits or do not and receive the non-college wage stream. Adults then sort into communities. The communities are assumed to differ in exogenous amenities. That is, some communities may have more old growth trees, or less risk of flood. All individuals rank these communities the same from best to worst. The individuals differ only on their willingness-to-pay for a better community.

Differences in willingness-to-pay are a normally distributed function of income. That is, on average individuals bid a given fraction of their income for amenities but some will bid a slightly higher fraction and some a slightly lower fraction

The bids are ranked from highest to lowest. The first F families go into community one, the second F into community two and so on until all families have been placed into communities. The students then become adults and new children are born into their respective communities.

The demand side structure is taken from Card and Lemieux.

I assume that the production function is Constant Elasticity of Substitution between aggregates of college and non-college labor of the form

$$y_t = (\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho})^{1/\rho}$$

Where yt is output in period t, H is high school or non-college labor and C is college labor. Following Card and Lemieux I also assume that aggregate high school and college are produced from high school and college labor of different generations indexed by j. So that:

$$H_t = [\sum_j (\alpha_j H_{jt}^{\eta})]^{1/\eta}$$

$$C_t = [\sum_j (\beta_j C_{jt}^{\eta})]^{1/\eta}$$

This implies that labor from different generations are not perfect substitutes for one another, so that an increase in the percentage of college goers in say 1965, not only has an impact on the wages for all college educated workers but has an additional impact on college educated workers in the 1965 cohort.

As usual, ρ equals $\frac{1-\sigma_E}{\sigma_E}$ where σ_E is the elasticity of substitution between education groups and η equals $\frac{1-\sigma_A}{\sigma_A}$ where σ_A is the elasticity of substitution between age groups.

Solving for the marginal productivity of labor, setting equal to the ratio of wages yields:

$$\ln(w^{c}_{jt} / w^{h}_{jt}) = \ln(\theta_{ct}/\theta_{ht}) + \ln(\beta_j/\alpha_j) + [(1/\sigma_A)-(1/\sigma_E)]\ln(C_t/H_t) - (1/\sigma_A)\ln(C_{jt}/H_{jt})$$

To simplify the model I assume that $\beta_j = \alpha_j = 1, \forall j$ implying that change in the log relative wage can be decomposed into year effects, aggregate supply effects and generation specific supply effects.

$$\ln(w^{c}_{jt} / w^{h}_{jt}) = \gamma_t + \varepsilon_A \ln(C_t/H_t) - \varepsilon_C \ln(C_{jt}/H_{jt})$$

Ultimately five parameter values are chosen exogenously

1. γ_t the year effect. Taken from Card and Lemiuex this is assumed to be a linear trend equal to 1.7% per year. This implies that in the absence of supply responses increases in demand for skilled workers will yield a 1.7% per year increase in the wage premium.
2. ε_A the aggregate demand elasticity. Taken from Card and Lemiuex this is assumed to be .414
3. ε_C the cohort demand elasticity. Taken from Card and Lemiuex this is assumed to be .202
4. D, the difficulty of college was set at 1609. D was calibrated to yield a 58% graduation rate for students with an aptitude at least 3 standard deviations above the mean. A 1994 Higher Education Research Institute found that students with SAT scores three standard deviations above the mean had a four year graduation rate of 58%. While aptitude in this model relates innate ability and SAT scores are influenced by education, they still remain the most widely available proxy for scholastic aptitude.
5. The discount rate i was set at 2.5%, to approximate the average TIPS long term yield.
6. Expenditures on housing were set to a mean of 32% with a standard deviation of 4 percentage points. The mean was taken from the consumer expenditure survey. The standard deviation of percentage expenditures was not available. The

standard deviation of expenditure by quintile was equal to 16% of total quintile expenditure. Using 4 percentage points assumes that roughly 75% of the variance in expenditure within quintiles is due to variance in income.

Fortunately, the model is not highly sensitive to this parameter. When the standard deviation exceeds roughly 11% there will be frequent realizations of negative willingness-to-pay which disrupts the model.

Results

We can compare the results with the five empirical facts mentioned above. First, broadly speaking students in lower ranked communities graduate college at lower rates. The lower ranked communities had graduation rates of approximately 10% in 1950, rising to a projected slightly less than 20% in 2009. In contrast, students in the highest ranked communities graduated at roughly 50% throughout.

Second, the response to wage premium is concentrated in community two. Community two goes from a little than 20% graduation rate to an approximately 50% graduation rate, identical to community one. Graduation rates in the other three communities rise but are much more muted. An interesting feature is that there seems to be three levels of graduation. Originally, community two is the sole community in the middle level. When its graduation rate rises it is replaced by community three.

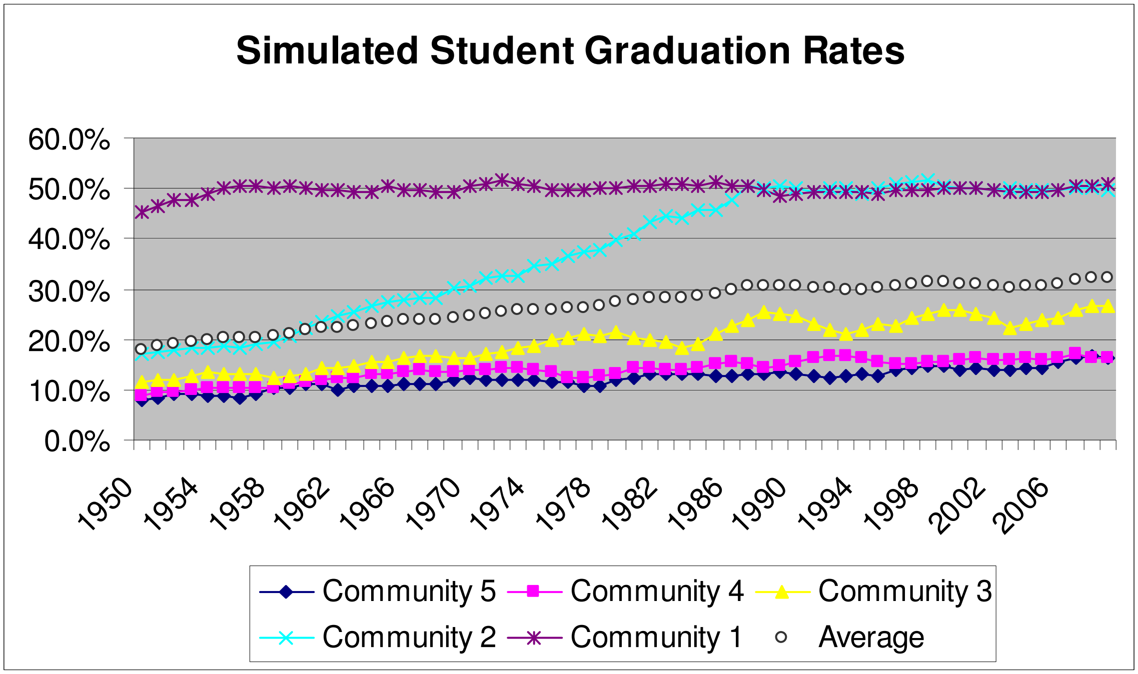

Figure 12 **Simulated Graduation Rates**

A similar response pattern can be seen in terms of the number of students who prepare for college. The rising return to education induces all of the students in community one to attempt college early on. Community two rises to meet community one and the other three communities show a more muted response.

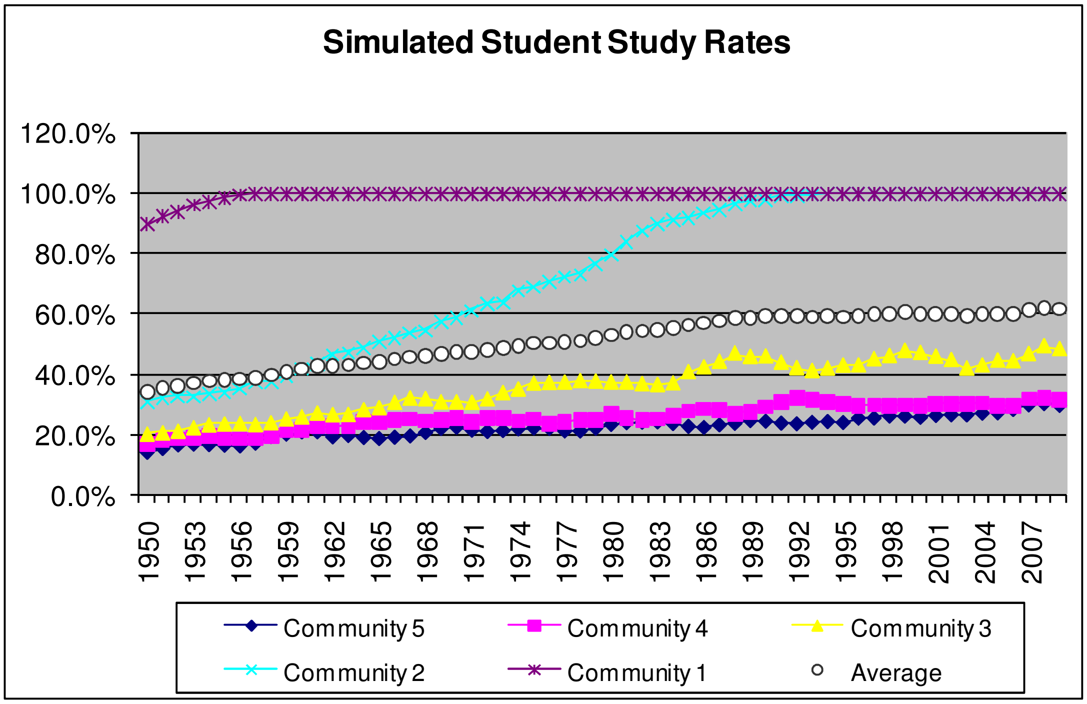

Figure 13 **Simulated Study Rates**

Third, both student preparation and student graduation rates trail off in the later part of the century. However, this trail off occurs 15 years later than in the real world data, a point I will return to later

Fourth, the wage premium begins to accelerate as college graduation rates level off. For real world data this happens in the late 70s for simulated data into the late 1980s early 1990s.

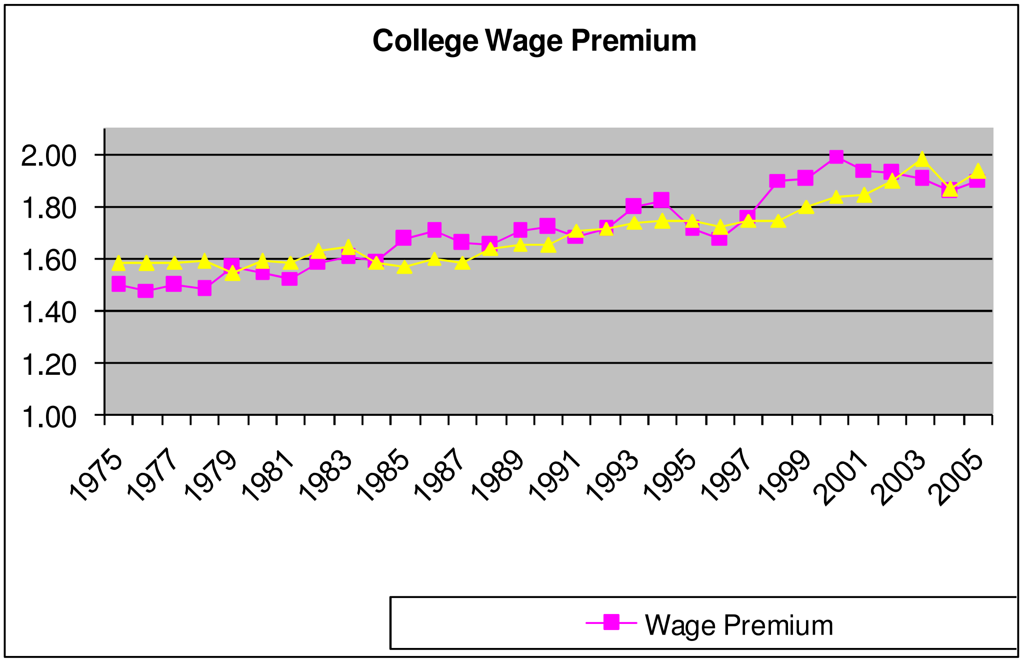

Figure 14 **Simulate Wage Premium**

Fifth, there is increasing inequality following the slowdown. In the model inequality only occurs because of differences in education. However, the graph below shows an dispersion in average discounted present value income between communities.

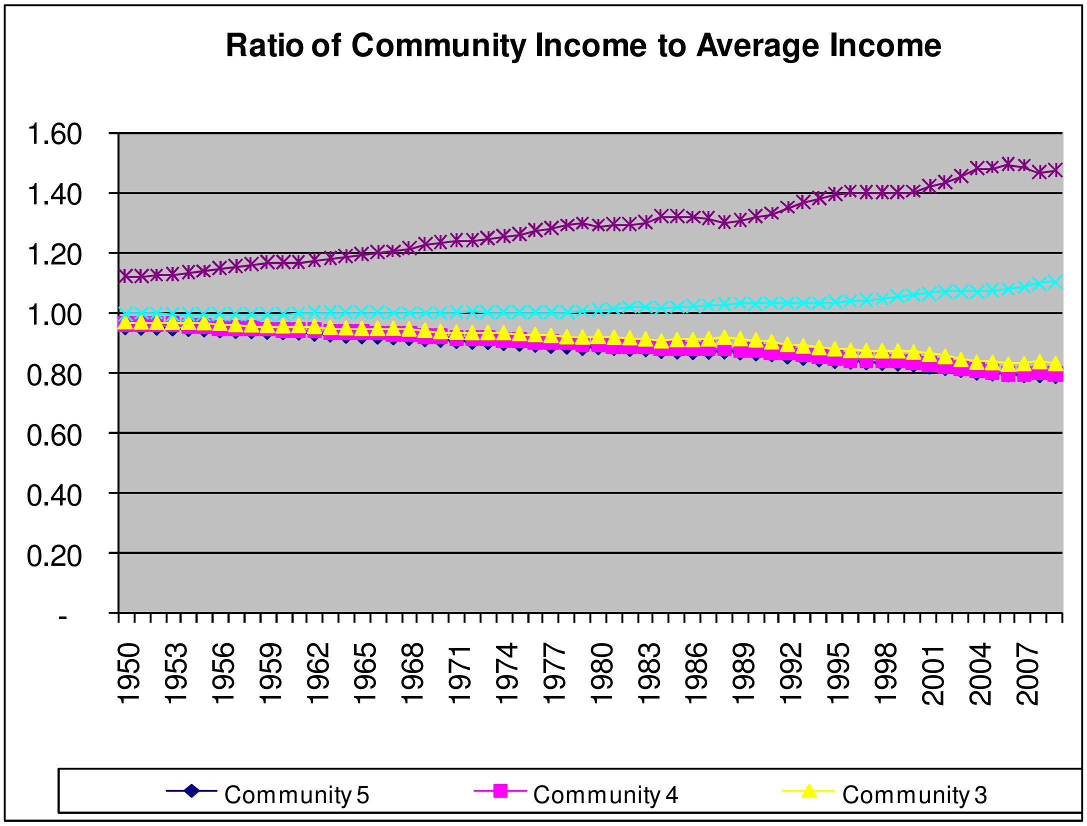

Figure 15 **Simulated Income Ratio**

I now return to a more detailed investigation of predicted graduation rates versus actual graduation rates. The numerical model provides a supply side to the analysis of Card and Lemieux. Therefore, if properly calibrated it should mimic the aggregate graduation rates of the entire population.. The model has no accounting for explicit racial or gender barriers and so I restrict my results to the data for white males.

The model also does not account for immigration. Immigration would affect the results in two ways. First, if adult immigrants are less educated than native born citizens then they will tend to drag down expected graduation levels. Second, if immigrants move into lower income communities then their children will graduate at lower levels than the population average.

In effect immigrants might tend to add additional communities to the bottom of the simulation ranking. Unfortunately separate data was not available for immigrants and native born Americans. However, since 1993 there is separate data on Non-Hispanic whites. To the extent that Hispanics tend to be immigrants or the children on immigrants the Non-Hispanic white numbers should offer a better check.

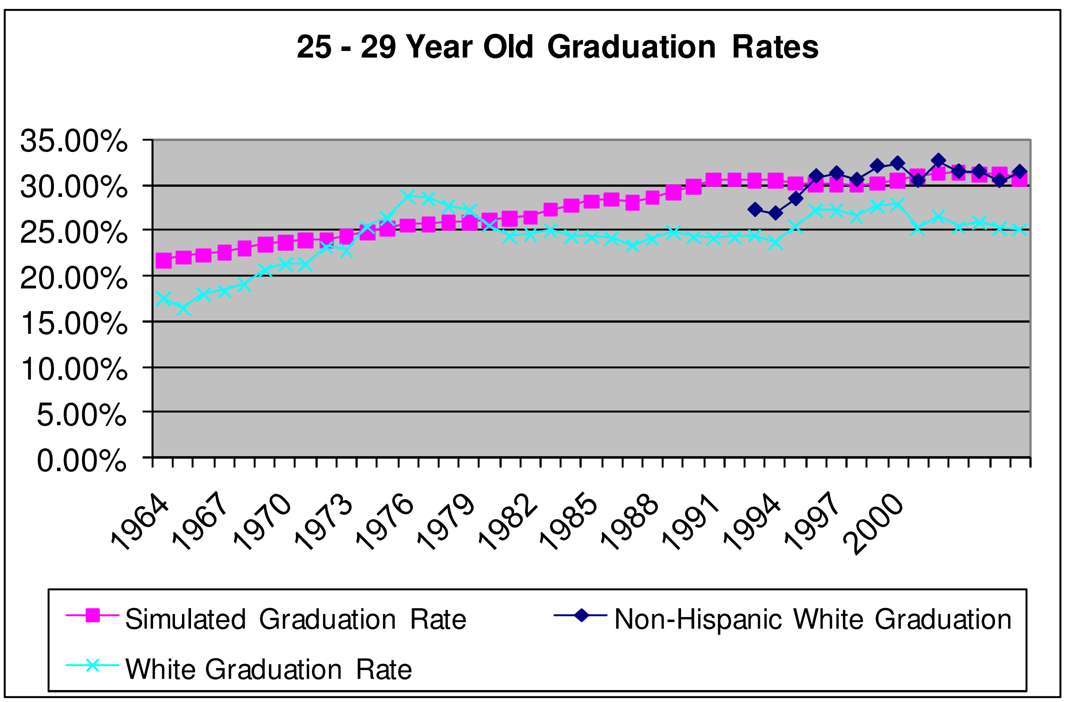

Figure 16 **Simulated Graduation Fraction**

For the limited period in which Non-Hispanic White data is available it tracks closely the simulated data. For most of the period the simulated graduation rates track above the white graduation rates. The exception is a bump in the wake of the Vietnam War. Since, students could gain draft deferment by entering college it might be reasonable to assume that the Vietnam War lowered the opportunity cost of college. A student could expect that rather than attending college there was a chance he would have to spend a year in combat.

It is difficult to estimate the implicit cost being drafted. However, as an approximation I introduce a one third reduction in cost phasing-in during college entrance year 1965-66 and phasing-out during entrance year 1972-73. The total time forced into part-time work because of schooling is 6 years (16 – 22) in this model. A one third reduction in costs presumes that service in Vietnam is equivalent to giving up one year of full time work. Thus the cost a student can spend two of the six years studying and lose nothing if he avoids the draft.

This structure implies that students foresaw an increase in the probability of being sent to war beginning in 1965-66 and a reduction in 1972- 1973. This may or may not be realistic but the results are striking.

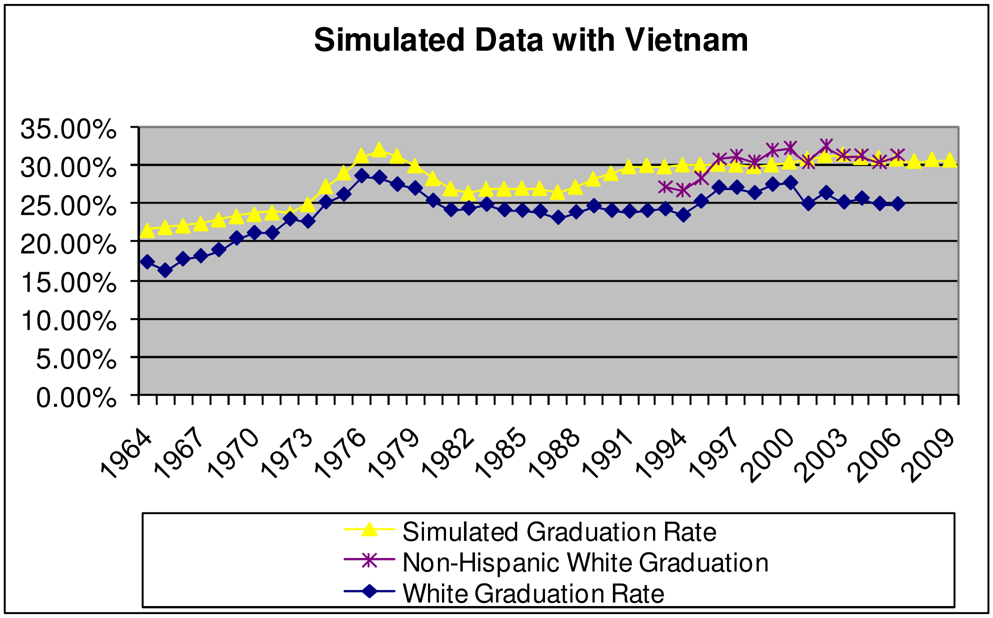

Figure 17 **Simulation with Vietnam Effect**

This striking transformation supports the idea that the bump in college graduation rates came from a reduction in the cost of attendance. Furthermore, the post Vietnam slide is a result of a reduced wage premium. The addition, of the Vietnam effect also causes the wage premium to more closely track the real data during the 1970s and early 1980s

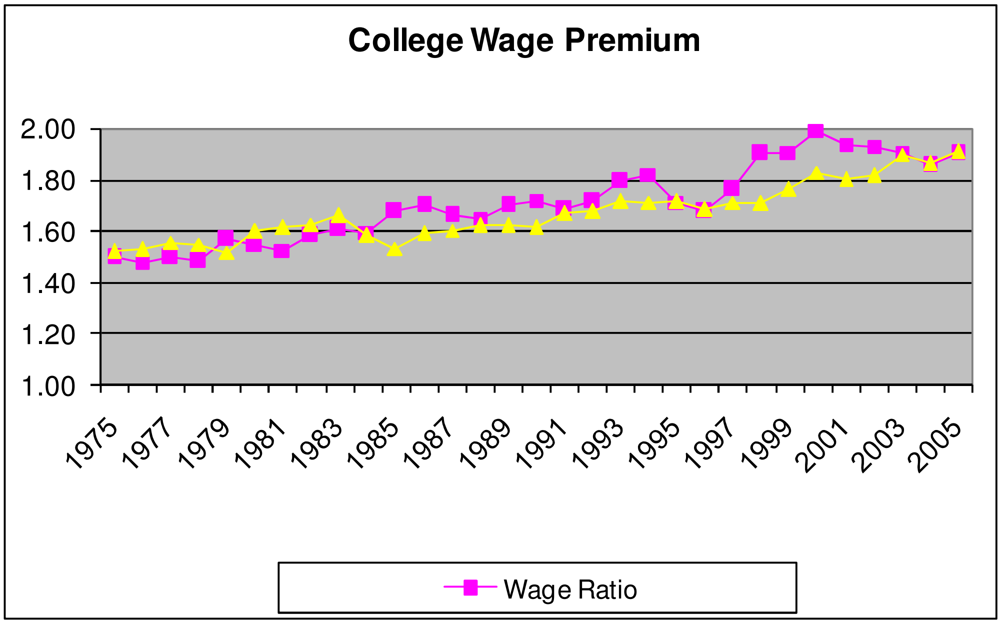

Figure 18 **Wages with Vietnam**

There is still an un-simulated bump in the wage premium during the late 1990s. Given that simulation track white graduation rates during the 1960s – 1980s and Non-Hispanic White Graduation rates in the 1990s to 2000s it could be the case that the wage premium is being generated by an influx of Hispanic immigrants which drive down the high school wage rate.

The model of the prior trap here is simple. There are only five communities. There is no accounting for racial, gender, or other social issues that might have influenced white male college going. There is no accounting for aggregate economic shocks, geographic variance in job growth or college quality, or education related policy initiatives. Yet, a prior trap is able to simulate much of the supply side slow down in college workers since the 1970s. This suggests that the mechanics underlying the prior trap are at least a candidate for explaining the slower growth in white male college going in the United States.

However, given how closely the model tracks the data it may be interesting to see what results we get in a model that does not exhibit a prior trap. Would any model that used the Card and Lemiuex demand side perform as well?

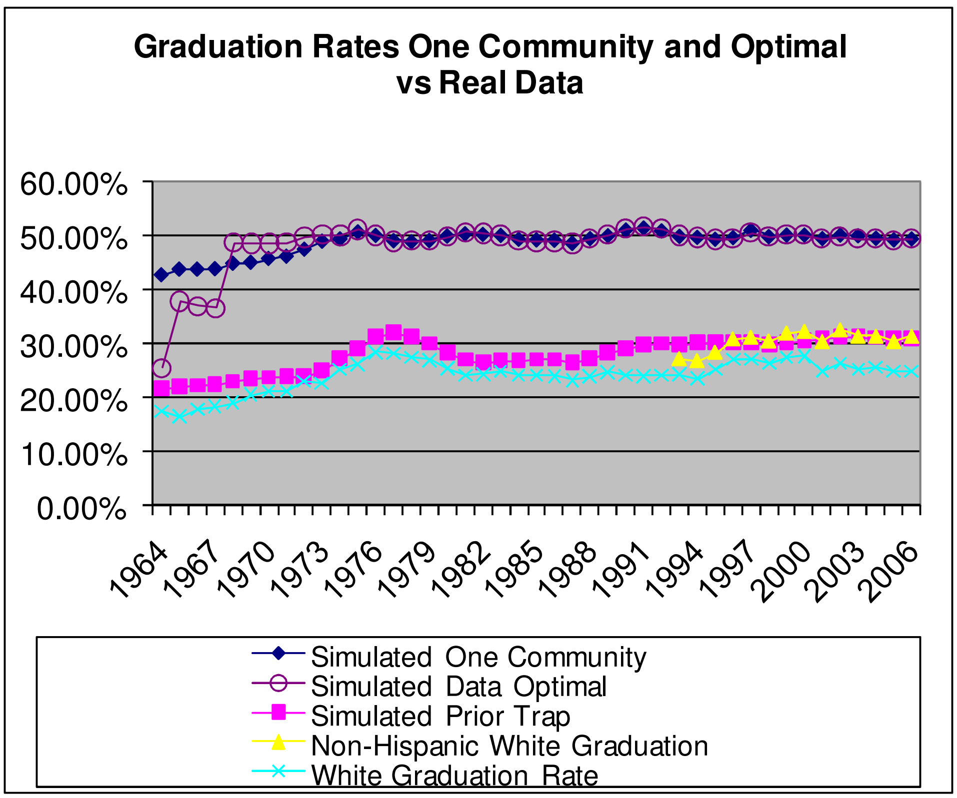

Figure 19 **Control Simulation**

To test that I simulated two different non-prior trap models. In the first there is only a single community. Since, all students can observe all adults there is no bias in the data and the community should arrive at the correct value for D and the corresponding threshold aptitude. In the second each student was given the actual value for D. The two series converge very quickly and at a much higher level than either the prior trap or the real world data. Indeed, an average graduation rate of approximately 50% implies that by 1975 virtually all students are preparing for and attempting college. That is, the return from college were large enough that it would have been worth it for all students to engage in a college prep education. It is also worth noting that this does not imply that 50% of college students flunked out.

The data from the Higher Education Institute on which this model was calibrated only measures how many students graduated, not how many flunked out. Some may leave early because of other opportunities, because they do not like the social or cultural environment of college or because they simply decide that college is not for them. The idea behind this model is that a student will not know these things for sure until he or she finishes high school and is ready to go to college. Thus, if the potential return from a college education is large it will make sense for almost all students to keep their options open by preparing for college during high school.

A model in which students have full information about their probability of success does not explain the data, even when the demand side of Card and Lemiuex is introduced. In particular, the secular rise in demand for college educated workers underpinning Card and Lemieux should have lead almost all students to attempt college. A large increase in the return to college without a large response in college going requires an inelasticity of supply that is not explained by a model with fully informed agents.

CONCLUSIONS

Low college going rates among economically disadvantaged and minority students have been explained by an appeal to either credit constraints or inadequate school funding. Yet, both of these explanations have come under empirical scrutiny. I offer a third explanation, that students less affluent have lower expectations. More specifically, lower income students underestimate the return to college preparation. Community observations lead those student to conclude that they are unlikely to complete college even if they study and thus the expected benefit from college preparation is low.

This approach fits into the literature in three ways. It provides a rational basis for the neighborhood effect. It explains why family income during primary schooling is a determining factor in a student educational career and it helps to explain the puzzle of low college graduation rates coupled with a rising return to college. The model is particularly successful in the third area. The dynamics which fall out this approach create a kinked supply curve. Beyond a certain point increases in residential segregation due to a rising return to college retard college going almost as much as the return itself encourages it.

Throughout this paper, the decision to prepare for college is couched as a binary choice. Students either study or they do not. Completing college is also treated as either/or consequence. In reality the more a student prepares for college the more years of college he or she is likely to complete. A next step will be developing a model in which increased preparation yields an increase in expected years of schooling completed.

Yet, even the analysis here cast the problem of low college going rates in a unique light. If this analysis is correct then increasing college going rates in less affluent communities requires getting students to disregard or at least discount the most readily available source of empirical data available to them. This is at what comforting and disheartening.

It is comforting because this analysis implies that the key is inspiring students. That is, getting them to believe they can literally beat the odds and achieve what they have little evidence to believe that they can achieve. It is comforting because this is what the educators in the field tell us the key is.

It is disheartening, however, because inspiring students has proven such a difficult task and as of yet there is no formula for it. To the extent this analysis quantifies the problem it may point us in the direction of prescription for inspiration. Or, at least that is my hope.

Reductions in residential segregation by income may be helpful, but only to the extent that encourage social integration. Families that live near one another tend to know one another when people freely choose where to live. However, if families were prevented from residentially segregating themselves there is a reasonable possibility that they would seek other means of social segregation. It is important to remember that no family in a community with lots of college graduates has an incentive to socialize with families in communities with low graduates.

Even if all families opened up, at best it could have no negative effect of families in the high graduation community and there is always the possibility that exposure to non-graduates will discourage children of graduates. In a sense segregation is a form of social pollution and preventing suffers from the same free rider problems as true pollution.

On a larger scale the techniques in this paper may explain behavior outside of the decision to prepare for college. Life is full of choices. Most of the time we have little direct information about what the consequences of those choices will be. Most of what we know, or at least what we think we know comes from observing the people around us. Using the behavior of others to deduce the consequences of our behavior own behavior is a necessary part of our lives. This implies that to the extent that people self-select into groups our interpretation of optimal behavior, of long run consequences and of the underlying mechanics of our world may tend to differ systematically.

This analysis may begin the examination of rational socialization. That is, even if individuals only care about their own consumption it is rational for them to base their actions on the actions of some reference group. Moreover, social groups can solve optimizations problems that would be impossible for any individual alone. Homo Economicus is an inherently social being.

BIBLIOGRAPHY

Angrist, Joshua D. and Alan B. Krueger "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics*, Vol. 106, No. 4. (Nov., 1991), pp. 979-1014

Bound, John, Michael Lovenheim and Sarah Turner, "Understanding the Increased Time to the Baccalaureate Degree" *Rough Draft*, Oct. 2006

Brock, William A and Steven N Durlauf "Discrete Choice with Social Interactions" *Review of Economic Studies* Vol 68 No. 2, pp 235–260

Card, David and Thomas Lemieux "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis" *The Quarterly Journal of Economics*, Vol. 116, No. 2. (May, 2001), pp. 705-746

Fryer Roland G. Jr, and Steven D. Levitt "Understanding the Black-White Test Score Gap in the First Two Years of School" *Review of Economics and Statistics* Vol. 86 No 2, (May 2004) pp 447 – 464

Hanushek, Eric A, "The Economics of Schooling: Production and Efficiency in Public Schools" *Journal of Economic Literature*, Vol. 24, No. 3. (Sep., 1986), pp. 1141-1177

Heckman, James and Pedro Carneiro, "Human Capital Policy" *NBER Working Paper* No. 9495 Feb 2003

Keane, Michael P and Kenneth I. Wolpin, "The Career Decisions of Young Men" *Journal of Political Economy* Vol. 105 No. 3 (Jun, 1997) pp. 473-522

Loury, Glenn "Intergenerational Transfers and the Distribution of Earnings" *Econometrica*, Vol. 49, No. 4, (Jul., 1981), pp. 843-867.

Mayer, Susan, "Income Inequality, Economic Segregation and Children's Educational Attainment" Working Paper 2000

Rivkin, Steven G., Eric A. Hanushek and John F. Kain "Teachers, Schools, and Academic Achievement" *Econometrica* Vol 73, No. 2 , (Mar. 2005), pp 417-458

Streufert, Peter "The Effect of Underclass Social Isolation on Schooling Choice" *Journal of Public Economic Theory* Vol. 2, No 4, (Oct, 2000) pp 461–482.