Neighborhood Effects Can Increase Inequality and Hurt Those at the Bottom

David M. Drukker

February 25, 1999

Abstract

This paper investigates the general equilibrium consequences of assuming that the time it takes an individual to obtain a skill depends on the average level of education in the neighborhood of his/her birth. These neighborhood effects provide an incentive for the high educated to break-off and form an exclusive neighborhood. Economic segregation is implemented by a zoning requirement which specifies that members of the better neighborhood must purchase an exogenously given amount of housing services. The investigation reveals that individuals born into better neighborhood prefer economic segregation while those born into the less well-off neighborhood prefer the case of one neighborhood. The analysis also shows that various measures of economic inequality increase with the change from one to two neighborhoods, and increase with the exclusivity of the zoning restriction. The behavior of the agents in this model is similar to a central hypothesis of *The Truly Disadvantaged* [19].

1 Introduction

This paper investigates the general equilibrium consequences of assuming that the time it takes an individual to obtain a skill depends on the average level of education in the neighborhood of his/her birth. These neighborhood effects provide an incentive for the high educated to break-off and form an exclusive neighborhood. Economic segregation is implemented by a zoning requirement which specifies that members of the better neighborhood must purchase an exogenously given amount of housing services. The investigation reveals that individuals born into better neighborhood prefer economic segregation while those born into the less well-off neighborhood prefer the case of one neighborhood. The analysis also shows that various measures of economic inequality are increasing in the exclusivity of the zoning restriction. The behavior of the agents in this model is similar to a central hypothesis of *The Truly Disadvantaged* [19].

Instead of assuming that education has external effects, previous studies have treated education as a local public good. Durlauf [4] and Benabou [1] have shown that if capital markets are imperfect and education is a local public good whose quality is determined by its resources, then a voting mechanism causes society to split up into distinct neighborhoods. The equilibrium of the multi-neighborhood economy is not Pareto optimal. Papers by Epple et al [5] and Fernandez and Rogerson [6] investigate the allocation of a local public good and the welfare properties of models in which individuals with exogenously determined incomes endogenously determine where to live. But these papers suffer from an empirical problem. The Coleman Report [3] noted that there was no simple relationship between a school's resources and average student outcomes. More recently, Hanushek [8] attributes inter-school differences in achievement to differences in teacher quality. The distribution of teacher quality is determined not just by salaries but also by the environment that the school has to offer. This environment is largely determined by the neighborhood. Further empirical support for the importance of a neighborhood based mechanism comes from a series of earnings regressions done by Betts [2]. Using the NLSY panel data set, he found that dummy variables for the high school attended were statistically significant in predicting labor market success, while the actual measures of school quality were not. Perhaps, the dummy variables were picking up the neighborhood effects. Minicozzi [12] has found that income based measures of neighborhood quality have significant nonlinear effects on social mobility.

The Gautreaux Program in Chicago is a natural experiment that suggests the importance of neighborhood effects. A court ordered attempt to redress previous racial discrimination in public housing resulted a sample of public housing residents receiving vouchers for better apartments in either the city or the suburbs of Chicago. Rosenbaum [13] and Rosenbaum and Popkin [14] argue that the selection process yielded a sample representative of a majority of the African American public housing residents. Since people admitted to the program were randomly assigned to city or suburban spots, the authors argue that the program serves as a natural experiment. Rosenbaum [13] finds that the children of those that received suburban apartment had significantly better educational achievement and early labor market success.

In the sociological literature, Wilson [19, 20] argues that adults "collectively socialize" the next generation. Wilson [19] found that the number of extreme poverty tracts in Chicago was increasing and that much of increase could be attributed to out-migration by middle- and working class families. This outflow increased the incidence of poverty in these areas giving rise to "concentration effects" and "social isolation" of its residents. According to Wilson [19, pp. 46-62], the high concentration of poverty deprived the children of the remaining residents of successful role models. Role models provide information. Wilson [19, p. 56] notes that

> the very presence of [middle- and working class] families provides mainstream role models that help keep alive the perception that education is meaningful, that steady employment is a viable alternative to welfare, and that family stability is the norm, not the exception.

Wilson also argues that a high incidence of poverty will affect the social norms governing the community. His point can be buttressed by a more economic logic. Social norms imposed and enforced by resident adults can be especially important in controlling drug use and criminal activity by adolescents. Because of the higher opportunity costs, higher levels of education among adults reduce drug use and criminal activity. An outflow of adults with high opportunity costs will increase the proportion of in the neighborhood adults engaging in such activities. This increase reduces the effectiveness of the group of all adults in collectively pressuring adolescents not to engage in such behavior.

Surveys by Jencks and Mayer [9] and Gephart [7] classify Wilson's theory as a theory of "collective socialization." Neighborhood adults outside of the family collectively play an important role in socializing the children of the neighborhood. Both surveys conclude that there is some evidence that growing up in a poor neighborhood reduces an individual's education and adversely affects her/his chances of labor market success.

This paper examines the general equilibrium effects of one formalization of collective socialization. Collective socialization is here reduced to the idea that the psychic and time costs of obtaining a given level of education are higher in poor neighborhoods than in better neighborhoods. Here, the psychic and time costs of obtaining a given level of education are decreasing in the quality of the neighborhood.

Marshall's [11] concept of local external effects provides the argument for the proposed extension and standard technique the effect into a formal model. Marshall [11, p. 271] argued that industries tend to agglomerate in given localities because, once established, their presence and activity greatly reduce the costs of training workers and advancing the production technology. The relevant part of this often quoted passage is that

> the mysteries of the trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously.

Children learn some of the skills of a trade through normal social interaction. This process of knowledge transfer from adults to children is not be limited to any specific skill. The general formulation of Marshall's idea has adults transferring general human capital to children through normal social interaction. The more educated the adults in a neighborhood are, the higher will be the rate of knowledge transfer that occurs from social interaction. Such transfer can be as simple as what a child picks up from a conversation between the parents of her/his friends. Alternatively, the process may be more formal. For instance, consider a youngster who is born to technologically challenged parents and is interested in computer programming. The higher the average level of education in the neighborhood, the greater the probability that the child will encounter an adult, or another child, willing and able to help. Wilson argues that collective socialization is a key determinant of the educational and labor market success of the children of a neighborhood. He cites middle class flight as an important cause in the sharp increase in the number of very bad neighborhoods and their serious decrease in quality.

The "collective socialization" mechanism modeled in this paper complements the work of Benabou [1] and Durlauf [4] because it evades Coleman type problems. It also extends the ideas of Epple et al [5] and Fernandez and Rogerson [6] because both incomes and the sorting of individuals over neighborhoods are endogenous.

Neighborhood effects are not the only factor that determines the time it takes a child to learn. All of the above arguments for neighborhood effects, and many more, imply that the time cost of education should be decreasing in the level of parental human capital. Individual ability is also an important factor in fixing an individual's opportunity.

Individuals in this model world choose their training, where to buy a house and how much to consume. One can remain low skilled and go straight to work at the low wage. Alternatively, an individual can spend some time studying to become high skilled before earning the high wage. The time required is strictly decreasing in the average skill level in the child's formative neighborhood, the parents' level of education and the child's ability level. Wages clear the labor markets. There are two cases. In the one neighborhood case there is no zoning and everyone lives in a single neighborhood. When there is zoning, individuals endogenously distribute themselves over the two neighborhoods. The equilibria of the one and two neighborhood cases are derived and compared.

In the two neighborhood case, only the high skilled can afford to purchase the exogenously zoned minimum house size and enter the high neighborhood. The average level of education in the low neighborhood is increasing in the number of high skilled people who purchase houses in this neighborhood. Numerical simulations of the theoretical model show that there are wide ranges of parameters for which several key results hold. Those born into the better neighborhood prefer the case of two neighborhoods to the single neighborhood case. In contrast, those born into the low neighborhood from low skilled parents prefer the single neighborhood case.

An increase in the exogenous zoning requirements generates a process similar to that described in *The Truly Disadvantaged*. The middle class of high skilled in the low neighborhood flees into the better neighborhood. This flight raises the cost of becoming high educated for those born into the low neighborhood. Those born into the low neighborhood, of either high or low skilled parents, are made worse off, while those born into the high neighborhood are made better off.

The middle class flight is due to a wealth effect. An increase in the exogenous zoning requirement raises the value of a spot in the better neighborhood relative to one in the worse neighborhood. Since children inherit the spots of their parents, an increase in this difference in inherited wealth makes it relatively easier for children of the high neighborhood to become high skilled than for children of the low neighborhood. The increase in the difference of relative difficulties causes some of the middle class to move from the low to the high neighborhood.

There is more economic inequality, less social mobility and the distribution of opportunity is more unequal in the two neighborhood case than in the one neighborhood case. Increasing the zoning requirement for entry into the high neighborhood, increases economic inequality, lowers social mobility and makes the distribution of opportunity more unequal.

This paper also provides a new framework for analyzing economic opportunity. The cost that an individual of a given ability must pay to become high skilled depends on the circumstances of her/his birth. Those born into better situations have to pay less to become high educated than those born into less fortunate surroundings. This paper identifies lower opportunity costs with more opportunity. Furthermore, the cost of education function is parameterized so as to permit the derivation of the analytic cumulative distribution function of opportunity. The behavior of this CDF is examined over several cases and it is found that there is greater inequality in opportunity when there are two neighborhoods than when there is only one. Furthermore, increases in the degree of exclusivity of the better neighborhood increase inequality in the distribution of opportunity.

2 The Model

2.1 The Inhabitants and Their Utilities

In this world, individuals must make three choices. People must decide where to purchase a house, how much to consume, and what skill level to obtain.

People care about their own consumption, the quality of their housing and the expected utilities of their children. They earn income from work and also inherit the house of their parent which they may either sell or choose to inhabit. People spend their incomes on consumption and housing.

Individuals are born from a single parent, who is either high or low skilled. Everyone is born low skilled and this level of human capital is normalized to 1. Some people become high skilled and this level of human capital is parameterized at $\tilde{H}$ units of human capital.

There may be either one or two neighborhoods. Houses yield dwelling services to their single inhabitants. All houses within a neighborhood are identical. When there is one neighborhood, all houses yield a single unit of dwelling services and neighborhood choice is trivial. When there are two neighborhoods, one of them contains a higher proportion of high skilled residents than the other. The neighborhood with a higher average level of skill is denoted as the high skilled or better neighborhood. The other is known as the low skilled or worse neighborhood. In the two neighborhood case, low neighborhood houses offer a single unit of dwelling services while high houses yield D_H units. D_H is set by exogenously determined zoning restrictions.

In the two neighborhood case, people must choose from the following menu

1. remaining low skilled and living in the low neighborhood;
2. staying low skilled but buying a house in the high neighborhood;
3. achieving the high level of skill but purchasing a house in the low neighborhood.
4. becoming high skilled and buying a house in the high neighborhood;

Any choice from this menu is a path. Each path is uniquely identified by a skill level and a neighborhood. Skill level is denoted by $h_t \in \{L, H\}$. A neighborhood choice is $N_t \in \{L, H\}$. The four possible states are LL, LH, HL, and HH, where the first letter gives the skill level, h_t and the second indicates neighborhood choice, N_t.

It is important to highlight the distinction between the skill level and the amount of human capital wielded by someone with that skill. Skill is a binary variable over Low and High. So an individual has a skill level $h_t \in \{L, H\}$. The parent's skill level, h_{t-1} is restricted to the same set. Human capital, denoted by H, is a binary variable over the levels of human capital that accrue to each state. The human capital of a low skilled individual is normalized to 1. A high skilled individual has $\tilde{H}$ units of human capital, where $\tilde{H}$ is a parameter of the model. An individual has $H_t \in \{1, \tilde{H}\}$ units of human capital. The parent's level, denoted by H_{t-1}, is also in $\{1, \tilde{H}\}$.

There are three important restrictions on the time cost of education function. The study time required to become high skilled should be strictly decreasing in the average level of education in the neighborhood. It should also be strictly decreasing in an individual's own ability and strictly decreasing in the parent's level of education. It will be necessary to integrate over the cost function, the following functional form was chosen,

$$c(H_{t-1}, N_{t-1}, \omega) = \frac{1}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega), \tag{1}$$

where

- N_{t-1} denotes the child's formative neighborhood, chosen by the parent
- ω is the ability level of the individual
- H_{t-1} is the amount of human capital managed by the parent

$$H_{t-1} = \begin{cases} 1 & \text{if the parent is low skilled} \\ \tilde{H} & \text{if the parent is high skilled} \end{cases}$$

- $\bar{H}_{N_{t-1}}$ is the average level of human capital in the child's formative neighborhood.

This function specifies the amount of time that a given child must study in order to become high skilled. Note that since $\omega \in [0,1]$ the range of possible costs is $[0, \frac{1}{H_{t-1}\bar{H}_{N_{t-1}}}]$, and that the upper bound is less than or equal to 1.

People have 1 unit of time which they must divide between work and study. Consider an individual with ability ω born to a parent with a skill level h_{t-1} in a neighborhood with an average human capital level of $\bar{H}_{N_{t-1}}$. If this person remains low skilled, then her/his income does not depend on the cost function. By going straight to work, s/he earns the low wage, W_L. However, if this person becomes high skilled then s/he has $\tau = 1 - \frac{1}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega)$ units of time to work at the high skilled wage. Becoming high skilled will result in an income of $\iota = W_H \left(1 - \frac{1}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega)\right)$.

Since both skill levels are in the choice sets of individuals from both neighborhoods, a new approach must be taken to the identification of opportunity. R. H. Tawney provides a colorful and useful definition of equal opportunity,

> [Equality of opportunity] obtains in so far as, and only in so far as, each member of a community, whatever his birth, or occupation, or social position, possesses in fact, and not merely in form, equal chances of using to the full his natural endowments of physique, of character, and of intelligence. In proportion as the capacities of some are sterilized or stunted by their social environment, while those of others are favored or

pampered by it, equality of opportunity becomes a graceful, but attenuated figment. [17]

This definition implies that equal opportunity occurs when a person with ability ω faces the same opportunity cost of becoming high skilled, regardless of her/his birth. If two people of the same ability level have different opportunity costs of becoming high skilled because they are born into distinct circumstances, then they have different opportunities in life. If a person is born into a social position with ability ω has a higher opportunity cost of becoming high skilled than another individual with the same ability born into a better social position, then the former has less opportunity. The above specification of the time cost of education function suggests that an individual's "opportunity" is $H_{t-1}\bar{H}_{N_{t-1}}$. In this formulation, opportunity varies over initial states but is unaffected by one's own ability.

This paper investigates the nature and degree of inequality that is purely a product of social and technological causes. Hence, the ability of a child is assumed to be independent of the ability level of the parent and ability is assumed to distributed Uniform $[0,1]$.

People have Cobb-Douglass preferences over dwelling services and consumption. Parents discount the value of their children's utility at a factor $\beta \in (0,1)$. Given this information, one can now write down maximization problem facing each individual in this economy. A person with ability ω, born to a parent with skill h_{t-1} and into a neighborhood with an average human capital level of $\bar{H}_{N_{t-1}}$ must solve

$$\max_{X,N_t,h_t} X^{\gamma} D_{N_t}^{1-\gamma} + \beta \tilde{V}_{h_t,N_t} \tag{2}$$

such that

$$P_X X + P_D D_{N_t} = W_{h_t}\left[1 - S_t \frac{1}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega)\right] + P_D D_{N_{t-1}},$$
$$N_t \in \{L, H\} \text{ and } N_{t-1} \in \{L, H\},$$

and

$$h_t \in \{1, \tilde{H}\} \text{ and } h_{t-1} \in \{1, \tilde{H}\}$$

where

- X is the amount of the consumption good

- N_t is the neighborhood chosen by the child
- N_{t-1} is the neighborhood chosen by the parent
- $\tilde{V}_{h_t,N_t}$ is the expected value function for the path h_t, N_t
- P_D is the price of a unit of dwelling services
- P_X is the price of a unit of the consumption good
- D_{N_t} is the amount of dwelling services in the neighborhood chosen by this person.
- $D_{N_{t-1}}$ is the amount of dwelling services in the neighborhood chosen by the parent of this person.
- h_t is the skill level chosen by this person
- H_t is the level of human capital that corresponds to the skill level chosen by this person
- W_{h_t} is the wage that corresponds to the chosen skill level.
- S_t is an indicator variable that equals 1 if $h_t = H$ and zero otherwise.

The combination of an initial state and a terminal state completely determine the levels consumption and dwelling services. For example, the budget constraint of someone of ability level ω who is born into the LL state and chooses the path LL is

$$W_L + P_D * 1 = X + P_D + 1.$$

Thus, if someone born into LL chooses to go LL then s/he will consume

$$X = W_L.$$

Hence, the lifetime utility of that this path offers someone from LL is

$$LU(\omega, LL, LL) = W_L^{\gamma} + \beta \tilde{V}_{LL}.$$

One can perform similar calculations on an individual from a general state $h_{t-1}N_{t-1}$. Let $LU(\omega, S_0, S_1)$ be the lifetime utility obtained by a person of ability ω born into state S_0 who opts for state S_1. The calculations imply that someone with ability level ω born into the state $h_{t-1}N_{t-1}$ must choose between the following four lifetime utilities,

$$LU(\omega, h_{t-1}N_{t-1}, LL) = \left(W_L + P_D(D_{N_{t-1}} - 1)\right)^{\gamma} + \beta\tilde{V}_{LL} \tag{3}$$

$$LU(\omega, h_{t-1}N_{t-1}, LH) = \left(W_L + P_D(D_{N_{t-1}} - D_H)\right)^{\gamma} D_H^{1-\gamma} + \beta\tilde{V}_{LH} \tag{4}$$

$$LU(\omega, h_{t-1}N_{t-1}, HL) = \left(W_H - \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega) + P_D(D_{N_{t-1}} - 1)\right)^{\gamma} + \beta\tilde{V}_{HL} \tag{5}$$

$$LU(\omega, h_{t-1}N_{t-1}, HH) = \left(W_H - \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega) + P_D(D_{N_{t-1}} - D_H)\right)^{\gamma} D_H^{1-\gamma} + \beta\tilde{V}_{HH} \tag{6}$$

Note that the first two choices are independent of the individuals ability level ω.

2.2 Which Paths are Chosen In Equilibrium

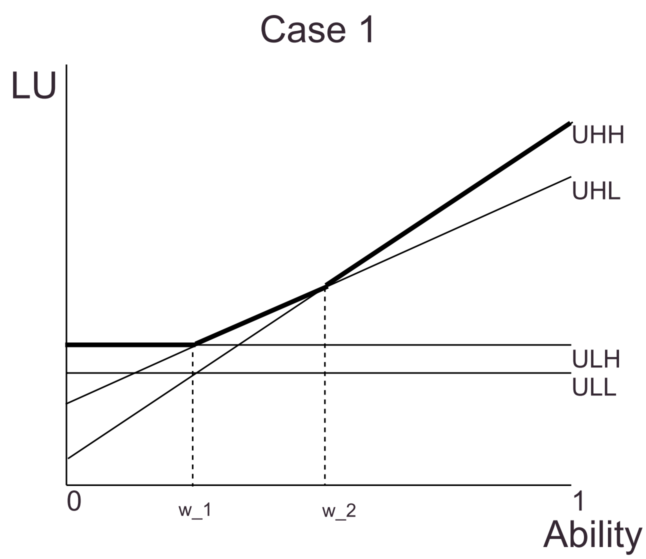
This section demonstrates that only 3 of the four paths will be chosen in any stable equilibrium.

Consider the problem of a person with an ability level ω who is born into some initial state. This person will choose the path that offers the highest lifetime utility. On a graph of the four possible paths over ability, each person ω will choose the path that offers him/her the highest level of lifetime utility. Thus, only the paths on the outer envelope will be chosen.

The Lifetime Utilities for each path derived in equations (3-6) have three key implications. First, as is shown in Appendix B, the HH path is steeper over ω than the HL path. Second, as is clear from the equations, both the LL and LH paths are independent of ability. Finally, that the vertical intercept for HL path is higher than the vertical intercept of the HH path is a necessary condition for the HL path to be chosen in some equilibrium. Were the intercepts reversed since the HH curve is steeper than the HL path, the HL path would never be chosen in equilibrium. These three results imply that there are four possible configurations of the Lifetime Utilities when graphed over ability.

Consider Case 1, depicted in figure 1. Ability ω is on the horizontal axis and lifetime utility is on the vertical axis. The line "UHH" in the graph gives the lifetime utility for each ability level ω for someone born into this general initial state. In case 1, the outer envelope touches the paths LH, HL and HH. There is an important general point that can be gleaned from this

Figure 1:

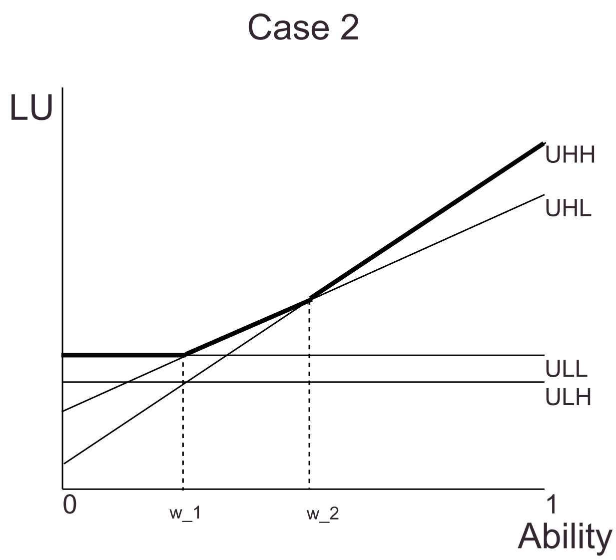


graph. Since the LH and LL are both independent of ability they both have zero slopes when graphed over ω. Thus, the outer envelope will only touch three of the four possible paths.[1] While for any particular initial path either ULH is always above ULL or ULL is always above ULH, this is not sufficient to show that one path dominates the other for all initial states. It may be the case that while ULL dominates ULH for one initial path, ULH dominates ULL for another initial path. This paper assumes that one path dominates the other for all initial paths. Hence, only three of the four states can exist in equilibrium.

Returning to the particulars of this specific case, were these three paths to be chosen in equilibrium, only high educated people would purchase homes in the"low" neighborhood while a mix of high and low skilled people enter the "high" neighborhood. If the labels mattered, then this situation could be ruled out the basis of consistency. In this case, the low neighborhood must have an average skill level that is less than or equal to that of the "high" neighborhood. However, microeconomic convention maintains that the labels are not so important. This convention implies that an equilibrium in which the low and high neighborhoods switch with each generation does exit. In this equilibrium, each new generation reverses the ranking of the neighborhoods.

[1]There is a set of parameters for which the two paths coincide. However, this extremely special case that can easily be ignored.

Figure 2:



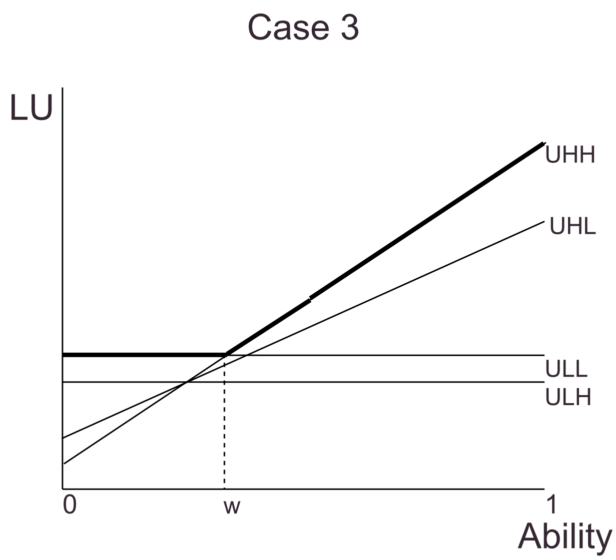
The good neighborhood becomes the bad and vice-versa. This equilibrium would require a degree of neighborhood rejuvenation and decline that is clearly counter-factual. For this reason, the equilibrium pictured in Case 1 is deemed to be "unstable", and ruled out.

Having ruled out the situation in Case 1, one can now turn Case 2 posited in Figure 2. Here the paths HH, HL and LL are all chosen in equilibrium, giving rise to a three state model. There is nothing to rule out this situation and it is examined in detail as the "Three Path Case." The states can be interpreted as the High, Middle and Low classes, respectively. This interpretation ranks the states in an intuitive fashion with HH at the top, HL in the middle and LL on the bottom. This ranking of the states is employed through out the paper. Of special interest are the conditions in which the "Middle class" of HL shrinks. And, where are the former residents of HL going?

Case 3, drawn in Figure 3, details a situation in which the outer envelope only touches two of four paths. In this case only the states HH and LH will be chosen for all initial states. If only these two paths are chosen, then everyone will go to the "high" neighborhood. If everyone goes to the "high" neighborhood then there will only be one neighborhood. The "One Neighborhood Case" is analyzed extensively in subsequent sections.

In Case 4, illustrated in Figure 4, the outer envelope only touches the paths HH and LL.

Figure 3:



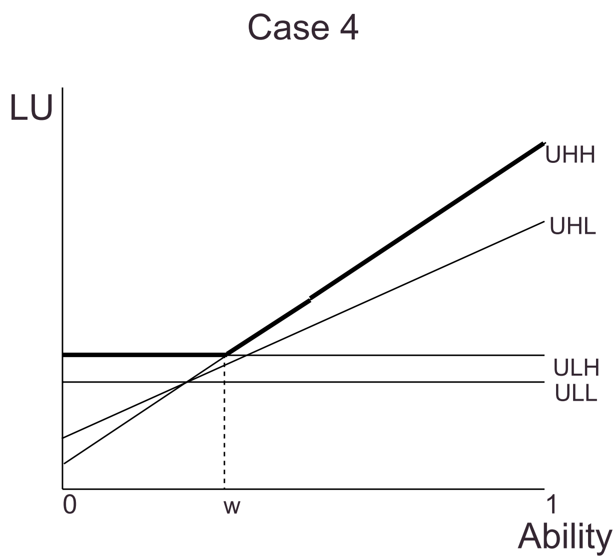
This would give rise to the case of complete economic segregation. In this two state model, individuals would either choose high education and a house in the high neighborhood or low education and a house in the low neighborhood. This situation may exist. However since this paper seeks to investigate the role of the middle class in generating "concentration effects", the two state model is not examined here.

2.3 Who Chooses Which State: The Three Path Case

Each person is born into a specific state with an ability level. In the Three Path Case, for each initial state, those with the lowest abilities will choose LL, the more able will prefer HL and the remaining most able will opt for HH. For each pair of contiguous states, there is a person with a unique ability level who is indifferent between the two states. This person is known as the cutoff person and her/his ability level is termed a cutoff ability level.

Another look at Figure 2 helps to clarify the point. The qualitative situation pictured in Figure 2 holds for all initial states, although the magnitudes and slopes differ. In the figure, those with an ability strictly below w_1 prefer LL, because it is gives them the highest lifetime utility. The person with ability w_1 is indifferent between the LL and HL paths because they offer exactly the same lifetime utility. Thus, w_1 is a cutoff level. Continuing out the ability

Figure 4:



axis, those with abilities strictly between w_1 and w_2 prefer the HL path, while w_2 is indifferent between HL and HH. Hence, w_2 is also a cutoff level.

In Figure 2 there are two cutoff levels. The generality of the initial state implies that there are two cutoff levels for each initial state. Thus, there are six cutoff levels for the Three State Case. The precise algebraic derivations of all six is found Appendix A.2.

While the cutoffs are defined as part of the solution to the utility maximization problem, the cutoffs also identify the conditional transition probabilities. Since ability is distributed uniform over $[0, 1]$ and i.i.d. over generations, the probability that someone born into LL will choose the LL path is ω_{LL1}. In general, let $Pr\{h_t N_t | h_{t-1} N_{t-1}\}$ be the probability of choosing the path $h_t N_t$ conditional on being born into $h_{t-1} N_{t-1}$. The assumption that ability is Uniform [0,1] and i.i.d. implies that

- $Pr\{LL|LL\} = \omega_{LL1}$
- $Pr\{HL|LL\} = \omega_{LL2} - \omega_{LL1}$
- $Pr\{HH|LL\} = 1 - \omega_{LL2}$
- $Pr\{LL|HL\} = \omega_{HL1}$

- $Pr\{HL|HL\} = \omega_{HL2} - \omega_{HL1}$
- $Pr\{HH|HL\} = 1 - \omega_{HL2}$
- $Pr\{LL|HH\} = \omega_{HH1}$
- $Pr\{HL|HH\} = \omega_{HH2} - \omega_{HH1}$
- $Pr\{HH|HH\} = 1 - \omega_{HH2}$

In any long run equilibrium, the number of people in each state must be given by the total number of people who choose to be in that state. For instance, the number of people in LL must be the sum of those who opt for LL and were born into the states LL, HL and HH. Define π_{LL} to be the fraction of the total population that chooses the LL path. Then

$$\pi_{LL} = Pr\{LL|LL\}\,\pi_{LL} + Pr\{LL|HL\}\,\pi_{HL} + Pr\{LL|HH\}\,\pi_{HH} \tag{7}$$

where π_{HL} is the fraction of the population that chooses the HL path and π_{HH} is the fraction of the population that chooses the HH path. The analogous constraints on the sizes of the other two states are

$$\pi_{HL} = Pr\{HL|LL\}\,\pi_{LL} + Pr\{HL|HL\}\,\pi_{HL} + Pr\{HL|HH\}\,\pi_{HH} \tag{8}$$

and

$$\pi_{HH} = Pr\{HH|LL\}\,\pi_{LL} + Pr\{HH|HL\}\,\pi_{HL} + Pr\{HH|HH\}\,\pi_{HH}. \tag{9}$$

Since the conditional probabilities have been found in terms of the cutoffs, these equations can be rewritten as

$$\pi_{LL} = \omega_{LL1}\pi_{LL} + \omega_{HL1}\pi_{HL} + \omega_{HH1}\pi_{HH}, \tag{10}$$

$$\pi_{HL} = (\omega_{LL2} - \omega_{LL1})\pi_{LL} + (\omega_{HL2} - \omega_{HL1})\pi_{HL} + (\omega_{HH2} - \omega_{HH1})\pi_{HH}, \tag{11}$$

and

$$\pi_{HH} = (1 - \omega_{LL2})\pi_{LL} + (1 - \omega_{HL2})\pi_{HL} + (1 - \omega_{HH2})\pi_{HH}. \tag{12}$$

Since the π's are population fractions, they must sum to one, i.e.

$$\pi_{LL} + \pi_{HL} + \pi_{HH} = 1. \tag{13}$$

Because the conditional probabilities for each state sum to one, equations (10) -(12) are linearly dependent. However, any two of these three equations and the condition that the population fractions sum to one, equation (13), uniquely determine the population fractions in terms of the cutoffs. The solutions are provided in Appendix A.3.

Since the population is normalized to 1, $\pi_{LL} + \pi_{HL}$ is the size of the low neighborhood and π_{HH} is the size of the high neighborhood. These population fractions also determine the labor supplies; π_{LL} is the supply of low skilled labor and $\pi_{HL} + \pi_{HH}$ is the supply of high skilled labor.

2.4 Who Chooses Which State: The One Neighborhood Case

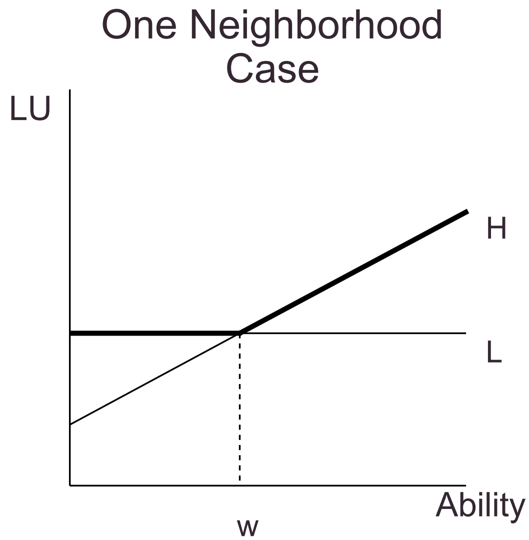
In the One Neighborhood Case, there are also cutoffs and transition probabilities. However, in the One Neighborhood Case, there are only two paths. Since everyone lives in the same neighborhood, the neighborhood effect is same for all and the cost functions only differ over the parent's skill choice.

Consider the problem of an individual with ability ω born into either initial state, in a One Neighborhood world. This person's problem is depicted in Figure 5. Being a Lifetime utility maximizer, this person will choose the higher of the two paths available. As in the Three Path Case, the highlighted outer envelope illustrates the choice for each person ω. Figure 5 also shows that there is a single ability level, w, at which the two lifetime utilities are equal. Everyone with an ability level less than w opts to remain low skilled, i.e. they choose the path L. Those with ability levels at least as large as w become high skilled by choosing the H path.

The equations that specify the lifetime utility available for each ability level and each initial state are given in Appendix A.4. This appendix also derives the equations that can be explicitly solved for the cutoffs for the low and high states which are ω_L and ω_H, respectively. As in the Three State Case, the cutoffs yield the conditional transition probabilities from one state to the other. Let $Pr\{h_t|h_{t-1}\}$ be the probability that someone born into the state h_{t-1} choose the state h_t. Then, in terms of the cutoffs, the conditional transition probabilities are

- $Pr\{L|L\} = \omega_L$

Figure 5:



- $Pr\{H|L\} = 1 - \omega_L$
- $Pr\{L|H\} = \omega_H$
- $Pr\{H|H\} = 1 - \omega_H$

In any long run equilibrium for the One Neighborhood Case, the number of people in each state must be given by those that choose to enter it. Thus, for the low state

$$\pi_L = \omega_L \pi_L + \omega_H \pi_H, \tag{14}$$

where π_L is the fraction of the population in the low state and π_H is the fraction of the population in the High state. There is a similar equation for π_H, but, like the Three Path Case, the fact that the conditional transition probabilities sum to one over each state makes it linearly dependent on Equation (14). However, the fact that the two populations fractions must sum to one provides a second linearly independent equation which is

$$\pi_L + \pi_H = 1. \tag{15}$$

The solution to these two equations and two unknowns is given in Appendix A.5. Since π_L is the fraction of the population that remains low educated, it gives the supply of low skilled labor. Similarly, π_H is the supply of high skilled labor.

2.5 The Expected Value Functions of the States in the Three Path Case

The people in this economy use dynamic programming to determine their optimal paths. From a solution standpoint, the key objects of interest are the expected value functions for each path. This section sets up the derivation of the equations that can be solved for the expected value functions in the Three Path Case. The actual equations are given in Appendix A.6.

The Optimality Equation of Dynamic Programming requires that the value of the LL path for person ω is

$$\begin{aligned} V_{LL}(\omega) = \max \Big\{ & W_L^\gamma + \beta \tilde{V}_{LL}, \\ & \left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega) \right)^\gamma + \beta \tilde{V}_{HL}, \\ & \left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega) + P_D(1-D_H) \right)^\gamma D_H^{1-\gamma} + \beta \tilde{V}_{HH} \Big\} \end{aligned} \tag{16}$$

Hence, the expected value function is

$$\tilde{V}_{LL} = E[V_{LL}(\omega)] = \int_0^1 V_{LL}(\omega) d\omega. \tag{17}$$

Since the cutoffs conveniently allow one to clear out the max function, the problem reduces to

$$\begin{aligned} \tilde{V}_{LL} = & \int_0^{\omega_{LL1}} \left[W_L^\gamma + \beta \tilde{V}_{LL} \right] d\omega \\ & + \int_{\omega_{LL1}}^{\omega_{LL2}} \left[\left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega) \right)^\gamma + \beta \tilde{V}_{HL} \right] d\omega \\ & + \int_{\omega_{LL2}}^1 \left[\left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega) + P_D(1-D_H) \right)^\gamma D_H^{1-\gamma} + \beta \tilde{V}_{HH} \right] d\omega. \end{aligned} \tag{18}$$

Similar arguments for the other two states imply that

$$\begin{aligned} \tilde{V}_{HL} = & \int_0^{\omega_{HL1}} \left[W_L^\gamma + \beta \tilde{V}_{LL} \right] d\omega \\ & + \int_{\omega_{HL1}}^{\omega_{HL2}} \left[\left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega) \right)^\gamma + \beta \tilde{V}_{HL} \right] d\omega \\ & + \int_{\omega_{HL2}}^1 \left[\left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega) + P_D(1-D_H) \right)^\gamma D_H^{1-\gamma} + \beta \tilde{V}_{HH} \right] d\omega. \end{aligned} \tag{19}$$

and

$$
\begin{aligned}
\tilde{V}_{HH} = \int_0^{\omega_{HH1}} & \left[(W_L + P_D(D_H - 1))^\gamma + \beta\tilde{V}_{LL}\right] d\omega \\
& + \int_{\omega_{HH1}}^{\omega_{HH2}} \left[\left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega) + P_D(D_H - 1)\right)^\gamma + \beta\tilde{V}_{HL}\right] d\omega \\
& \qquad + \int_{\omega_{HH2}}^{1} \left[\left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega)\right)^\gamma D_H^{1-\gamma} + \beta\tilde{V}_{HH}\right] d\omega. \qquad (20)
\end{aligned}
$$

Performing the integration in equations (18)-(20) yields a system of three equations and three unknowns that can be uniquely solved for the $\tilde{V}_{LL}, \tilde{V}_{HL}, \tilde{V}_{HH}$ for given values of the parameters and the other endogenous variables. The equations that result from the integration are given in Appendix A.6.

2.6 The Expected Values of the States: The One Neighborhood Case

As in the Three Path Case, the methodology for obtaining the Expected Value functions is to find the Value Function for each state using the Optimality Equation of Dynamic Programming and then use the cutoffs to eliminate the maximum operator leaving a simple integration calculation.

Using the Optimality Equation of Dynamic Programming, the value to person ω of being born into the low state is

$$V_L(\omega) = \max\left\{W_L^\gamma + \beta\tilde{V}_L, \left(W_H - \frac{W_H}{\bar{H}}(1-\omega)\right)^\gamma + \beta\tilde{V}_H\right\}. \qquad (21)$$

Hence, the Expected Value Function is

$$\tilde{V}_L = E[V_L(\omega)] = \int_0^1 V_L(\omega)d\omega. \qquad (22)$$

The cutoffs can be used to clear out the maximum operator, leaving

$$\tilde{V}_L = \int_0^{\omega_L} \left[W_L^\gamma + \beta\tilde{V}_L\right] d\omega + \int_{\omega_L}^1 \left[\left(W_H - \frac{W_H}{\bar{H}}(1-\omega)\right)^\gamma + \beta\tilde{V}_H\right] d\omega. \qquad (23)$$

Carrying out the integration yields that

$$
\begin{aligned}
\tilde{V}_L = \omega_L W_L^\gamma + \omega_L\beta\tilde{V}_L & \\
& + \frac{\bar{H}}{W_H(\gamma+1)}\left[W_H^{\gamma+1} - \left(W_H - \frac{W_H}{\bar{H}}(1-\omega_L)\right)^{\gamma+1}\right] + (1-\omega_L)\beta\tilde{V}_H \qquad (24)
\end{aligned}
$$

Similarly, the Optimality Equation of Dynamic Programming implies that the value of the high state to person ω is

$$V_H(\omega) = \max\left\{W_L^\gamma + \beta\tilde{V}_L, \left(W_H - \frac{W_H}{\tilde{H}\bar{H}}(1-\omega)\right)^\gamma + \beta\tilde{V}_H\right\}. \tag{25}$$

The corresponding Expected Value Function is

$$\tilde{V}_H = E[V_H(\omega)] = \int_0^1 V_H(\omega)d\omega. \tag{26}$$

Again, the cutoffs can be used to clear out the maximum operator, yielding

$$\tilde{V}_H = \int_0^{\omega_H}\left[W_L^\gamma + \beta\tilde{V}_L\right]d\omega + \int_{\omega_H}^1\left[\left(W_H - \frac{W_H}{\tilde{H}\bar{H}}(1-\omega)\right)^\gamma + \beta\tilde{V}_H\right]d\omega. \tag{27}$$

The result is that

$$\begin{aligned}\tilde{V}_H &= \omega_H W_L^\gamma + \omega_H\beta\tilde{V}_L \\ &\quad + \frac{\tilde{H}\bar{H}}{W_H(\gamma+1)}\left[W_H^{\gamma+1} - \left(W_H - \frac{W_H}{\tilde{H}\bar{H}}(1-\omega_H)\right)^{\gamma+1}\right] + (1-\omega_H)\beta\tilde{V}_H.\end{aligned} \tag{28}$$

Equations (24) and (28) can be uniquely solved for $\tilde{V}_H$ and $\tilde{V}_L$ for given values of the parameters and the other endogenous variables.

2.7 The Market for Dwelling Services

The production and pricing of dwelling services is assumed to be the same in either case. In contrast, the different partitions of the population over the states and distinct levels of dwelling services purchased per individual give rise to two aggregate demand equations. In the market for dwelling services, only the aggregate demand equation changes over the two cases.

The market for dwelling services is assumed to be perfectly competitive. The production side of the market is modeled by a single representative producer, known as "the Developer". The Developer has a Cobb-Douglass technology for providing dwelling services, D from low and high skilled labor, L_L and L_H, respectively. The production function is

$$D = S_D L^{\alpha_D} L_H^{1-\alpha_D} \tag{29}$$

where S_D is a scale factor for the production function and α_D is expenditure share of low skilled labor. S_D should be interpreted as a fixed capital endowment for the dwelling services sector.

Because the addition of a savings decision would significantly complicate the model, physical capital is taken as fixed.

Since the technology is constant returns to scale, profit maximization requires a zero profit assumption. Thus, it is assumed that this cost minimizing developer prices at unit cost:

$$P_D = \frac{1}{S_D}\left(\frac{W_L}{\alpha_D}\right)^{\alpha_D}\left(\frac{W_H}{1-\alpha_D}\right)^{1-\alpha_D} \tag{30}$$

In the Three Path Case, there are two levels of dwelling services. Without loss of generality, the level for the low neighborhood is normalized to one. It is assumed that exogenous zoning laws have set the level dwelling services for the high neighborhood at $D = D_H$. The level for D_H was set exogenously to facilitate the analysis of how different degrees of exclusivity in the high neighborhood affect the model. For nearly every level of D_H there is some mechanism via which the inhabitants would choose D_H. Although, some mechanisms are more credible than others, the goal is to focus the effects of raising the degree of exclusivity rather than the mechanism itself.

Since the size of the total population is normalized to 1, aggregate demand for dwelling services in the Three Path Case is given by

$$D_{D3} = (\pi_{LL} + \pi_{HL}) + \pi_{HH}D_H \tag{31}$$

Market clearing implies that

$$(\pi_{LL} + \pi_{HL}) + \pi_{HH}D_H = D. \tag{32}$$

In the One Neighborhood Case, everyone purchases an identical quantity of dwelling services. Convention and the need to compare this case with the Three Path Case, require that this quantity be normalized to one. Hence, aggregate demand for the One Neighborhood Case is

$$D_{D1} = 1 \tag{33}$$

So market clearing for this case requires that

$$1 = D. \tag{34}$$

2.8 The Market for the Consumption Good

The production and pricing of the Consumption is also the same in both the Three Path Case and the One Neighborhood Case. The complication that the aggregate demands are different is avoided by using Walra's Law to drop the consumption market clearing condition in each case. What remains of the Consumption Good sector is identical over the two cases.

The perfectly competitive representative firm that produces the consumption good X has Cobb-Douglass technology that enables it to produce X according to

$$X = S_X L^{\alpha_X} L_H^{1-\alpha_X} \tag{35}$$

where S_X is a scale factor for the production function and α_X is expenditure share of low skilled labor. As in the Dwelling services sector, the scale factor should be interpreted as a fixed capital endowment.

The consumption good is used as the numeraire. The zero profit condition implies that the producer prices at unit cost. Hence, cost minimization implies that

$$1 = \frac{1}{S_X} \left(\frac{W_L}{\alpha_X} \right)^{\alpha_X} \left(\frac{W_H}{1-\alpha_X} \right)^{1-\alpha_X} \tag{36}$$

2.9 The Labor Markets

Since the production side of both sectors is the same in each case, the factor demands are also the identical. The supplies of high and low skilled labor, however, differ over the two cases. Consider the Three State Case first. After deriving the demands for low and high skilled labor for each of the two sectors, one can incorporate the market clearing conditions for each type of labor. Since only those in state LL supply low skilled labor, market clearing requires that

$$\pi_{LL} = \frac{\alpha_D P_D D}{W_L} + \frac{\alpha_X X}{W_L}. \tag{37}$$

The first term on the RHS of equation (37) gives the demand for low skilled labor from the dwelling services sector. The second term is the demand for low skilled labor from the consumption sector. Since people in both HL and HH supply high skilled labor, market clearing for high skilled labor requires that

$$\pi_{HL} + \pi_{HH} = \frac{(1-\alpha_D) P_D D}{W_H} + \frac{(1-\alpha_X) X}{W_H}. \tag{38}$$

The first term on the RHS of equation(38) gives the demand for high skilled labor from the dwelling services sector. The second term is the demand for high skilled labor from the consumption sector.

In the One State Case, those in state L supply low skilled labor while those in H supply high skilled labor. Since the factor demands are the same as in the Three Path Case, market clearing for low skilled labor in the One Neighborhood Case is

$$\pi_L = \frac{\alpha_D P_D D}{W_L} + \frac{\alpha_X X}{W_L}. \tag{39}$$

And market clearing for high skilled labor requires that

$$\pi_H = \frac{(1-\alpha_D) P_D D}{W_H} + \frac{(1-\alpha_X) X}{W_H}. \tag{40}$$

2.10 Equilibrium

In both the Three Path Case and the One Neighborhood Case, the models are two sector general equilibrium models in which the constant returns to scale representative firms earns zero profits. An equilibrium for this class of models is a set of prices and quantities in which consumers maximize their utilities, firms minimize costs, firms earn zero profits and all markets clear. This section identifies the equations that are used in computing an equilibrium for each case. However, it is important to note that, in each case, the systems of equations are necessary but not sufficient conditions for an equilibrium. In both cases, there are restrictions on the endogenous variables that are only implicit in the equations. Specifically, utility maximization in the Three Path Case implies that the high cutoff must be greater than the low cutoff for each initial state. Also, in both cases, all the cutoffs must be in $[0, 1]$. If, for some parameters, the approximate numerical solution to the system of equations necessary for an equilibrium for that case violates either of these additional restrictions then that case can not be an equilibrium to the model for those parameters.

First, consider the equations necessary for an equilibrium of the Three State case. Consumer maximization places restrictions on the six cutoffs and the three expected value functions. The logic of the restrictions is developed in the text while the actual equations are given in equations (51)-(56), (71), (74), and (77) in Appendices A.2 and A.6. As shown in section 2.3, optimal behavior by the individual consumers implies that the population divides itself over the states according to the population fractions given in equations (57)-(59). These population fractions determine the total labor supplies for low and high skilled workers. The implications of cost minimization and market clearing are derived in sections 2.7 -2.9. The equations that must hold in equilibrium are (30), (32), (36), (37) and (38). This case is subject to both of the additional restrictions: the high cutoff must be greater than the low cutoff for each initial path and all the cutoffs must be in $[0, 1]$.

The numerical solutions discussed in the following section prove that an equilibrium exits for those parameters considered. The numerical experiments also show that there are parameterizations for which the Three Path Case is not an equilibrium for the model. For these parameters, one of the other cases, e.g. the Two Path Case, is an equilibrium.

Now consider the equations necessary for an equilibrium of the One Neighborhood Case. Consumer maximization places restrictions on the two cutoffs and the two expected Value functions. The logic of the restrictions is derived in the text. Equations (65), (68), (24), and (28) impose these conditions algebraically. Section 2.4 shows that equations (14) and (15) describe how the rational agents will allocate themselves over the two states. These population fractions determine the supplies of low and high skilled labor. The implications of cost minimization and market clearing are derived in sections 2.7 -2.9. The restrictions are imposed algebraically in equations (30), (34), (36), (39) and (40). This case is only subject to the additional restriction that the cutoffs be in $[0, 1]$.

There are 11 equations and 11 unknowns in the system of equations necessary for an equilibrium in the One Neighborhood Case. The numerical solutions discussed in the next section demonstrate that there are equilibria for the parameters considered. There are other parameterizations for which the One Neighborhood Case can not an equilibrium for this model because the cutoff lie outside $[0, 1]$.

3 Empirical Aspects

The net returns to education are increasing in the average level of human capital in the child's neighborhood. This result follows directly from the assumption that the time cost of education is decreasing in the average level of human capital in the child's neighborhood and the existence of an equilibrium. This assumption is an empirically testable hypothesis. An empirical project using data from the Panel Study on Income Dynamics, PSID, and its confidential geocode supplement is now underway to test this hypothesis.

In addition, if neighborhood choice and lifetime parental income are sufficiently correlated, then this model also implies that the net returns to education should be increasing in the lifetime income of one's parents. This hypothesis directly contradicts standard labor economic theory. According to the traditional theory, higher parental lifetime incomes allow children to loosen financial constraints and obtain more education than children of equal ability from less well off parents. Since there are decreasing returns to schooling, these children of better off parents should have more education but earn a lower return.[2] The author is currently investigating whether the returns to education are increasing or decreasing in parental lifetime income using data from the PSID.

4 Numerical Aspects

4.1 Solution Method

The equations for both cases were coded in GAUSS and solved using the Constrained Optimization module. This module uses the Sequential Quadratic Programming method.[3] The model was solved with the equations as a set of nonlinear equality constraints to a dummy optimization problem because it avoids the re-parameterization that is necessary when using a simple systems of equations solver and Cobb-Douglas production functions.[4] Inequality constraints were imposed to keep the solution algorithm from seeking outside of legal domains. The small number of endogenous variables and a good choice of additional constraints permitted the use of a Newton's method with full calculated Hessian matrix.

[2]See Willis [18] for the outline of a formal model that produces this result.

[3]For further details on the algorithm used see Schoenberg [16].

[4]When trying to solve a system of equations that involve endogenous variables to a fractional exponential power, the model must be re-parameterized to keep these variables from appearing in a form that might be negative. For instance, the equation

$$P_D = \frac{1}{S_D}\left(\frac{W_L}{\alpha_D}\right)^{\alpha_D}\left(\frac{W_H}{1-\alpha_D}\right)^{1-\alpha_D}$$

might be coded as

$$P_D = \frac{1}{S_D}\left(\frac{exp(W_L)}{\alpha_D}\right)^{\alpha_D}\left(\frac{exp(W_H)}{1-\alpha_D}\right)^{1-\alpha_D}$$

. The values for W_L and W_H can then be easily calculated afterwards.

Good starting values were found to important in quickly finding a solution. For this reason, an initial solution vector was found from a set of random starting values. This starting vector was then used to start each set of simulations. However, it was verified that, for assorted parameter values, random starting values produce the same solution vector as the initial solution vector.

4.2 Choice of Parameters

There are eight parameters which must be chosen before the numerical simulations can be performed. Five of the eight specify the two production technologies. Two of the five technological parameters are the expenditure shares of low skilled labor for the housing and consumption, α_D and α_X, respectively. The fixed capital endowments for the housing and consumption sector S_D, S_X, respectively, must also be set. Finally, $\tilde{H}$ is amount of human capital required to become high skilled. The budget share of consumption and the factor by which parents discount the expected the utility of their children pin down the preferences. The last parameter, D_H is socially determined.

There are wide ranges of the parameters that yield equilibrium solutions for both the One Neighborhood Case and the Three Path Case. However, there are parameters for which is not possible to solve either the One Neighborhood Case nor the Three Path Case. In addition, there are parameter choices for which it is possible to obtain numerical solutions to the Three Path Case that are not equilibria because they violate the restrictions on the cutoffs. Specifically, there are parameters which cause the Three State Model to reduce to the Two Path Case. The numerical simulations indicate that are four parameters which must be properly scaled with each other to obtain an equilibrium solution. They are $\tilde{H}$, D_H, S_D, and S_X. This result is to be expected. The population is scaled to one and is endogenously divided between low and high skilled labor. If D_H exceeds 1 by too much for the specified capital endowments, then it will not be feasible to produce the required amount of housing. $\tilde{H}$ scales the difference between low and high skilled labor. If $\tilde{H}$ is too high for specific values of the other three, then those born into the HH state will have too much of an advantage and the solution will not be an equilibrium. In this case the Three Path Case becomes the Two Path Case.

The fixed capital endowments determine which industry is more capital intensive. Similarly, the expenditures shares of low skilled labor specify that one industry is relatively intensive in low skilled labor and that the other is relatively intensive in high skilled labor. Extensive experiments found that as long as the capital endowments remain properly scaled to the other parameters, switching the capital intensity had no qualitative impact on the model. Reversing the relative intensities in low and high skilled labor reversed the changes in the high and low skilled wages for particular simulations but it did not alter any of the fundamental results on ranking the cases or the degree of inequality. These results are documented in supplemental numerical result appendix, available from the author upon request.

The expenditure shares of low skilled labor must be less than .5. Aside from this restriction, they are free. Values of .3 for α_D and .4 for α_X were chosen for the base case. Housing services was thus assumed to be relatively intensive in high skilled labor. This assumption can be reversed without changing the fundamental results of the model. Values of 4 and 3 were chosen for the fixed capital endowments of the housing and consumption sectors, respectively. Again, the assumption that housing is capital intensive relative to consumption can be reversed without altering any of the fundamental conclusions. The high skill level was set at 1.3. How the model behaves when $\tilde{H}$ is changed is the subject of further investigation.

Convention requires that the budget share of housing .3. The rate at which parents discount the utility of their children , β , was set to .8. Changes of about 25% in either of these values had no effect on any of the results. Finally, 1.2 was chosen as a base case value for the amount of housing services that must be purchased to enter the high neighborhood. How changes in the value affect the model is discussed in the next several sections.

5 Results

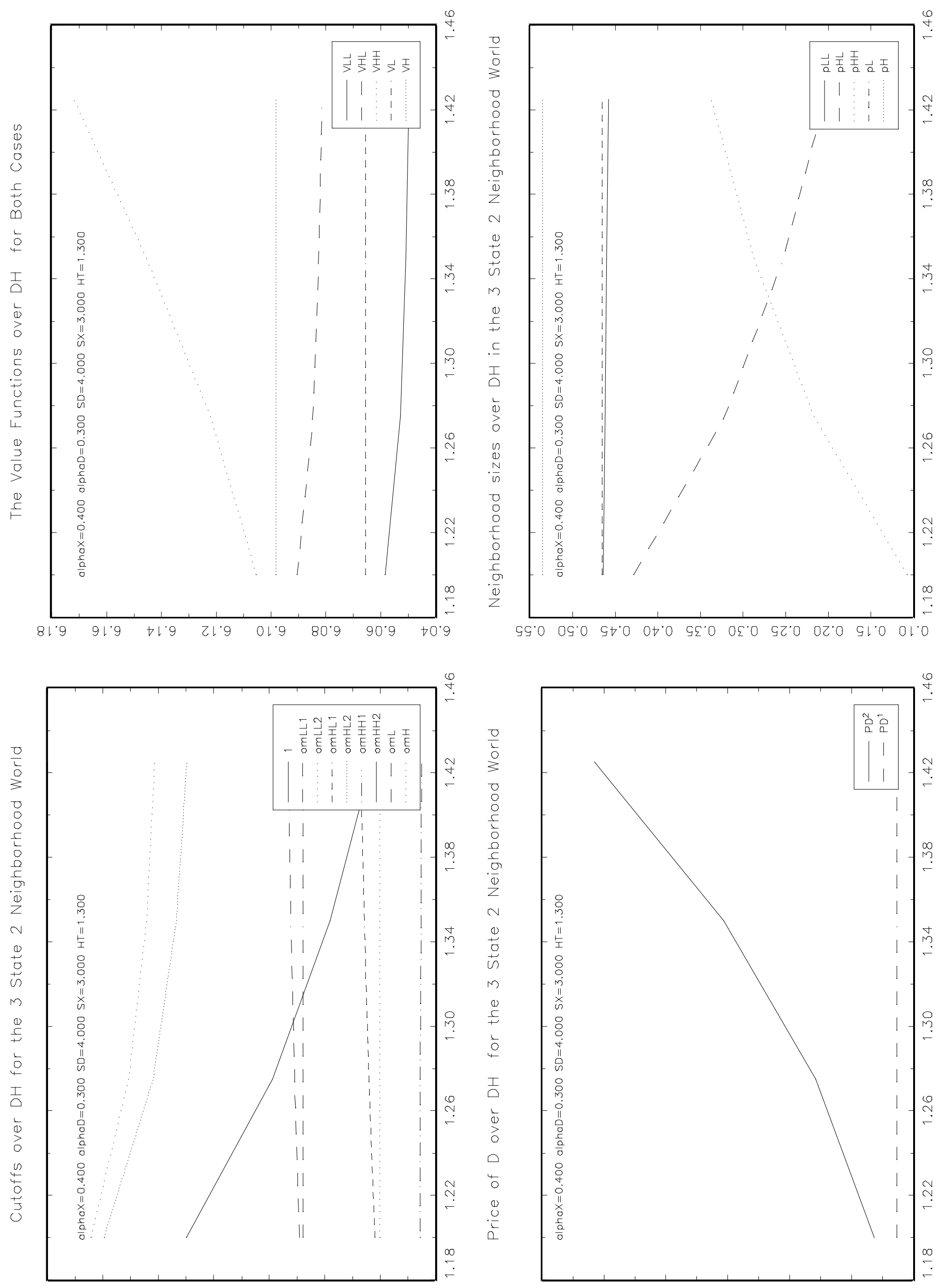
There are four sets of results, all which can be obtained for a wide range of parameters. The first set of results concerns the cutoffs. Both the low and the high cutoffs are strictly decreasing in the initial state. This implies that average level of ability in each skill level is strictly decreasing in the initial state. The second set revolves around who prefers which case. Those born into the low state or neighborhood prefer the single neighborhood case to the case of two neighborhoods. Furthermore, those born into the low neighborhood prefer that the zoning requirement for the high neighborhood be smaller. In contrast, those born into the high neighborhood strictly prefer high levels of exclusion. The third set asks whether there is more or less inequality. Inequality is measured by Lorenz dominance, Social mobility and second order stochastic dominance in the distribution of opportunity. It is found that there is more inequality in the Three Path case than when there is only one neighborhood. Moreover, inequality increases with increases in the zoning requirement for the high neighborhood. The fourth set of questions characterizes and explains the behavior of the model. The major result is that an increase in the exclusivity of the high neighborhood generates results like those described in *The Truly Disadvantaged* [19].

5.1 The Cutoffs and Average Ability

Employing the ranking of the states discussed in section 2.2, both the low and the high cutoffs are strictly decreasing in the initial state. To see this point,consider Figure 6. It is clear from the graph that ω_{LL1}, which is labeled as omLL1, is strictly greater than ω_{HL1}, which is labeled as omHL1. Similarly, ω_{HL1} is always above ω_{HH1}, which is labeled omHH1. The graph also shows an analogous ranking for $\omega_{LL2} > \omega_{HL2} > \omega_{HH2}$, which bear similar labels. These ranking results imply that the average level of ability in each state is decreasing in the resident's initial state. Since ability is distributed Uniform [0,1] in all the initial states, the average ability level of those in HH conditional on starting in state $h_t N_t$ is $a = \frac{1}{2}(1 + \omega_{h_t N_t})$. Thus, the average ability level of those in HH conditional on their starting in HH is $\frac{1}{2}(1 + \omega_{HH2})$. In contrast the avenge ability levels for those in HH from HL and LL, respectively, are $\frac{1}{2}(1 + \omega_{HL2})$ and $\frac{1}{2}(1 + \omega_{LL2})$. Since $\omega_{LL2} > \omega_{HL2} > \omega_{HH2}$ the result follows immediately. Similar results also hold for those in the HL and LL states.

There are two important points contained in these results. First, this model produces a situation in which a child of lesser ability born into better circumstances can go farther in life a child born into more difficult surroundings. This situation raises the issue of fairness. Second, the declining average ability level suggests that future versions of the model should consider production functions that depend on ability. If lower ability level of those in high skilled positions negatively affects production more than the presence of high ability people in low skilled positions raises production, then there will be efficiency gains in a fairer sorting process.

Figure 6:



The Value Functions over DH for Both Cases
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
VLL
VHL
VHH
VL
VH
Neighborhood sizes over DH in the 3 State 2 Neighborhood World
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
pLL
pHL
pHH
pL
pH
Cutoffs over DH for the 3 State 2 Neighborhood World
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
1
omLL1
omLL2
omHL1
omHL2
omHH1
omHH2
omL
omH
Price of D over DH for the 3 State 2 Neighborhood World
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
PD2
PD1

5.2 Who Likes Which Case?

If person ω, born into some state S, gets a higher lifetime utility in the One Neighborhood Case than in the Three Path Case, then person ω will prefer the former case. The only issue with these comparisons is that the states change between the cases. Someone born into the LL state in the Three Path Case would be born into L in the One Neighborhood Case. Hence, person ω born into LL will prefer whichever of the two cases that offers her/him a higher lifetime utility. Similarly, someone born into HH in the Three State Case would be born into H in the One Neighborhood Case. Person ω born into HH will also make her/his choice based lifetime utility. There is no direct analogue for HL in the one state case. Hence, there is no way of comparing the lifetime utility paths over the two cases for this group of people.

The graph in Figure 7 shows that for these parameters, those born into the LL state would almost unanimously prefer the One Neighborhood Case to the Three Path Case. To see why the very able from LL prefer the Three Path Case, consider the lifetime utility of those from LL who choose HH, given in equation (44). For those with an ability level near 1, the fall in lifetime utility due to a decrease in current consumption is less than the utility increase from more dwelling services and a better situation for their children. But, except for these values very close to one, the increase in the cost of becoming high skilled drives down the lifetime from consumption by more than enough to leave them strictly worse off. In particular, two points should be highlighted. First, those born into LL who stay in LL are worse off. Second, it is more difficult to escape LL in the Three Path Case than it is to escape L in the One Neighborhood Case. Thus, not only do higher levels of exclusivity make people born into LL who stay in LL worse off, the increases in D_H also make it increasingly difficult to escape LL.

In this parameterization, the wage for a low skilled worker is increasing in D_H. Qualitatively similar results are generated when dwelling services are relatively intensive in low skilled labor, implying that increases in D_H raise the low skilled wage. The fall in the lifetime utilities of the those born into LL who stay in LL is explained by the fact that $V_L L$ is smaller than V_L and that $V_L L$ is decreasing in D_H. This result is shown in the top left graph in Figure 6. The top left graph in Figure 6 illustrates why it is harder to escape LL in the Three Path Case than it is L in One Neighborhood case. The graph shows that the low cutoff ω_{LL1} for the Three Path Case is always above the One Neighborhood Case cutoff ω_L. The graph also shows that ω_{LL1} is strictly increasing in D_H, implying that increases in D_H make it more difficult escape LL.

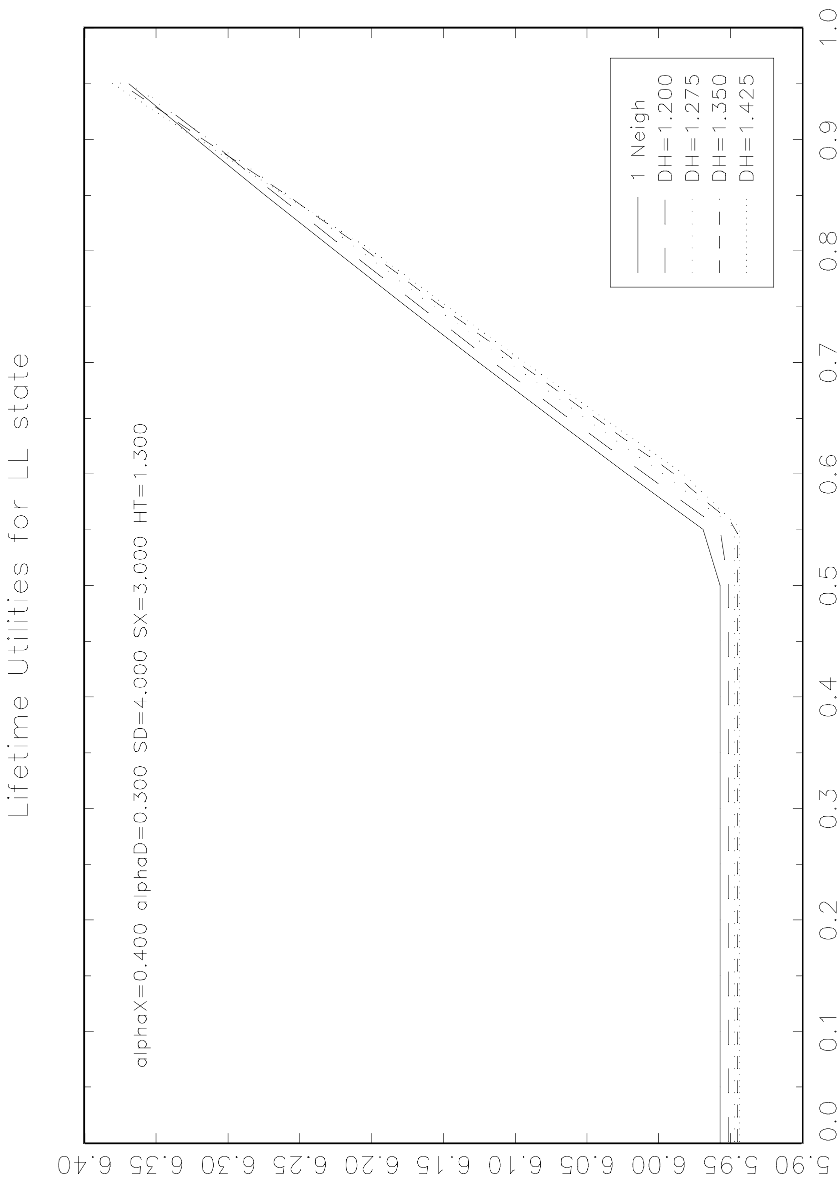
Increases in D_H cause the middle class of HL to flee the low neighborhood. The bottom left graph in Figure 6 shows that π_{HL} is strictly decreasing in D_H. This middle class flight from the low neighborhood lowers the average level of human capital in the low neighborhood, making it more difficult for the children of LL to move up. This result is one of the conclusions reached by Wilson [19]. This paper contributes to his discussion by adding a formal model of rational agents that generates the same process.

In contrast to those born into LL, the graph in Figure 8 shows that everyone born into HH would prefer the Three Path Case to the One Neighborhood Case. The wealth effect is the primary reason for this result. The wealth effect has two components an increase in the stock of the investment and an increase in its price. The quantity increase is due to exogenous change in the zoning laws. The bottom left graph in Figure 6 shows that the price of a unit of dwelling services is always higher in the Three Path Case than in the One Neighborhood Case. Furthermore, this price is strictly increasing in D_H. Clearly, these results are do to the increased demand for dwelling services. Both the increases in P_D and D_H raise the initial wealth of those born into HH. As can be seen from equation (48), since W_L and $V_L L$ are falling, this wealth effect is responsible for the upward shift in the lifetime utility paths of those born into HH.

A change from the One Neighborhood Case to the Three Path Case makes it significantly easier for someone born into HH to end up in one of the two higher states. Both the shape of the paths in Figure 8 and the fact that ω_{HH1} is below ω_H is the top left graph in Figure 6. Both of these graphs also show that changes in D_H have no significant changes on the low cutoff ω_{HH1}.

The dramatic decreases in the high cutoff for those born into HH seen in the top left graph in Figure 6 is deserves an explanation. Reconsider the Figure 2. For the HH initial state, lines UHH and UHL graph equation (50) and equation (49), respectively. Since increases in D_H have small effects on W_H, the slope of UHL is not going to change by much. Furthermore, as can be seen from the top right graph in Figure 6, the increase in the intercept of UHL due to increases in D_H is somewhat offset by a falling V_{HL}. In contrast to UHL, increases in D_H raise both the intercept and the slope of UHH. Furthermore, the top right graph in Figure 6 shows that V_{HH} is increasing in D_H. The result is that the distance between the intercepts of UHL and UHH is shrinking and that the slope of UHH is significantly increasing while the slope of

Figure 7:



Lifetime Utilities for LL state
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
1 Neigh
DH=1.200
DH=1.275
DH=1.350
DH=1.425
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
5.90 5.95 6.00 6.05 6.10 6.15 6.20 6.25 6.30 6.35 6.40

UHL rises only slightly. Imagining these changes to Figure 2 illustrates why increases in D_H cause such a large decreases in ω_{HH2}.

Unlike LL and HH, the HL path has no analogous counterpart in the One Neighborhood Case. Still, as can be seen from comparing Figures 9, 7, and 8, the HL is strictly below the both of the high and low paths in the One Neighborhood Case. Hence, those born into HL must have been better off under the One Neighborhood Case. Figure 9 also shows that, except for the very able, those born into HL prefer smaller levels of D_H. It also shows that increases in D_H increase the number of people born into HL that end up in LL. This last point can also be seen from the top left graph in Figure 6 that shows ω_{HL1} rising with D_H. So increases in the exclusivity of the high neighborhood make the majority of those who remain in the middle class worse off. These increases in ω_{HL1} are driven primarily by the rising cost of becoming high skilled as D_H rises. Children of high skilled parents that grow up in the low neighborhood have higher time costs of education because the average level of human capital in the neighborhood is falling with D_H. To see that the fall in the average level of human capital in the neighborhood is the culprit, reconsider Figure 2. The graphs of ULL and UHL are given by equations (45) and (46), respectively. Figure 9 shows that the intercept of ULL has moved down and that UHL has become steeper. Figure 2 implies that the only way to increase the low cutoff when ULL falls and the slope of UHL rises is for the intercept of UHL to fall by more than the intercept of ULL. The top right graph of Figure 6 shows that the distance between V_{LL} and V_{HL} is not noticeably changing. Furthermore, since the graph in Figure 9 does not change qualitatively by reversing the relative factor intensities, the changes in the wages are not driving the result.[5] Hence, neither the changes in wages nor in V_{LL} and V_{HL} can account for the fact that the intercept of UHL must fall by more that the intercept of ULL. But equation (46) shows that a fall in $\bar{H}_L$ will increase the slope and lower the intercept. Hence, it is the fall in the average level of human capital in the low neighborhood that explains the increase in the low cutoff for those born into HL.

So the vast majority of those born into LL are worse off in the Three Path Case than in the One Neighborhood Case. A standard question in Welfare Economics is whether or not there exists a compensation scheme that can leave everyone better off. In this case, the question is whether or not those born into HH could transfer some resources to those born into LL and

[5]The graph that verifies this statement is in an appendix available from the author upon request.

Figure 8:

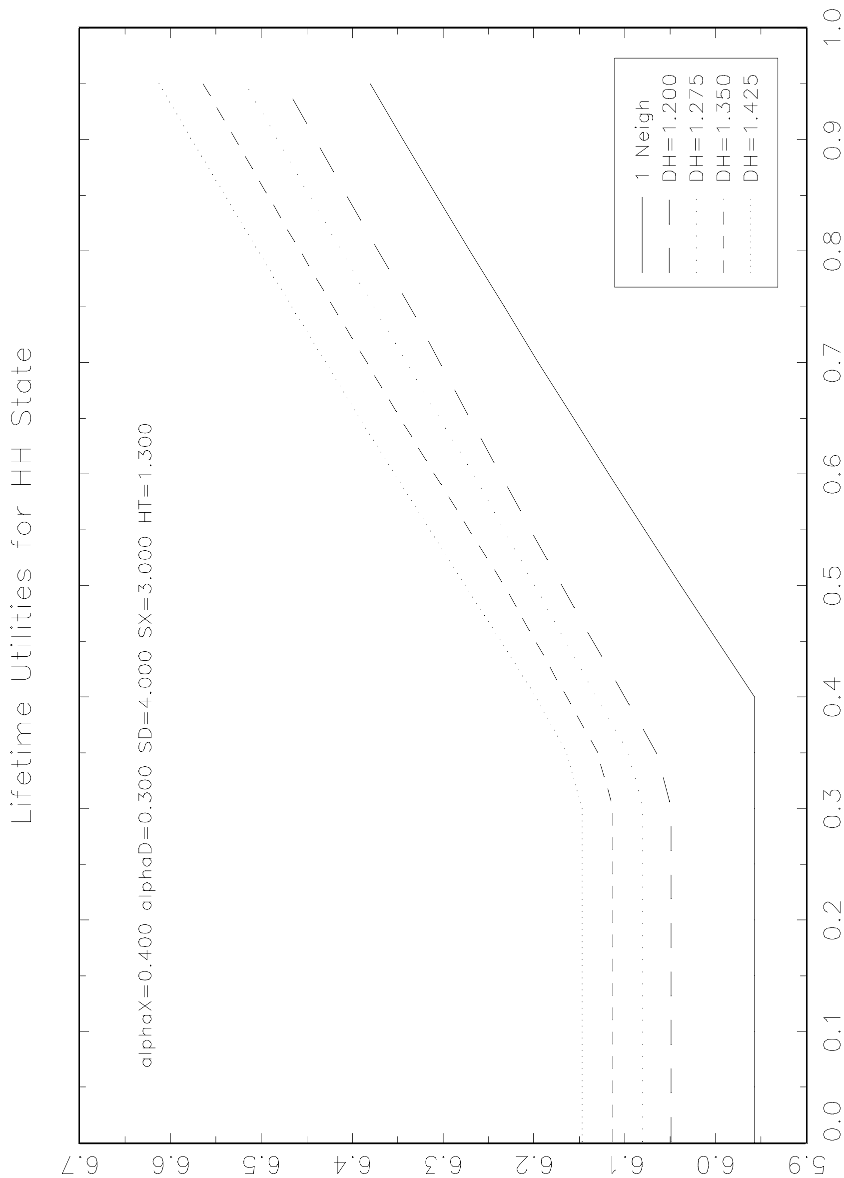
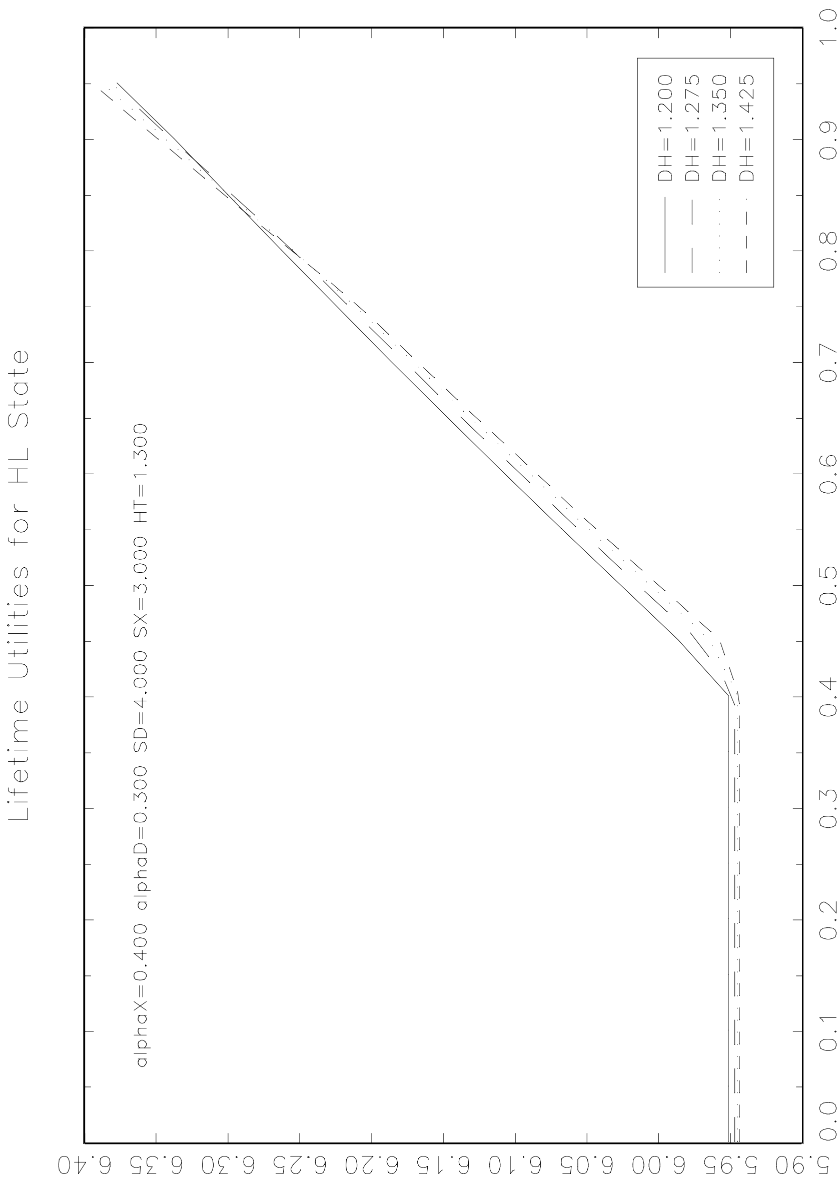

Lifetime Utilities for HH State
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
1 Neigh
DH=1.200
DH=1.275
DH=1.350
DH=1.425
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
5.9
6.0
6.1
6.2
6.3
6.4
6.5
6.6
6.7

Figure 9:



Lifetime Utilities for HL State
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
DH=1.200
DH=1.275
DH=1.350
DH=1.425
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
5.90 5.95 6.00 6.05 6.10 6.15 6.20 6.25 6.30 6.35 6.40

HL so that those born into HH would still prefer the Three Path Case while those born into LL and HL would be indifferent between the two cases. The large upward shifts in the HH and the small downward shifts in the LL and HL paths make the existence of such a scheme likely. However, if the residents of HH are fixing the level of D_H then the existence of such transfers mechanisms is not relevant. If the residents of HH can fix D_H on their own, then they do not need to compensate the other agents and will not unnecessarily relinquish any resources. Hence, assuming that the residents of HH unilaterally set D_H, one must conclude that those born into HH prefer the Three Path Case while the vast majority of the rest of the population prefers the One neighborhood Case.

5.3 Inequality Results

The question of how to measure inequality is difficult.[6] This model is robust enough to make strong statements about three measures of inequality. They are inequality in the distribution of income, social mobility and inequality in the distribution of opportunity. The Lorenz curve is used to measure income inequality. A measure of social mobility is derived using first order stochastic dominance of the conditional income distributions. A measure of the distribution of opportunity is derived based on second-order stochastic dominance of the distribution of the opportunity costs.

5.3.1 Lorenz Inequality of Lifetime Income

This paper considers inequality in the distribution of lifetime income, including an individual's initial wealth. Someone with an ability ω from state $h_{t-1}N_{t-1}$ that chooses the path h_tN_t has a lifetime income of

$$I_t = P_D D_{N_{t-1}} + W_{h_t}\left(1 - S_{h_t}\frac{1}{H_{h_{t-1}}\bar{H}_{N_{t-1}}}(1-\omega)\right) \tag{41}$$

where

$$S_{h_t} = \begin{cases} 1 & \text{if } S_{h_t} = H \\ 0 & \text{Otherwise} \end{cases}.$$

Inequality in the distribution of income is measured by the Lorenz Curve. For any two parameterizations of the model, A and B, if the Lorenz curve that results from A strictly dominates

[6] See [15, Sen] or [10, Lambert] for good introductions to this issue.

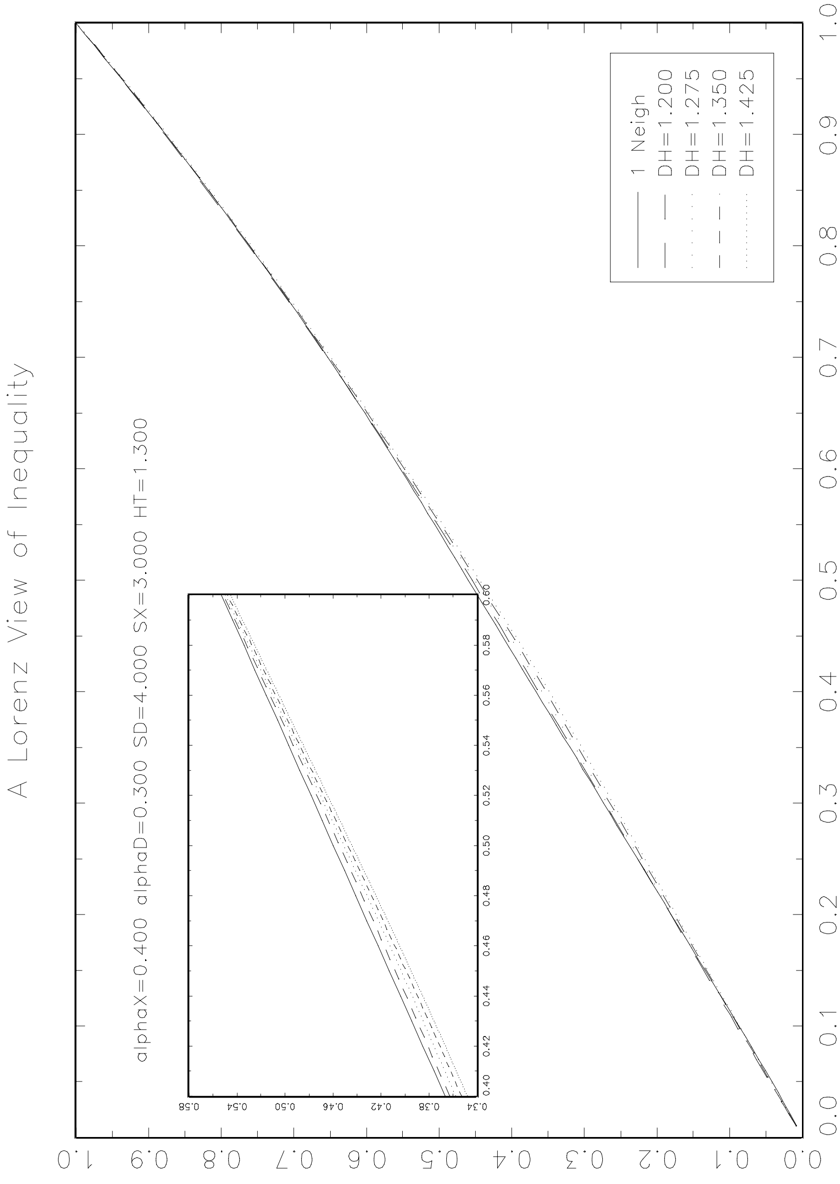
that of B, then the distribution of income is less equal under B than A.

Since ability affects income and ability is continuously distributed, the distribution of income is continuous. The unconditional distribution was approximated and the Lorenz calculated by treating this approximation as an empirical income distribution. The approximation algorithm is described in Appendix A.7. The resulting graph of the Lorenz Curves, Figure 10, shows that the Lorenz Curve for the One Neighborhood case strictly dominates that of the Three Path Case. It also shows that as D_H grows, each experiment generates a Lorenz Curve which is strictly dominated by the previous one.

5.3.2 Social Mobility

Income mobility involves the distribution of the child's income, conditional on that of the parent. If there is more upward movement from the lower incomes and more downward movement from the upper incomes in A than in B, then there is more social mobility in A than in B. However, there is no generally agreed upon fashion to make these comparisons for continuous income distributions. While a general solution to this problem lies beyond the scope of this paper, a practical solution is readily available. There is general agreement on how to how measure upward and downward movement conditional upon a parental income. If the cumulative distribution of income conditional upon a parent's income at X under the parameterization A first order stochastically dominates the same conditional income distribution under B, then there is more upward movement in A at X. Equivalently, the same condition implies that there is more downward movement in B than in A. Hence, first order stochastic dominance provides the comparison of upward and downward movement, conditional on any particular parental income. The question is what parental high and low incomes to choose. The top and bottom quartiles suggest themselves as good candidates. But these centiles would frequently pick a member of the HL path as the representative of the top part of the distribution. This choice would create an "apples and oranges" problem because the children of parents from H do better in a One Neighborhood world than the children of parents from HL in the Three Path Case. It is important to pick the upper centile so that it identifies a member of the highest class in each case, i.e. a member of H in the One Neighborhood Case and someone from HH in the Three Path Case. For this reason, the bottom quartile and the top tenth are chosen as the parental low

Figure 10:


A Lorenz View of Inequality
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
1 Neigh
DH=1.200
DH=1.275
DH=1.350
DH=1.425

and high incomes, respectively.[7] To simplify the final statement, let $F_A(I_t|I_{t-1} = X\%)$ be the cumulative distribution of the child conditional on the parent earning the Xth percentile under parameterization A. Similarly define $F_B(I_t|I_{t-1} = X\%)$ be the cumulative distribution of the child conditional on the parent earning the Xth percentile under parameterization B. Finally, for any two parameterizations of the of the model A and B, there is less social mobility under the parameterization B than under A if

1. if $F_A(I_t|I_{t-1} = 25\%)$ first order stochastically dominates $F_B(I_t|I_{t-1} = 25\%)$
2. if $F_A(I_t|I_{t-1} = 90\%)$ is first order stochastically dominated by $F_B(I_t|I_{t-1} = 90\%)$.

One needs two types of information to implement this social mobility ranking, the centiles and the graphs of the appropriate conditional cumulative distribution functions (or conditional CDF's). The 25th and 90th centiles are calculated from the approximated unconditional distribution of income found above. The analytic conditional CDF's for each parental income are derived in Appendix A.8.1 and Appendix A.8.2. The incomes from the 25th and 90th centiles determine which of the conditional CDF's is appropriate. Since the conditional CDF's depend on ω, they are continuous. Hence the graphs of these conditional CDF's are approximations.[8]

Applying the above social mobility ranking, there is less social mobility in the Three Path Case than in the One Neighborhood Case if

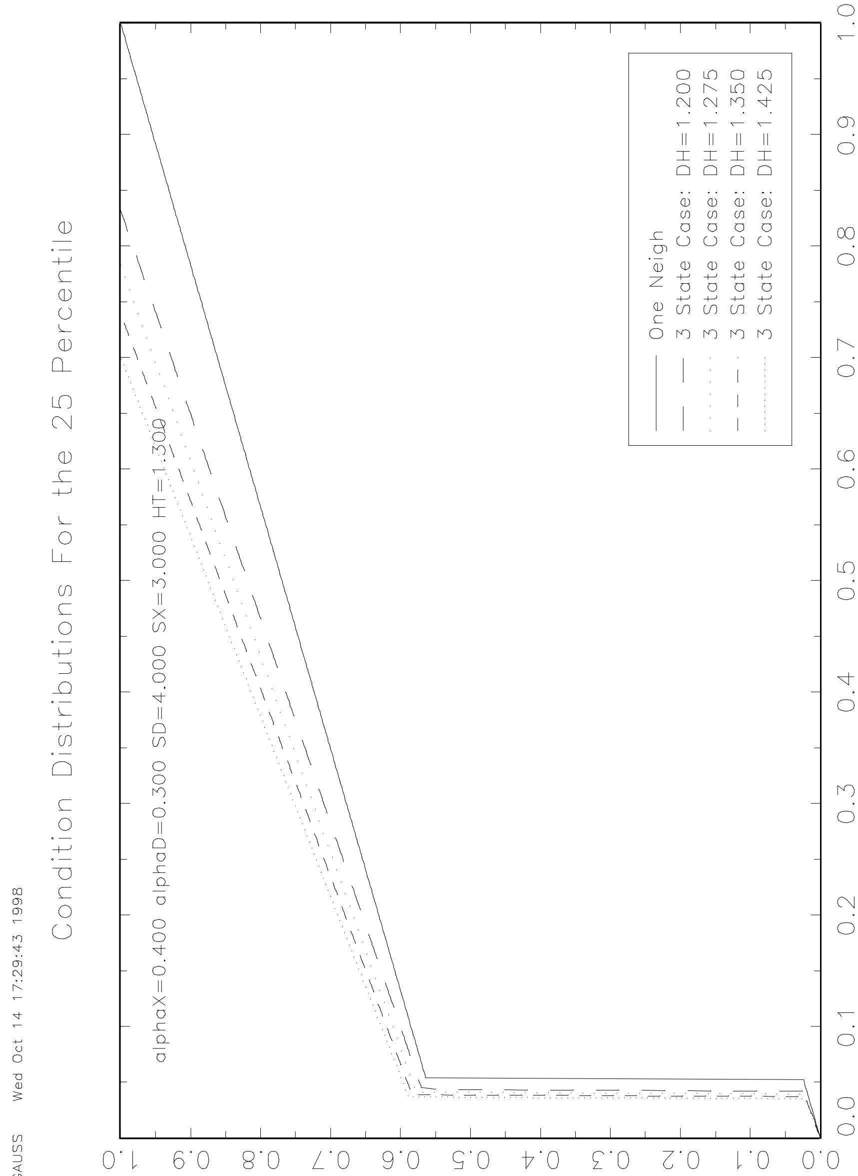
1. the conditional CDF of the child of the parent who earns 25th centile income in the One Neighborhood Case first order stochastically dominates the same conditional CDF derived from the Three Path Case
2. the conditional CDF of the child of the parent who earns the 90th centile income in the One Neighborhood Case is first order stochastically dominated by the same conditional CDF derived from Three State Case.

Figures 11 and 12 shows that both conditions hold. In Figure 11 the CDF of a child's income, conditional on the parent earning the 25th percentile income for the One Neighborhood

[7] There are parameterizations of the model for which the 90th centile comes from the H state in the One Neighborhood Case and from the HL in state in the Three Path Case. For these parameterizations, the second condition does not hold because the upper centile has identified a member of the middle, not the upper, class. It does hold at higher centiles in which the income is from the HH state.

[8] 1,000 points in [0,1] are used in these approximations.

Figure 11:


GAUSS Wed Oct 14 17:29:43 1998
Condition Distributions For the 25 Percentile
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
One Neigh
3 State Case: DH=1.200
3 State Case: DH=1.275
3 State Case: DH=1.350
3 State Case: DH=1.425
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

case, $F_1(I_t|I_{t-1} = 25\%)$ first order stochastically dominates those from the Three Path Case for assorted values of D_H. Figure 12 shows that second condition also holds. In Figure 12, the conditional CDF of children born to parents in the 90th percentile in the Three State Case, for various values of D_H, stochastically dominate the same conditional CDF in the One Neighborhood case. Since both conditions hold, there is less social mobility in the Three Path Case than in the One Neighborhood Case.

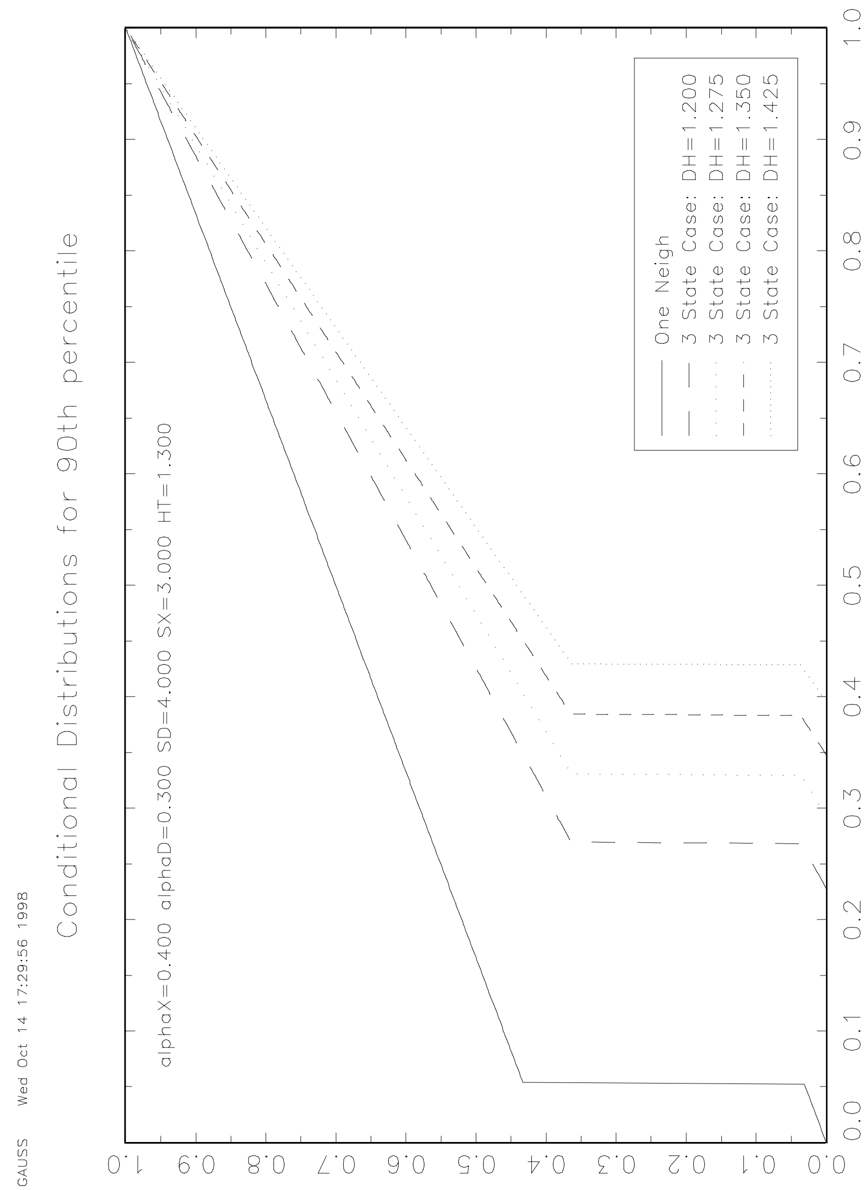
But these graphs also show that increases in D_H lower social mobility. Figure 11 shows that an increase in D_H yields a CDF conditional on the parent earning the 25th percentile which is first order stochastically dominated by the same conditional CDF for the lower value of D_H. Figure 12 shows that an increase in D_H yields a CDF conditional on the parent earning the 90th centile which is first order stochastically dominate the same conditional CDF at a lower value of D_H. Thus, the figures show that both conditions hold for increases in D_H.

5.3.3 Inequality of Opportunity

As noted above, this paper identifies the "opportunity" of an individual with the multiplicative inverse of her/his opportunity cost of becoming high educated. This definition permits the derivation of the analytic cumulative distribution functions of opportunity for the One Neighborhood and Three Path cases. These derivations can be found in Appendix A.9. These CDF's are graphed in Figure 13. The graph shows that those at the top have much more opportunity in the Three Path Case than when there is only one neighborhood. In contrast, those at the bottom end have significantly less opportunity in the Three Path Case than in the One Neighborhood Case. Moreover, increases in D_H augment the differences in both of these comparisons. However, the mean is slightly higher in the Three Path Case than in the One Neighborhood Case. Furthermore, the mean in the Three Path Case is slowly increasing in D_H.

If one were to perform linear transformations on the Three Path CDF's so that they had the same mean as the One Neighborhood CDF, then the CDF for the One Neighborhood Case would second order stochastically dominate the transformed CDF's from the Three Path Case. Ranking these cases requires some assumption about relative weights in a mean-variance tradeoff. Moreover, it is not possible to make a statement about "inequality" per se, only how society would rank the outcomes. Since the increases in the mean are so small relative to the standard deviation, if the variance weight is slightly larger than the mean weight, then then society

Figure 12:



GAUSS Wed Oct 14 17:29:56 1998
Conditional Distributions for 90th percentile
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
One Neigh
3 State Case: DH=1.200
3 State Case: DH=1.275
3 State Case: DH=1.350
3 State Case: DH=1.425
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

would prefer the Three Path Case to the One Neighborhood Case. Furthermore, these same assumptions imply that an increase in D_H generates a less preferable distribution of opportunity.

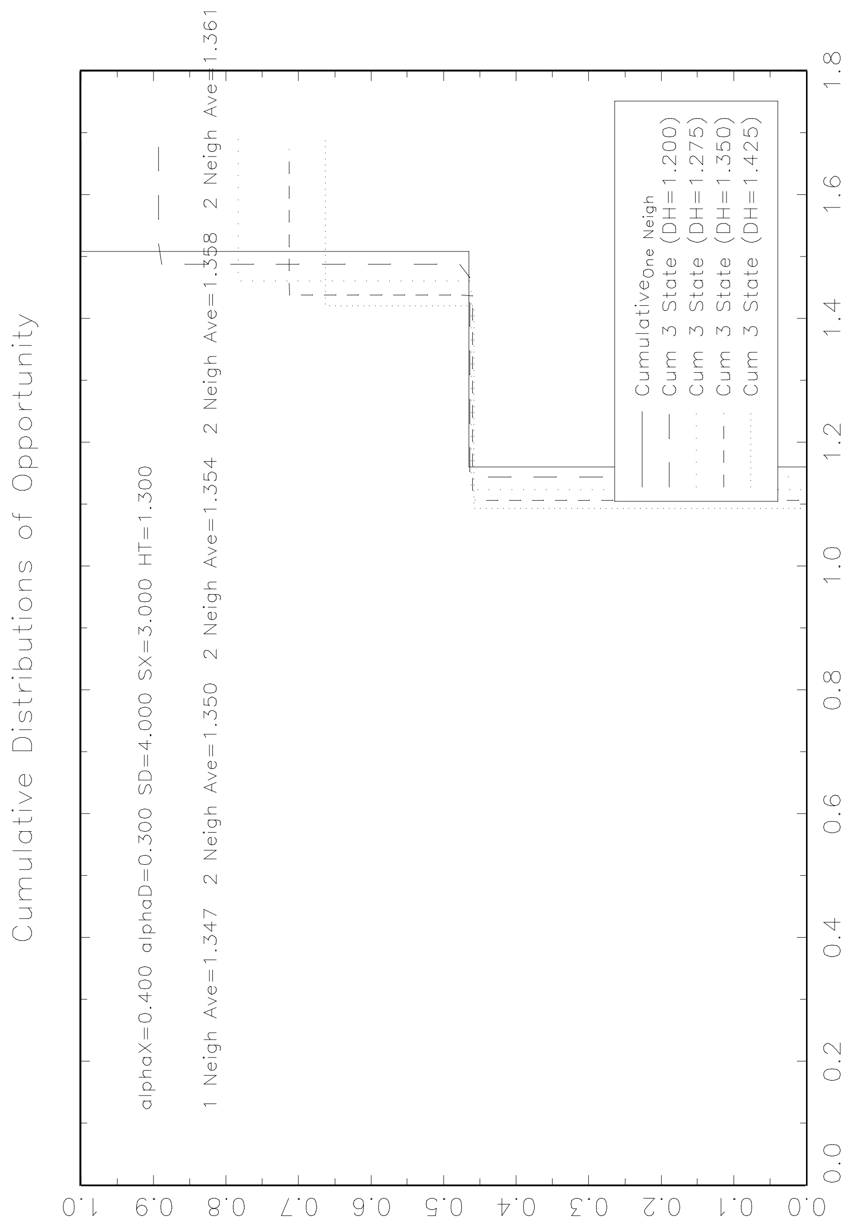
5.4 General Characterization

The behavior of the model is similar in many respects to the account in the *The Truly Disadvantaged*. The bottom right hand graph in Figure 6 graphs the fractions of the population in each state over the degree of restrictiveness of the high neighborhood. In the graph, pLL is the fraction of the population in the LL state, pHL gives fraction of the population in the HL state and pHH is defined in similar manner. The graph shows that as D_H increases the middle class of HL gets smaller. The increases in D_H are causing the middle class to flee the low neighborhood. The bottom left graph of Figure 6 shows that the price of dwelling services is rising with increases in D_H. So part of the movement is due to a wealth effect. The difference in the initial wealths of those born into the low neighborhood and those born into the better neighborhood is growing because of an increase in the difference of the initial stocks and an increase in the price of the asset. The top left graph in Figure 14 shows the difference between the equilibrium high wage in the One Neighborhood Case and the equilibrium high wage in the Three Path Case for increasing values of D_H. The negative values on the vertical axis imply that the high wage is higher in the Three Path Case than in the One Neighborhood Case. The downward slope implies that the wage of the high skilled is rising as D_H increases. Since dwelling services is relatively intense in high skilled labor compared to the consumption good and both of these changes entail an producing more dwelling services, both of these results on the behavior of the real wage are to be expected.

The top right graph in Figure 14 shows the difference between the equilibrium low wage in One Neighborhood Case and the equilibrium low wage in the Three Path Case. The positive values on the vertical axis indicate that the low wage is lower in the Three Path Case than in the One Neighborhood Case. The upward slope implies that the wage of the low skilled is falling as D_H increases. Again, since the changes shift production into the good which uses high skilled labor relatively intensively, these results are to be expected.

Hence, the difference between the low and high skilled wages in growing, giving rise to an earned income effect. The wealth effect and the earned income effect combine to raise the expected value of a spot in the high neighborhood and to lower the expected value of a spot in

Figure 13:


Cumulative Distributions of Opportunity
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
1 Neigh Ave=1.347 2 Neigh Ave=1.350 2 Neigh Ave=1.354 2 Neigh Ave=1.358 2 Neigh Ave=1.361
Cumulative One Neigh
Cum 3 State (DH=1.200)
Cum 3 State (DH=1.275)
Cum 3 State (DH=1.350)
Cum 3 State (DH=1.425)
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

the low neighborhood. In the *The Truly Disadvantaged*, this growing wage gap is an important factor in the analysis. It is interesting to note that none of the fundamental results regarding the preference rankings of cases, nor the results on increasing inequality depend on the presence of the earned income effect.[9]

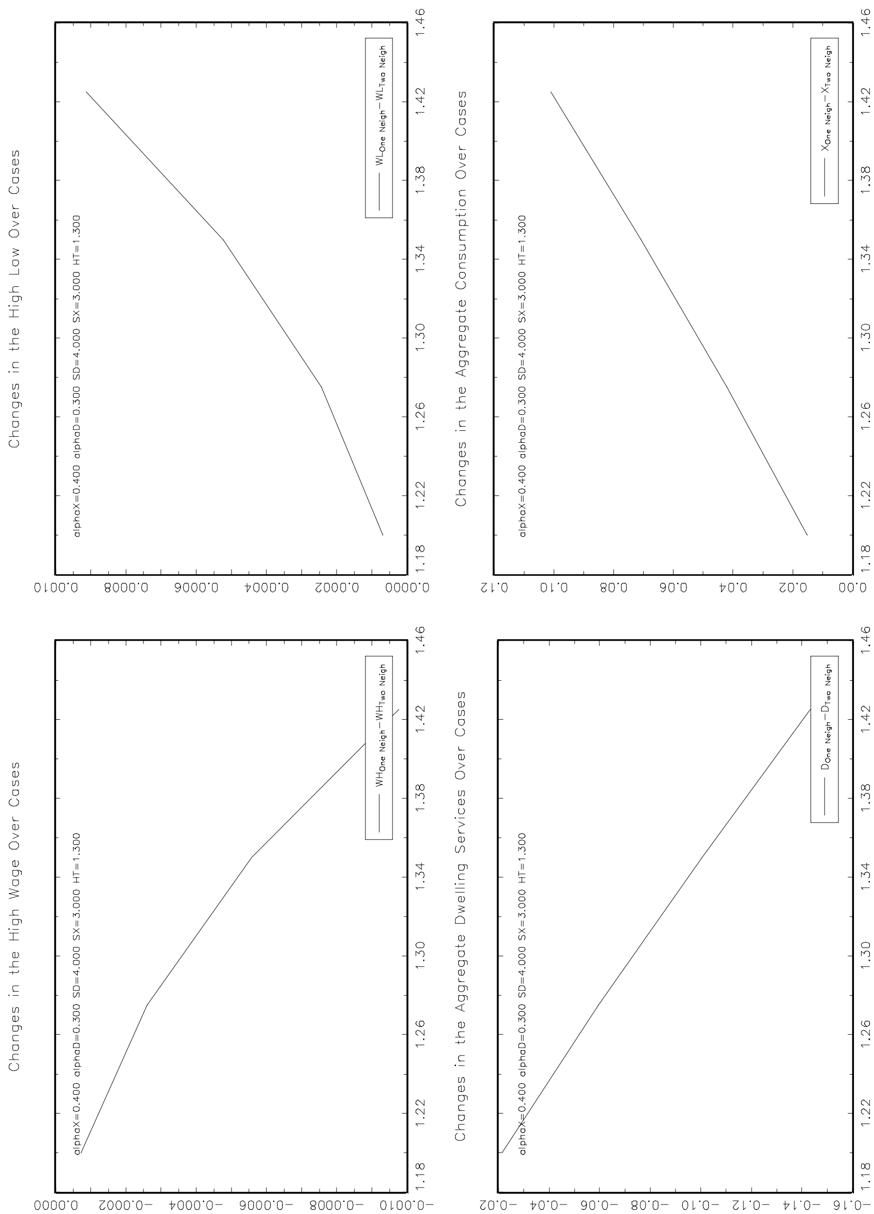
6 Conclusion and Further Work

For a wide range of parameters, the simulations show that nearly everyone born into the low neighborhood prefers the case of one neighborhood to the case in which there are two neighborhoods. Also shown that those born into the better neighborhood prefer the case of two neighborhood to the case on only one. The middle class present in the Three Path Case prefer the One Neighborhood Case. With respect to inequality, the two neighborhood case has more inequality in the distribution of income, less social mobility and more inequality in the distribution of opportunity than the one neighborhood case. Furthermore, increases in the exclusivity of the high neighborhood, increase inequality in the distribution of income, increase inequality in the distribution of opportunity and reduce social mobility. Finally, this model can produce results similar to the account in *The Truly Disadvantaged.* Increases in the exclusivity of the high neighborhood cause the middle class of high skilled residing in the low neighborhood to flee, raising the cost of education for those left behind. This result works through a wealth effect in the quantity and price of housing. It does not depend on there being an increase in the difference between the low and high skilled wages.

This model can be extended to include an endogenous savings decision. Not only will this extension remove the assumption of fixed capital stocks but it will open up the application of this model to questions of growth and development. The conjecture to be investigated is that countries with more exclusive neighborhood structures restrict social mobility and thereby inhibit growth, even when physical capital is endogenously accumulated.

[9] Graphs confirming this result are in an appendix available from the author on request.

Figure 14:



Changes in the High Low Over Cases
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
WL One Neigh−WL Two Neigh
Changes in the Aggregate Consumption Over Cases
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
X One Neigh−X Two Neigh
Changes in the High Wage Over Cases
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
WH One Neigh−WH Two Neigh
Changes in the Aggregate Dwelling Services Over Cases
alphaX=0.400 alphaD=0.300 SD=4.000 SX=3.000 HT=1.300
D One Neigh−D Two Neigh

A Derivations

A.1 Lifetime Utilities in the Three Path Case

Employing the definitions in the text, this subsection details the lifetime utilities that accrue to each person in each initial state.

Consider first those born into the LL state. If someone born into LL chooses to the LL path then s/he will obtain a the lifetime utility of

$$LU(\omega, LL, LL) = W_L^\gamma + \beta \tilde{V}_{LL}. \tag{42}$$

As noted in the text, the level of utility is independent of ability. If someone from LL opts for HL then their lifetime utility is given by

$$LU(\omega, LL, HL) = \left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega) \right)^\gamma + \beta \tilde{V}_{HL}. \tag{43}$$

Finally, those from LL who choose the HH path obtain

$$LU(\omega, LL, HH) = \left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega) + P_D(1-D_H) \right)^\gamma D_H^{1-\gamma} + \beta \tilde{V}_{HH}. \tag{44}$$

Those born into the HL state have the benefit of a lower time cost of becoming high skilled. However, note that someone from HL who choose LL earns the same lifetime utility as some from LL. The utilities are the same because the states only differ in there time cost functions which are not relevant for those who choose to remain low skilled. Hence,

$$LU(\omega, HL, LL) = W_L^\gamma + \beta \tilde{V}_{LL}. \tag{45}$$

Someone born into HL who chooses HL obtains

$$LU(\omega, HL, HL) = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega) \right)^\gamma + \beta \tilde{V}_{HL}. \tag{46}$$

Person ω born into HL who opts for HH earns

$$LU(\omega, HL, HH) = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega) + P_D(1-D_H) \right)^\gamma D_H^{1-\gamma} + \beta \tilde{V}_{HH}. \tag{47}$$

Those born into the HH state have lower time cost of becoming high skilled and the Wealth effect of inheriting a house in the high neighborhood. The wealth effect causes the lifetime utility of those that remain low skilled to be higher for those born into HH than for those who begin life in either of the other two initial states. This can be seen by comparing the lifetime utility for those from HH who remain low skilled, which is

$$LU(\omega, HH, LL) = (W_L + P_D(D_H - 1))^\gamma + \beta\tilde{V}_{LL} \tag{48}$$

with the utilities of those that opt for the LL that start off in LL and HL. Both the wealth effect and the lower time cost benefit those born into HH that choose HL. An person ω from HH who chooses HL obtains

$$LU(\omega, HH, HL) = \left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega) + P_D(D_H - 1)\right)^\gamma + \beta\tilde{V}_{HL}. \tag{49}$$

Both benefits also evident for those from HH who choose HH. Person ω from HH who goes HH obtains

$$LU(\omega, HH, HH) = \left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega)\right)^\gamma D_H^{1-\gamma} + \beta\tilde{V}_{HH}. \tag{50}$$

A.2 The Cutoffs for the Three Path Case

In this section the six cutoffs for the Three Path Case are derived. As noted in the text, for each initial state there are two cutoffs. The first cutoff equates the lifetime utilities from the LL path and the HL path. For the initial state $H_{t-1}N_{t-1}$, this first cutoff is denoted $\omega_{H_{t-1}N_{t-1}1}$. For instance, the fist cutoff for the LL is denoted ω_{LL1} and it is defined as the real number between 0 and 1 that solves

$$W_L^\gamma + \beta\tilde{V}_{LL} = \left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega_{LL1})\right)^\gamma + \beta\tilde{V}_{HL}. \tag{51}$$

For each initial state $H_{t-1}N_{t-1}$, the second cutoff is denoted $\omega_{H_{t-1}N_{t-1}2}$. This cutoff equates the lifetime utilities of the HL and HH paths. For instance, ω_{LL2} is defined to be the real number between zero and one that solves

$$\left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega_{LL2})\right)^\gamma + \beta\tilde{V}_{HL} = \left(W_H - \frac{W_H}{\bar{H}_{N_L}}(1-\omega_{LL2}) + P_D(1-D_H)\right)^\gamma D_H^{1-\gamma} + \beta\tilde{V}_{HH}. \tag{52}$$

The cutoffs for the HL state are similarly defined. In particular, ω_{HL1} solves,

$$W_L^\gamma + \beta\tilde{V}_{LL} = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega_{HL1})\right)^\gamma + \beta\tilde{V}_{HL}. \tag{53}$$

The second cutoff, ω_{HL2} solves

$$\left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega_{HL2})\right)^{\gamma} + \beta\tilde{V}_{HL} = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}_{N_L}}(1-\omega_{HL2}) + P_D(1-D_H)\right)^{\gamma} D_H^{1-\gamma} + \beta\tilde{V}_{HH}. \quad (54)$$

The first cutoff for the HH state, ω_{HH1}, solves

$$(W_L + P_D(D_H-1))^{\gamma} + \beta\tilde{V}_{LL} = \left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega_{HH1}) + P_D(D_H-1)\right)^{\gamma} + \beta\tilde{V}_{HL}. \quad (55)$$

The second, , ω_{HH2}, solves

$$\left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega_{HH2}) + P_D(D_H-1)\right)^{\gamma} + \beta\tilde{V}_{HL} = \left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega_{HH2})\right)^{\gamma} D_H^{1-\gamma} + \beta\tilde{V}_{HH}. \quad (56)$$

A.3 Solving for the Population Fractions

Solving equations (10), (12) and (13) for the population fractions in terms of the cutoffs yields the following solutions:

$$\pi_{LL} = \frac{\omega_{HH1} + \omega_{HH2}\omega_{HL1} - \omega_{HH1}\omega_{HL2}}{Z} \quad (57)$$

$$\pi_{HL} = \frac{-\omega_{HH1} + \omega_{HH2} - \omega_{HH2}\omega_{LL1} + \omega_{HH1}\omega_{LL2}}{Z} \quad (58)$$

$$\pi_{HH} = \frac{1 + \omega_{HL1} - \omega_{HL2} - \omega_{LL1} + \omega_{HL2}\omega_{LL1} - \omega_{HL1}\omega_{LL2}}{Z} \quad (59)$$

where

$$Z = 1 + \omega_{HH2} + \omega_{HL1} + \omega_{HH2}\omega_{HL1} - \omega_{HL2} - \omega_{HH1}\omega_{HL2} - \omega_{LL1} - \omega_{HH2}\omega_{LL1} + \omega_{HL2}\omega_{LL1} + \omega_{HH1}\omega_{LL2} - \omega_{HL1}\omega_{LL2} \quad (60)$$

A.4 The Utilities and Cutoffs for the One Neighborhood Case

Consider first the lifetime utilities that someone born to a low educated might earn. Since remaining low skilled implies that the time cost function is not invoked, the lifetime utility of someone born in state L who remains low skilled is independent of her/his ability level. Let $LU(\omega, H_{t-1}, H_t)$ be the lifetime utility of a person with an ability level ω who is born into the state H_{t-1} and chooses the state H_t. Then the lifetime utility of someone born into L that stays in L is

$$LU(\omega, L, L) = W_L^{\gamma} + \beta \tilde{V}_L. \tag{61}$$

In contrast, someone born into the low state with an ability level ω who becomes high skilled obtains

$$LU(\omega, L, H) = \left(W_H - \frac{W_H}{\bar{H}}(1-\omega) \right)^{\gamma} + \beta \tilde{V}_H \tag{62}$$

In contrast to the Three Path Case, since houses are all of the same type, there is no wealth effect. The lack of a wealth effect explains why someone born into the high state that does not become high educated will obtain the same the lifetime utility as someone born into the low state. This equality can be gleaned from comparing equation (61) and the lifetime utility of someone born into the high state who remains low skilled which is

$$LU(\omega, H, L) = W_L^{\gamma} + \beta \tilde{V}_L. \tag{63}$$

Since someone born to high skilled parent has a lower time cost of obtaining skill than someone born to a low skilled parent, the lifetime utility of someone born into H who becomes high skilled,

$$LU(\omega, H, H) = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}}(1-\omega) \right)^{\gamma} + \beta \tilde{V}_H, \tag{64}$$

is higher than someone becomes high skilled but is born to a low skilled parent, for the same ability level.

The cutoff ability level for the low state is the ability level that equates the lifetime utilities of the low and high skilled paths for someone born into the low state. Hence, ω_L solves

$$W_L^{\gamma} + \beta \tilde{V}_L = \left(W_H - \frac{W_H}{\bar{H}}(1-\omega_L) \right)^{\gamma} + \beta \tilde{V}_H \tag{65}$$

Similarly, the cutoff for the high state equates the lifetime utilities for the two paths for someone with that ability level who is born into the high state. Hence,

$$W_L^\gamma + \beta \tilde{V}_L = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}}(1-\omega_H) \right)^\gamma + \beta \tilde{V}_H \tag{66}$$

However, the fact that the lifetime utilities of people who remain low skilled is the same over the two states implies that

$$\left(W_H - \frac{W_H}{\bar{H}}(1-\omega_L) \right)^\gamma + \beta \tilde{V}_H = \left(W_H - \frac{W_H}{\tilde{H}\bar{H}}(1-\omega_H) \right)^\gamma + \beta \tilde{V}_H \tag{67}$$

Performing a few lines of algebra on this equation yields that

$$\omega_H = 1 + \tilde{H}\omega_L - \tilde{H}. \tag{68}$$

A.5 The Population Fractions for the One Neighborhood Case

Solving the two equations and two unknowns given in equations (14) and (15) yields

$$\pi_H = \frac{1-\omega_L}{1-\omega_L+\omega_H} \tag{69}$$

and

$$\pi_L = \frac{\omega_H}{1-\omega_L+\omega_H} \tag{70}$$

A.6 The Expected Value Functions

This section gives the system of equations that results from carrying the integration required in equations (18)-(20). From equation (18) one obtains

$$\begin{aligned} \tilde{V}_{LL}(1-\omega_{LL1}\beta) = \omega_{LL1} W_L^\gamma + Z_{LL}^{HL}(\omega_{LL2}) - Z_{LL}^{HL}(\omega_{LL1}) \\ + Z_{LL}^{HH}(\omega_1) - Z_{LL}^{HL}(\omega_{LL2}) + \beta(\omega_{LL2}-\omega_{LL1})\tilde{V}_{HL} + \beta(1-\omega_{LL2})\tilde{V}_{HH} \end{aligned} \tag{71}$$

where

$$Z_{LL}^{HL}(\omega) = \frac{\bar{H}_L}{W_H(\gamma+1)} \left(W_H - \frac{W_H}{\bar{H}_L}(1-\omega) \right)^{\gamma+1} \tag{72}$$

and

$$Z_{LL}^{HH}(\omega) = \frac{\bar{H}_L}{W_H(\gamma+1)} \left(W_H - \frac{W_H}{\bar{H}_L}(1-\omega) + P_D(1-D_H) \right)^{\gamma+1} D_H^{1-\gamma}. \tag{73}$$

From equation (19) one obtains

$$\tilde{V}_{HL} = \omega_{HL1} W_L^{\gamma} + \omega_{HL1}\beta\tilde{V}_{LL} + Z_{HL}^{HL}(\omega_{HL2}) - Z_{HL}^{HL}(\omega_{HL1}) + Z_{HL}^{HH}(\omega_1) - Z_{HL}^{HL}(\omega_{HL2}) + \beta(\omega_{HL2} - \omega_{HL1})\tilde{V}_{HL} + \beta(1-\omega_{HL2})\tilde{V}_{HH} \quad (74)$$

where

$$Z_{HL}^{HL}(\omega) = \frac{\tilde{H}\bar{H}_L}{W_H(\gamma+1)}\left(W_H - \frac{W_H}{\tilde{H}\bar{H}_L}(1-\omega)\right)^{\gamma+1} \quad (75)$$

and

$$Z_{HL}^{HH}(\omega) = \frac{\tilde{H}\bar{H}_L}{W_H(\gamma+1)}\left(W_H - \frac{W_H}{\tilde{H}\bar{H}_L}(1-\omega) + P_D(1-D_H)\right)^{\gamma+1} D_H^{1-\gamma}. \quad (76)$$

From equation (20) one obtains

$$\tilde{V}_{HH} = \omega_{HH1}\left(W_L + P_D(D_H - 1)\right)^{\gamma} + \omega_{HH1}\beta\tilde{V}_{LL} + Z_{HH}^{HL}(\omega_{HH2}) - Z_{HH}^{HL}(\omega_{HH1}) + Z_{HH}^{HH}(\omega_1) - Z_{HH}^{HL}(\omega_{HH2}) + \beta(\omega_{HH2} - \omega_{HH1})\tilde{V}_{HL} + \beta(1-\omega_{HH2})\tilde{V}_{HH} \quad (77)$$

where

$$Z_{HH}^{HL}(\omega) = \frac{\tilde{H}^2}{W_H(\gamma+1)}\left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega) + P_D(D_H-1)\right)^{\gamma+1} \quad (78)$$

and

$$Z_{HH}^{HH}(\omega) = \frac{\tilde{H}^2}{W_H(\gamma+1)}\left(W_H - \frac{W_H}{\tilde{H}^2}(1-\omega)\right)^{\gamma+1} D_H^{1-\gamma}. \quad (79)$$

A.7 Approximating the Unconditional Income Distribution

The approximation algorithm has two parts. First, set a number of base points for the approximation. All the approximations in this paper used 10,000 points. Then use the population fractions to partition this number of points into "subset numbers" that sum to the original base number. For a One Neighborhood Case simulation there will be two numbers, while for a Three Path Case there will be three numbers. For instance, in a Three Case State case simulation with 10,000 base points and population fractions of $\pi_{LL} = .5$, $\pi_{HL} = .2$ and $\pi_{HH} = .3$ the subset numbers would be 5,000, 2,000 and 3,000 respectively. These subset numbers are the number of grid points for the [0,1] interval for each initial state. Continuing the above example, one would break the [0,1] interval up into 5,000 points for the LL state. This process yields a grid of points that approximate the ω's in that state and the size of grid is correctly weighted to the whole population. Then, for each point in each grid use the cutoffs to determine the income that would accrue to person with an ω given by that point. The resulting set of 10,000 incomes is approximates the unconditional income distribution. Treating this approximation of the unconditional income distribution as an empirical distribution it is trivial to calculation the Lorenz Curve and the centiles used in the paper and in the derivation of the condition distributions.

A.8 Deriving the Conditional Cumulative Income Distribution Functions

A.8.1 The One State Case

A person's income depends on their random draw of ability and the terminal state of one's parent. The trick to deriving the conditional cumulative distribution function is to determine the parent's state from her/his income and then use this information to calculate the cumulative income distribution for someone born into that state. In the One Neighborhood case, it is straight forward to determine the state from the observed income. If the parent had a lifetime income of $W_L + P_D$ then s/he finished in the low state with probability one. Any other income was generated by a high skilled parent.[10] Thus, there are two objects of interest, $Pr\left\{I_t|I_{t-1} = W_L + P_D\right\}$ and $Pr\left\{I_t|I_{t-1} \neq W_L + P_D\right\}$.

Consider first, $Pr\left\{I_t|I_{t-1} = W_L + P_D\right\}$. In this case, the parent was low skilled. Because the expected value of the high state is greater than the expected value of the low state, the cutoff person will earn less than the could be obtained from the low state.[11] Hence, there is a range of incomes accruing to people that choose to become high skilled that is actually lower than those that remain low skilled. Define ω_{LM} to be the ability level that equates the high and low skilled incomes of someone born into the low state, i.e. ω_{LM} solves

$$W_L + P_D = W_H - \frac{W_H}{\bar{H}}(1 - \omega_{LM}) + P_D. \tag{80}$$

A few lines of algebra imply that

$$\omega_{LM} = \bar{H}\left(\frac{W_L}{W_H} - 1\right) + 1. \tag{81}$$

[10] There is a small caveat that must be mentioned. There are two high skilled individuals that earn the income of $W_L + P_D$. However, the probability of two zero probability events is still zero. So the probability that someone with this income is low skilled is 1.

[11] To verify this fact, see the cutoff equation (65).

Note that ω_{LM} is strictly greater than ω_L.

Those born into the low state with an ability below ω_L will become low skilled an have a lifetime income of $W_L + P_D$. Individuals born into this state with an ability at least as large as ω_L will become high skilled and will have incomes of $W_H(1 - \frac{1}{\bar{H}}(1 - omega) + P_D$. Since the ω ability levels are distributed uniform over [0,1], the cumulative income distribution, condition on a parental income of $W_L + P_D$ is

$$Pr\left\{I_t \leq W_H(1 - \frac{1}{\bar{H}}(1-\omega) + P_D | I_{t-1} = W_L + P_D\right\} = \begin{cases} \omega - \omega_L & \text{if } \omega_L < \omega < \omega_{LM} \\ \omega_{LM} & \text{if } \omega = \omega_{LM} \\ \omega & \text{if } 1 \geq \omega > \omega_{LM} \end{cases} \tag{82}$$

The derivation of $Pr\{I_t | I_{t-1} \neq W_L + P_D\}$ proceeds in a similar fashion. The fundamental difference is that the parent was high educated. Define ω_{HM} to be the ability level that equates the lifetime incomes from the high and low states. Solving the equation implies that

$$\omega_{HM} = \tilde{H}\bar{H}\left(\frac{W_L}{W_H} - 1\right) + 1. \tag{83}$$

Those born to a high educated parent with an ability below ω_H will become low skilled an have a lifetime income of $W_L + P_D$. Individuals born into this state with an ability at least as large as ω_H will become high skilled and will have incomes of $W_H(1 - \frac{1}{\tilde{H}\bar{H}}(1 - \omega) + P_D$. Since the ω ability levels are distributed uniform over [0,1], the cumulative income distribution, conditional on the parental income not being $W_L + P_D$ is

$$Pr\left\{I_t \leq W_H(1 - \frac{1}{\tilde{H}\bar{H}}(1-\omega) + P_D | I_{t-1} \neq W_L + P_D\right\} = \begin{cases} \omega - \omega_H & \text{if } \omega_H < \omega < \omega_{HM} \\ \omega_{HM} & \text{if } \omega = \omega_{HM} \\ \omega & \text{if } 1 \geq \omega > \omega_{HM} \end{cases} . \tag{84}$$

Since the goal is compare the above distributions with those from the Three Path Case, some normalization is necessary. Convention dictates that $[0,1]$ should be chosen. It is useful to recall that the cutoff equations, above, imply that

$$W_H(1 - \frac{1}{\bar{H}}(1 - \omega_L) + P_D = W_H(1 - \frac{1}{\tilde{H}\bar{H}}(1 - \omega_H) + P_D. \tag{85}$$

Furthermore, they are both the minimum income. The largest income is $W_H + P_D$ obtained by individuals of unit ability born into either state. Carrying out the normalization yields the distributions

$$Pr\left\{I_{Nt} \leq \frac{\omega-\omega_L}{1-\omega_L} | I_{N(t-1)} = \frac{\bar{H}}{1-\omega_L}\left(\frac{W_L}{W_H}-1\right)+1\right\} =$$

$$s\begin{cases} \omega-\omega_L & \text{if } \omega_L < \omega < \omega_{LM} \\ \omega_{LM} & \text{if } \omega = \omega_{LM} \\ \omega & \text{if } 1 \geq \omega > \omega_{LM} \end{cases} \tag{86}$$

and

$$Pr\left\{I_{Nt} \leq \frac{\omega-\omega_H}{1-\omega_H} | I_{N(t-1)} \neq \frac{\bar{H}}{1-\omega_L}\left(\frac{W_L}{W_H}-1\right)+1\right\} =$$

$$\begin{cases} \omega-\omega_H & \text{if } \omega_H < \omega < \omega_{HM} \\ \omega_{HM} & \text{if } \omega = \omega_{HM} \\ \omega & \text{if } 1 \geq \omega > \omega_{HM} \end{cases} . \tag{87}$$

A.8.2 The Three Path Case

The derivation of the $Pr\left\{I_t \leq I \middle| I_{t-1}\right\}$ is potentially very complicated. The reason is that it is conceivable that there are regions in which the incomes from distinct states overlap implying that the parent's terminal state is not known with probability one. Fortunately, the nature of the problem rules this out. Let $I(\omega_S)$ be the lifetime income of person omega born into state S. Figure 15 gives a table of the lifetime incomes that accrue to each of the six cutoff individuals and the incomes of those that remain low skilled.

This table numerically verifies, for the parameter values discussed in the text, several important facts:

- $I(\omega_{LL1}) = I(\omega_{HL1})$
- $I(\omega_{LL1}) < W_L + P_D < I(\omega_{HH1}) < W_L + P_D D_H$
- $W_L + P_D D_H < I(\omega_{LL2}) = I(\omega_{HL2}) = I(\omega_{HH2})$.

Other parameterizations yielded the same qualitative results.

Hence, a parental income of $W_L + P_D$ or $W_L + P_D D_H$ indicates that the parent was low skilled with probability one. A parental income in $[I(\omega_{LL1}), I(\omega_{LL2}))$ that is not equal to either of the low skilled incomes indicates that the parent chose HL. Finally, an income at least as

Figure 15:

$W_L + P_D$	$W_L + P_D D_H$	$I(\omega_{HH1})$	$I(\omega_{HH2})$	$I(\omega_{LL1})$	$I(\omega_{LL2})$	$I(\omega_{HL1})$	$I(\omega_{HL2})$
1.8916	2.0764	2.0368	2.3763	1.8538	2.3763	1.8538	2.3763
1.8913	2.1500	2.1099	2.2894	1.8537	2.2894	1.8537	2.2894
1.8909	2.2234	2.1827	2.2567	1.8532	2.2567	1.8532	2.2567

large as $I(\omega_{LL2})$ indicates that the parent finished up in the HH state. These results imply that there are only three distinct functions that need to be derived.

Consider $Pr\left\{I_t \leq W_H(1 - \frac{1}{\bar{H}_L}(1-\omega)) + P_D \middle| I_{t-1} = W_L + P_D \text{ or } I_{t-1} = W_L + P_D D_H\right\}$. Either of these two parental incomes indicates that the parent was in state LL with probability one. Given this parental state, $I_t = W_L + P_D$ with probability ω_{LL1}. As in the One Neighborhood Case, because the expected value of the HL state is higher than that of the LL state, there are some individuals in the HL state that have a lower lifetime income than $W_L + P_D$. Let ω_{LLM} be the ability that equates the LL lifetime income with the HL lifetime income. Similar to the cases above, $\omega_{LLM} > \omega_{LL1}$. A few lines of algebra imply that

$$\omega_{LLM} = \bar{H}_L\left(\frac{W_L}{W_H} - 1\right) + 1. \tag{88}$$

Using the cutoffs then yields the desired condition cumulative income distribution function. To simplify notation, let $I_{t-1} = LL$ stand for $I_{t-1} = W_L + P_D$ or $I_{t-1} = W_L + P_D D_H$. Then the object of interest is

$$Pr\left\{I_t \leq W_H(1 - \frac{1}{\bar{H}_L}(1-\omega)) + P_D \middle| I_{t-1} = LL\right) = \begin{cases} \omega_{LL1} - \omega & \text{if } \omega_{LL1} \leq \omega < \omega_{LLM} \\ \omega_{LLM} & \text{if } \omega = \omega_{LLM} \\ \omega & \text{if } \omega_{LLM} < \omega \leq 1 \end{cases}. \tag{89}$$

Since a parental income in $[I(\omega_{LL1}), I(\omega_{LL2})) = I_{HL}$ that is not equal to either of the low skilled incomes indicates that the parent was in the HL, one needs to construct

$$Pr\left\{I_t \leq W_H(1 - \frac{1}{\tilde{H}\bar{H}_L}(1-\omega)) + P_D \middle| I_{t-1} \in I_{HL} \text{ and } I_{t-1} \neq I_{LL}\right\}.$$

Similar to the previous cases, ω_{HLM} is the ability level that equates the lifetime incomes from the LL and HL states. Some algebra implies that

$$\omega_{HLM} = \tilde{H}\bar{H}_L\left(\frac{W_L}{W_H} - 1\right) + 1. \tag{90}$$

To simplify the notation let Q_{HL} be the conditions that $I_{t-1} \in I_{HL}$ and $I_{t-1} \neq I_{LL}$. Using the cutoffs, one can quickly ascertain that

$$Pr\left\{I_t \leq W_H(1 - \frac{1}{\tilde{H}\bar{H}_L}(1-\omega)) + P_D \middle| Q_{HL}\right\} = \begin{cases} \omega - \omega_{HL1} & \text{if } \omega_{HL1} \leq \omega < \omega_{HLM} \\ \omega_{HLM} & \text{if } \omega = \omega_{HLM} \\ \omega & \text{if } \omega_{HLM} \leq \omega < 1 \end{cases}. \tag{91}$$

If the parent's income was greater than or equal to $I(\omega_{LL2})$ then s/he was in state HH. Let ω_{HHM} be the income that equates the incomes from the LL and HH state to a decent of an HH parent. Algebraically,

$$\omega_{HHM} = \tilde{H}^2\left(\frac{W_L}{W_H} - 1\right) + 1. \tag{92}$$

As in the previous cases, using the cutoffs yields that

$$\Pr\left\{I_t \leq W_H(1 - \frac{1}{\tilde{H}}(1-\omega)) \middle| I_{t-1} \geq I(\omega_{LL2})\right\} =$$
$$\begin{cases} \omega - \omega_{HH1} & \text{if } \omega_{HH1} \leq \omega < \omega_{HHM} \\ \omega_{HHM} & \text{if } \omega = \omega_{HHM} \\ \omega & \text{if } \omega_{HHM} \leq \omega < 1 \end{cases}. \tag{93}$$

So that they will be comparable with those from the One Neighborhood Case. all of these distributions are normalized to [0,1]. Since the smallest income in the Three Path Case is $W_H(1 - \frac{1}{\bar{H}_L}(1-\omega_{LL1}) + P_D$ and the largest is $W_H + P_D D_H$ the normalized distributions are

$$Pr\left\{I_{Nt} \leq \frac{\omega - \omega_{LL1}}{(1-\omega_{LL1}) + \frac{\bar{H}_L}{W_H}P_D(D_H - 1)} \middle| I_{t-1} = LL\right) = \begin{cases} \omega_{LL1} - \omega & \text{if } \omega_{LL1} \leq \omega < \omega_{LLM} \\ \omega_{LLM} & \text{if } \omega = \omega_{LLM} \\ \omega & \text{if } \omega_{LLM} < \omega \leq 1 \end{cases}. \tag{94}$$

$$Pr\left\{I_{Nt} \leq \frac{\omega - \omega_{HL1}}{(1-\omega_{HL1}) + \frac{\tilde{H}\bar{H}_L}{W_H}P_D(D_H - 1)} \middle| Q_{HL}\right\} = \begin{cases} \omega - \omega_{HL1} & \text{if } \omega_{HL1} \leq \omega < \omega_{HLM} \\ \omega_{HLM} & \text{if } \omega = \omega_{HLM} \\ \omega & \text{if } \omega_{HLM} \leq \omega < 1 \end{cases}. \tag{95}$$

and

$$\Pr\left\{ I_{Nt} \leq \frac{(1-\omega_{LL1}) + \frac{\bar{H}}{W_H} P_D(D_H - 1) - \frac{\bar{H}}{\tilde{H}^2}(1-\omega)}{(1-\omega_{LL1}) + \frac{\bar{H}}{W_H} P_D(D_H - 1)} \middle| I_{t-1} \geq I(\omega_{LL2}) \right\} = \begin{cases} \omega - \omega_{HH1} & \text{if } \omega_{HH1} \leq \omega < \omega_{HHM} \\ \omega_{HHM} & \text{if } \omega = \omega_{HHM} \\ \omega & \text{if } \omega_{HHM} \leq \omega < 1 \end{cases} . \quad (96)$$

A.9 The Cumulative Distribution Functions of Opportunity

There are two cumulative distribution functions to be derived, one for each case. The random variable of opportunity is denoted OP. Consider first the One Neighborhood Case. In this case there are only two "opportunities" with positive probability, $\bar{H}_L$ and $\bar{H}_L\tilde{H}$. The probability that someone in the population has opportunity $\bar{H}_L$ is π_L. The remained of the population has opportunity $\bar{H}_L\tilde{H}$. So the CDF is

$$F_1(OP) = \begin{cases} 0 & \text{if } OP < \bar{H} \\ \pi_l & \text{if } \bar{H} \leq OP < \bar{H}\tilde{H} \\ 1 & \text{if } OP \geq \bar{H}\tilde{H} \end{cases} . \quad (97)$$

In the three state OP can take on three values. The probability that an individual in the population has an opportunity of $OP = \bar{H}_L$ is π_{LL}. With probability π_{HL} $OP = \bar{H}_L\tilde{H}$. Finally, with probability π_{HH} an individual has an opportunity of $\tilde{H}^2$. The corresponding CDF is

$$F_2(OP) = \begin{cases} 0 & \text{if } OP < \bar{H}_L \\ \pi_{LL} & \text{if } \bar{H}_L \leq OP < \bar{H}_L\tilde{H} \\ \pi_{LL} + \pi_{HL} & \text{if } \bar{H}_L\tilde{H} \leq OP < \tilde{H}^2 \\ 1 & \text{if } OP \geq \tilde{H}^2 \end{cases} . \quad (98)$$

B Proofs

Claim 1 *For an initial state, the graph of the Lifetime utilities over ability is steeper for those who choose to become High Educated and purchase a house in the High neighborhood than the graph for those who become High Educated but purchase a house in the low neighborhood.*

Proof 1

The hypothesis is actually the claim that

$$\frac{\partial LU(\omega, H_{t-1}\bar{H}_{N_{t-1}}, HH)}{\partial \omega} > \frac{\partial LU(\omega, H_{t-1}\bar{H}_{N_{t-1}}, HL)}{\partial \omega}.$$

Using equations 5 and 6 in the text, one can compute the required derivatives and plug them into the desired relation to obtain,

$$\left(W_H - \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega) + P_D(D_{N_{t-1}} - D_H)\right)^{\gamma-1} D_H^{1-\gamma} \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}} > \left(W_H - \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega) + P_D(D_{N_{t-1}} - 1)\right)^{\gamma-1} \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}. \quad (99)$$

The above holds if

$$D_H^{\gamma} > \left(\frac{\left(W_H - \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega) + P_D(D_{N_{t-1}} - D_H)\right)}{\left(W_H - \frac{W_H}{H_{t-1}\bar{H}_{N_{t-1}}}(1-\omega) + P_D(D_{N_{t-1}} - 1)\right)}\right)^{1-\gamma} \quad (100)$$

Since $D_H > 1$ the above holds if the denominator is greater than the numerator. But a quick calculation shows that the numerator is in fact less than the denominator as long as $D_H > 1$ which is true by the parameterization of D_H. □

References

[1] Ronald Bénabou. Equity and efficiency in human capital investment: The local connection. *Review of Economic Studies*, 63:237–264, 1994.

[2] Julian R. Betts. Does school quality matter? evidence from the national longitudinal survey of youth. *Review of Economics and Statistics*, 77(2):231–250, May 1995.

[3] James Coleman. Equality of educational opportunity. *Government Printing Office*, 1966.

[4] Steven N. Durlauf. Neighborhood feedbacks, endogenous stratification, and income inequality. In William A. Barnett and Claude Hillinger, editors, *Dynamic disequilibrium modeling: Theory and applications:Proceedings ofthe Ninth International Symposium in Economic Theory and Econometrics*, pages 505–534. Cambridge University Press, Cambridge, 1996.

[5] Dennis Epple, Radu Filmon, and Thomas Romer. Housing voting, and moving: Equilibrium in a model of local public goods with multiple juisdictions. In *Research in Urban Economics*, volume 3. JAI Press, 1983.

[6] Raquel Fernandez and Richard Rogerson. Keeping people out: Income distribution, zoning, and the quality of public education. *International Economic Review*, 38(1):23–42, February 1997.

[7] Martha A. Gephart. Neighborhoods and communities as contexts for development. In Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber, editors, *Neighborhood Poverty*, volume 1. Russell Sage Foundation, New York, 1997.

[8] Eric A. Hanushek, Joh F. Kain, and Steven G. Rivkin. Teachers, schools, and academic achievement. *NBER Working Paper*, (6691), 1998.

[9] Christopher Jencks and Susan E. Mayer. The social consequences of growing up in a poor neighborhood. In Laurence E. Lynn Jr. and Michael Gl. H. McGeary, editors, *Inner-City Poverty in the United States*. National Academy Press, Washington, D.C., 1990.

[10] Peter Lambert. *The Distribution and Redistribution of Income*. Basil Blackwell, Cambridge, 1989.

[11] Alfred Marshall. *Principles of Economics*. Macmillan and Co, London, eighth edition, 1920.

[12] Alexandra Minicozzi. Nonparametric analysis of intergenerational income mobility: A tale of two families. *University of Texas at Austin, Department of Economics*, September 1998.

[13] James E. Rosenbaum. Changing the geography of opportunity by expanding residential choice: Lessons from the gautreaux program. *Housing Policy Debate*, 6:231–269, 1995.

[14] James E. Rosenbaum and Susan J. Popkin. Employment and earnings of low-income blacks who move to middle-class suburbs. In Christopher Jencks and Paul E. Peterson, editors, *The Urban Underclass*, pages 342–356. The Brookings Institution, Washington DC, 1991.

[15] Amartya Sen. *On Economic Inequality*. Clarendon Press, Oxford, 1997.

[16] Ronald Shoenberg. Constrained optimization. In *Constrained Optimization*. Apteck Systems Inc, Maple Valley WA, 1996.

[17] R. H. Tawney. *Equality.* Capricorn Books, New York, 1931.

[18] Robert Willis. Wage determinants: A survey and reinterpretation of human capital earnings functions. In A. Ashenfelter and R. Layard, editors, *Handbook of Labor Economics.* Elsevier Science Publishers, 1986.

[19] William J. Wilson. *The Truly Disadvantaged.* University of Chicago, 1987.

[20] William J. Wilson. *When Work Disappears.* Alfred A. Knopf, 1996.