The Effect of Underclass Social Isolation on Schooling Choice by Peter Streufert

Diamond Group

February 19, 2015

Contents

1 Introduction

As the author notes, this paper is a demonstrative response to The Truly Disadvanaged by Wilson (1987). Namely, Streufert attempts to model and demonstrate through a simulation the effect of poverty's concentration on the schooling choices of kids. More precisely, each student observes the employment status and education of the adults in her immediate environment and infers the value of schooling from the total set of observations. Consequently, students who are mired in neighborhoods with lower levels of education or income will fail to gather sufficient data about the relationship between higher education attainments and income. Consequently, she will underestimate the value of schooling and choose a level of lower than an individual who has been exposed to a representative sample of the relationship between high educational attainment and income.

1.1 The Benchmark Model

Consider a young person who has to make a decision on her level of schooling based on her inferred relationship between schooling and income. Let a role model be one observation of the connection between schooling and income. From the role models, the individual learns the relationship between a particular level of schooling s and its relation to income.

Mathematically, we denote the above as follows: Let $S = \{7, 8, \ldots, 19\}$[1] be the level of schooling an individual can choose. Let F_s be the cumulative distribution of income y at a particular choice of schooling s, such as the one depicted in **Figure 1.1**. Let the $\{F_s\}_S$ be the set of all such cumulative distributions. Finally, as a simplification of the model, through the observation of role models the individual learns perfectly the distribution for each $F_s(y)$, as opposed to needing to acquire representative large samples and the like. Consequently, when the young person learns the mean of F_s, $E[F_s(y)]$, it coincides with the population mean, μ.

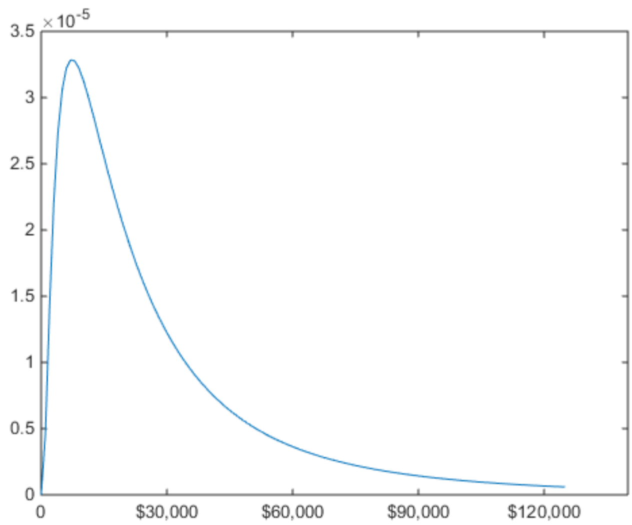


Figure 1.1: Example of $F_s(y)$

Given that the individual now knows $\{E[F_s(y)]\}$ for every element in set S, she maximizes utility, or in this case income, over s given a discount factor $\delta \in (0, 1)$ and annual disutility of

[1] Although it isn't legal to quit school at grade 7, the assumption is that the individual stops liking school in grade 7, fails the subsequent two grades and by age 16 drops out of school.

attending school $\theta \in \mathbb{R}$. Mathematically, she does

$$^{2}\max\{\arg_s \max\{\sum_{a=\sigma+1}^{59} \delta^{a-1} E[F_\sigma] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta | \sigma \in S\}\}$$

The $\sum_{a=\sigma+1}^{59} \delta^{a-1} E[F_\sigma]$ is the expected earnings over the individual's lifetime discounted to the present. The latter term $\sum_{a=1}^{\sigma} \delta^{a-1}\theta$ is sum of disutility of schooling over a person's potential further schooling years discounted to the present.

1.2 Incremental Characterization of schooling

This section augments the above model to reflect the fact that individuals update their preferences every year, as opposed to doing a single optimization problem in grade 7. This is done by taking the additional expected lifetime earnings that the person will get by investing in one more year of school and dividing that by the income that she would earn if she were to instead work during that year. Mathematically, the amount of expected income for an additional year of schooling is $[E[F_\sigma] - E[F_{\sigma-1}]]$ and the amount of income she can earn during that year is given by $E[F_{\sigma-1}]$. Now, to convert these terms into the present value, multiply them by $\frac{\delta(1-\delta^{59-\sigma})}{1-\delta}$. Thus, we get the expression:

$$\delta \frac{(1-\delta^{59-\sigma})}{1-\delta}(E[F_\sigma] - E[F_{\sigma-1}]) - E[F_{\sigma-1}]$$

Let B_σ be the aforementioned incremental benefit the person receives. According to Steufeurt's **Theorem 1**[3], as σ increases, the benefits will be weakly decreasing (ie: if $\sigma_{k+1} > \sigma_k \implies B_{\sigma_{k+1}} \leq B_{\sigma_k}$) if two conditions are met:

1. $\{E[F_s]\}$ is concave, such as in **Figure 1.2**.
2. $\{E[F_s]\}$ is increasing, such as in **Figure 1.2**.

If condition 1 fails B_σ can still be decreasing if foregone income outweighs future income –typically be true for low values of δ.

1.3 The Benchmark Simulation

The details of the simulation are uninformative as they pertain more to the manner in which it was programmed rather than any theoretical underpinning. However, there are a few noteworthy remarks. First, most simulated individuals in the dataset have *negative* disutility of going to school which means that they find school more enjoyable than work. The curves and distributions of income and disutility have been matched to fit observed data in the US. The figures from the simulation can be observed in **Figure 1.3**.

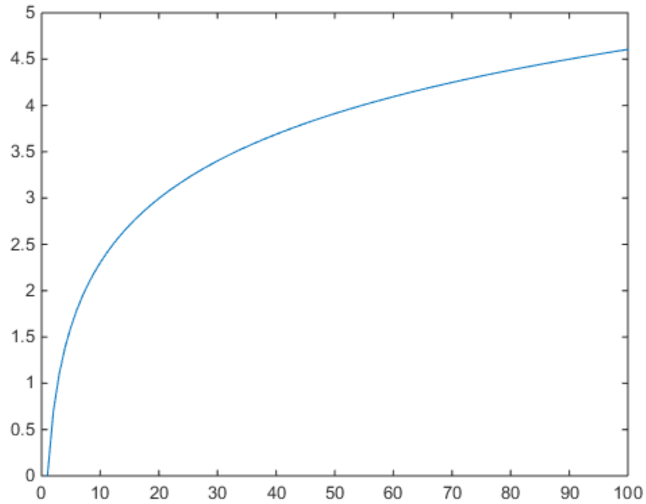
The takeaways from the model are as follows:

1. As the grounding for the simulation, Steufert checks the extreme parameters. The return to schooling for completing only grade 8 is in the interval [\$230, \$300]. In the simulation, two individuals were assigned discount coefficients δ of \$402 and \$301 respectively so they only complete grade 7 since both are greater than the return to completing grade 8. On the other hand, all individuals whose disutility of attending 8th grade is smaller than 230 go onto 8th grade. These benchmark results match the prediction of *Theorem 1.*

[2] Note, the difference between arg max and max. max picks the largest element from a set, potentially function output. So $\max\{3, 4, 84, -1, .3\} = 84$. arg max, on the other hand, maximizes a function output given an input variable. For example, for $\arg_x \max f(x) = -x^2 + 5$, $x = 0$

[3] proof shown in the Appendix

Figure 1.2: An increasing, concave curve



2. Once the individuals choose their schooling, the resulting incomes for that level of schooling are drawn randomly for the corresponding F_s distribution. For example, individuals with the two lowest disutility for school complete the most amount of schooling, 19 years. Correspondingly, their income will be drawn randomly from F_{19} which is a lognormal distribution of incomes for 19 years of schooling.

2 Social Isolation

2.1 Theory

An important assumption of this model is that, for each schooling s, the students observe a distribution of incomes that is bounded from above. This bound is meant to represent those that achieve a certain level of income, α and leave the underclass neighborhood. This means that the students only observe role models with income below level α.

$E[F_s^\alpha]$ is the expectation of distribution F_s^α, which is obtained by bounding the distribution F_s above at $\alpha \geq 0$. A student with disutility parameter θ observing the regression $\langle E[F_s^\alpha]\rangle_s$ chooses schooling $S^\theta(E[F_s^\alpha]\rangle_s)$. Streufert tests whether a decrease in α causes a decrease in $S^\theta(E[F_s^\alpha]\rangle_s)$. He finds that it does, as further shown in the Isolation Simulation in his *Section 3.2*. However, he adds that a bound from above does not always decrease schooling and provides a theoretical counterexample in his Section 3.3.

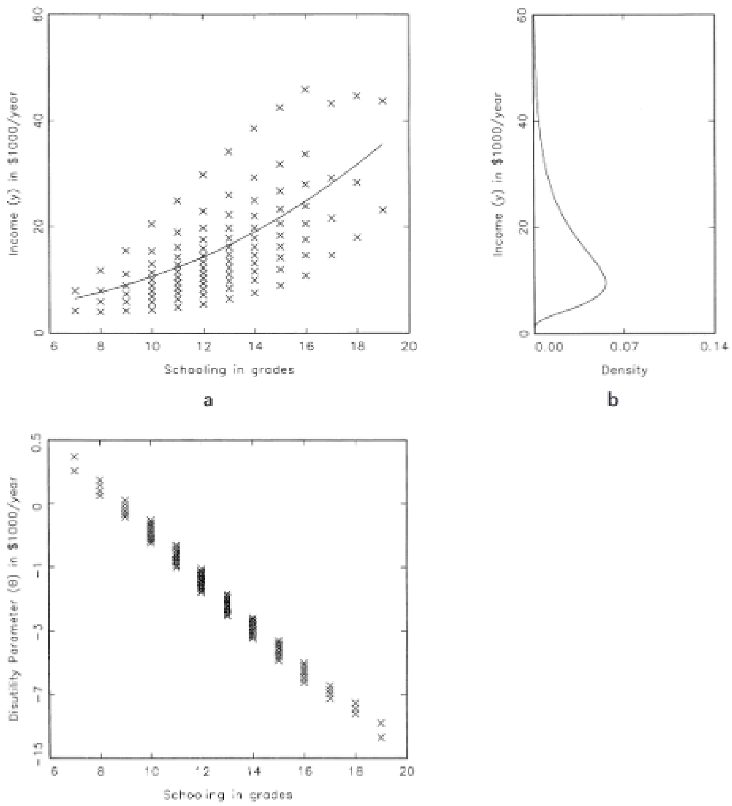
2.2 Isolation Simulation

The Isolation Simulation illustrates the isolation effect by setting the upper bound at $\alpha = \$15,000$/year (i.e. young people have no role models earning more than α).

The results can be seen in **Figure 2.1**:

- **Figure 2.1a** shows the \$15,000 line, shifting the regression from $\langle E[F_s]\rangle_s$ (the dotted curve) to $\langle E[F_s^{15000}]\rangle_s$ (the solid curve). The stars above the bound ("truncation line" in Streufert's terms) represent the 12 people that leave the underclass area, and the other

Figure 1.3: Benchmark Simulation Results



88 stars below the bound represent those that will remain as role models for the next generation.

- **Figure 2.1c** shows the inward shift of schooling (from the small stars to the large stars) because the perceived incremental benefit of each grade falls and it affects everyone's schooling choice.
- **Figure 2.1d** shows the lifetime utility loss that each person suffers from this decline in schooling.
- **Figure 2.1b** shows a decline in the distribution of incomes.

2.3 Theoretical Counterexample

It was proven in Theorem 1 and demonstrated in the Isolation Simulation that truncation causes a decline in schooling when it decreases the perceived incremental benefit.

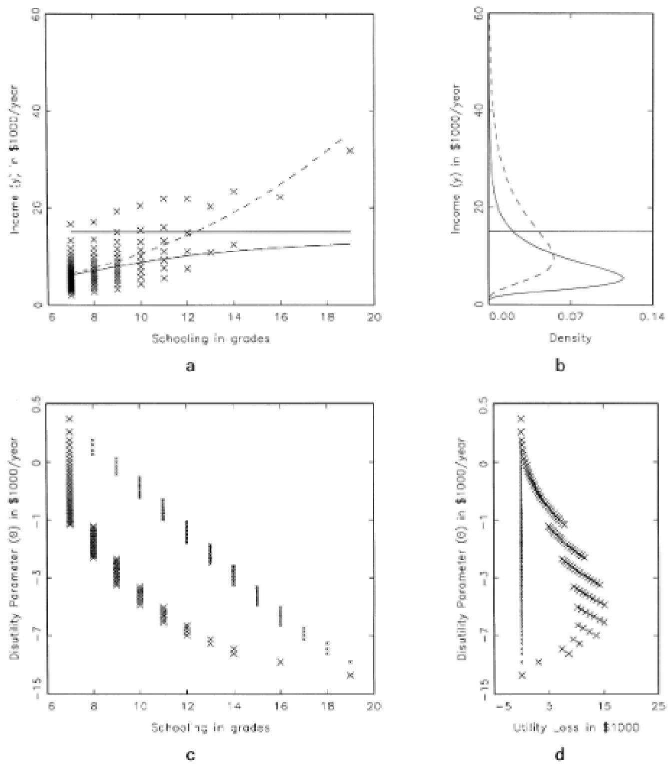
$$B_\sigma(\langle E[F_s^\alpha]\rangle_s) = \delta(E[F_\sigma^\alpha] - E[F_{\sigma-1}^\alpha])(1 - \sigma^{59-\sigma})/(1 - \sigma) - E[F_{\sigma-1}^\alpha]$$

The incremental benefit, shown above, will decrease through truncation because its effect on the second term is always negative and ambiguous on the first term is ambiguous.

The second term, $E[F_{\sigma-1}^\alpha]$ is the income a student gives up while attending grade σ instead of beginning work after grade $\sigma - 1$. Truncation decreases the perceived level of forgone income, thereby decreasing the perceived incremental benefit of grade σ.

In the first term, the focus is on $E[F_\sigma^\alpha] - E[F_{\sigma-1}^\alpha]$, which is the slope of the regression line between $s = \sigma - 1$ and $s = \sigma$. This is the additional income an individual receives in each

Figure 2.1: Isolation Simulation Results



working year from attending grade σ rather than grade $\sigma - 1$. Though it did in the Isolation Simulation, truncation doesn't necessarily

The theoretical counterexample is as follows, and refers to the simulation shown in **Figure 2.2**. We start with **Figure 2.2a**. Each F_s is defined to place 1/3 probability of each of the three triangles (which correspond to incomes of \$50, 000/year, \$30, 000/year, and $\$20,000/year$) shown at each level of schooling s. Setting the bound at \$20, 000/year, represented by the flat solid line, the top two triangles from every F_s are cut off. Therefore, the truncated regression is steeper than the true regression (the dotted upward-sloping line). In English, this means that because of the truncation, students underestimate the income forgone by another year of schooling and overestimate the resulting incremental income. This is an overestimation of the benefit of additional schooling.

In **Figure 2.2b**, we the effects of this overestimation on choices made. When perceiving the true regression, everyone chooses 7 grades of schooling (i.e. the benefit of grade 8 is less than the lowest disutility parameter of $\theta = -\$11,801$/year). Under the truncated regression, everyone with $\theta < -\$1,300$/year chooses more schooling if they would if they perceived the true benefits.

3 Polarization

3.1 Theory

This section extends the Truncation Model by letting the truncation cutoff vary over time. The student's decision is now a sequential problem in which she continues her schooling if and only if her observed role models show benefits of additional schooling large enough to continue school.

It is assumed that α increases over time. The intuition behind that is that as the underclass youth advances to higher levels of schooling, she will encounter stronger role models and the effects initial social isolation will be mitigated. However, this does not preclude a social isolation so powerful that she drops out of school before observing higher income role models.

Let $\{\alpha_t\}_{t\in\{7,8...,19\}}$ be a sequence of cutoffs. Streufert assumes that a person with disutility parameter θ and truncation cutoffs $\langle\alpha_t\rangle_{t\in\{7,8...,19\}}$ chooses schooling:

$$min\{t \geq 7 | S_t^\theta(\langle E[F_s^{\alpha_t}]\rangle_s) = t\},$$

where

$$S_t^\theta(\langle E[F_s^{\alpha_t}]\rangle_s = max\{argmax\{\sum_{a=\alpha+1}^{59} \delta^{a-1} E[F_\sigma^{\alpha_t}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta | \sigma \in \{t, t+1, ..., 19\}\}\}.$$

The first expression says that an individual will stop attending school in the first grade t such that the role models she is exposed to are under α_t, which leads her to choose t grades of schooling. The second expression defines S_t such that it is identical to S except that an individual can't choose a level of schooling lower than t.

An important point is that the student does not anticipate changes in her role models, and therefore won't continue school in order to gather information.

3.2 Polarization Simulation

An important feature of this simulation is that $\alpha_t = \$15,000$/year for $t \leq 12$, and $\alpha_t = +\$\infty$/year for $t > 13$.

The results can be seen in **Figure 3.1**. Isolation influences a student to choose a less than optimal level of schooling (like in the Isolation simulation) unless she makes it to the university level, which then exposes her to role models that cause her to choose the optimal level of schooling (like in the Benchmark simulation).

The critical point of entering university partitions the underclass youth into the school-loving youth ($\theta < -\$7,000$/year and postgraduate training) and the school-averse youth ($\theta > -\$7,000$/year and no more than a high-school education). The former overcome their initial isolation, while the latter don't.

4 Income Support

4.1 Theory

In this section, Streufert modifies the benchmark model by incorporating an income support program that guarantees everyone an income of at least β. This creates a scenario where the underclass youth observe no role models with an income below β and a significant amount with an income of exactly β.

$E[F_{s,\beta}]$ is the mean distribution $F_{s,\beta}$ and is derived from F_s, by defining $F_{s,\beta}(y) = 0$ if $y < \beta$ and $F_{s,\beta}(y) = F_s(y)$ if $y \geq \beta$ (Streufert calls this "censoring below at $\beta \geq 0$). Then let a person with disutility parameter θ use the regression $\langle E[F_{s,\beta}]\rangle_s$ to choose the schooling $S^\theta(\langle E[F_{s,\beta}]\rangle_s)$, where S is defined as it was in the benchmark model.

He then states a theorem that shows the income support program must decrease schooling. This effect happens through the incremental benefit of grade, σ:

$$B_(\sigma(\langle E[F_{s,\beta}]\rangle_s) = \delta(E[F_{\sigma,\beta}] - E[F_{\sigma-1,\beta}])(1-\delta^{59-\sigma})/(1-\delta) - E[F_{\sigma-1,\beta}].$$

In the first term, the slope of the regression, support diminishes additional future income. In the second term, the level of the regression, support increases forgone income. In both terms, the incremental benefit of grade σ is decreased, and therefore schooling is too.

Theorem 2: Suppose that $(\forall s \in \{8, 9, ..., 19\}) F_s$ first-order stochastically dominates F_{s-1}. Then $\beta' \geq \beta$ implies $S^\theta(\langle E[F_{s,\beta'}]\rangle_s) \leq S^\theta(\langle E[F_{s,\beta}]\rangle_s)$.

Streufert also shows that income support harms society as a whole in Theorem 3. The utility gained by the person from the income support β is on the left-hand side of the following inequality. The cost of the support, paid by the taxpayer, is on the right-hand side. Utility being measured in units income allows the two to be comparable.

Theorem 3: Let $\sigma* = S^\theta(\langle E[F_s]\rangle_s)$ be the schooling in the benchmark model and let $\sigma = S^\theta(\langle E[F_{s,\beta}]\rangle_s)$ be schooling given income support β. Then

$$(\sum_{a=\sigma+1}^{59} \delta^{a-1} E[F_{\sigma,\beta}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta) - (\sum_{a=\sigma*+1}^{59} \delta^{a-1} E[F_{\sigma*}] - \sum_{a=1}^{\sigma*} \delta^{a-1}\theta)$$

$$\leq \sum_{a=\sigma+1}^{59} \delta^{a-1}(E[F_{\sigma,\beta}] - E[F_\sigma]).$$

4.2 Support Simulation

This simulation alters the Benchmark Simulation by censoring all distributions F_s below $\beta = \$5,000$/year. This means that all role models make $\geq \$5,000$/year, including those that would have made less without the support.

For the results of the simulation, see **Figure 4.1**. **Figure 4.1c** shows that 34 out of the 100 people complete only 7 grades (the large stars), whereas only 2 out of the 100 people do this in the benchmark (the small stars).

Unlike the Isolation Simulation, which decreases nearly everyone's chosen level of schooling, the effects in this simulation are mostly concentrated in the lower portion of the population. The most severely harmed person, with $\theta = -\$996$/year, completes only 7 grades, while she would have completed 11 without the income support.

5 Combination Simulation

This simulation is a combination of the Polarization and the Support Simulations. Social isolation is modeled with $\alpha_t = \$15,000$/year for $t \leq 12$ and $\alpha_t = +\$\infty$/year for $t \geq 13$, and income support is modeled with $\beta = \$5,000$/year.

The results can be seen in **Figure 5.1**. Social isolation and income support "squeeze" the regression line from both sides, so they decrease schooling through distinct and self-reinforcing ways. No one chooses more schooling than they do in either simulation that these two effects are tested separately, and 43 people choose less. For example, a person with $\theta = -\$2,752$/year chooses 14 grades in the Benchmark simulation, 9 in the Isolation and Polarization Simulations, 14 in the Support Simulation, and 7 in the Combination Simulation.

6 Streufert's Conclusions

Though there exists the theoretical possible that it doesn't, the simulations suggest that social isolation can have severe negative effects, including dividing the underclass youth into two classes. Streufert argues that making neighborhoods more representative of society as a whole can help, but that it may be more effective to make the focus even more local. The examples he gives are families, friends, and cultural and recreational groups.

Income support depresses schooling, and the simulations suggest that it may disproportionately affect the lower end of the income distribution. Streufert asserts that this is not a problem of information but an unintended consequence

He also hopes that this paper can help inform empirical work on social isolation and schooling choice. A key take-away from this is that role models in their social networks, not merely mean neighborhood income, may play a key role in the development of underclass youth.

7 Critique

1. In these models, we assume that the children choose their schooling. This may be true in select cases but seems unreasonable outside of particular instances. For example, suppose an offspring stems from a family with two parents both of whom possess PhDs, reside near Stanford, and have her attend one of the US's best high schools. Under such circumstances, it seems unreasonable to assume that it ever enters the mind of the student to not finish high school because attaining a Bachelor's, or even a PhD, is likely to be an internalized expectation. In fact, thinking of not finishing may never enter her mind in earnest. On the other hand, for a student stemming from a less fortunate background, alternative paths are observed and thus conscious choice of the length of schooling may kick in. Consequently, the model may be appropriate for modeling or exhibiting behaviour in underserved communities but the results and comparisons may lose validity in juxtaposition to more privileged circumstances.

2. The model may be antiquated, as it was disseminated in the 1989. For one, we consider the proliferation of the Internet over the past 2 decades as a vehicle to information, including the relationship between schooling and income. By facilitating access to articles, news, studies, pamphlets, and other sources pertaining to the above relationship, the effect of the role model in the agent's schooling choice may be diminished, instead being replaced by the aforementioned sources. Consequently, with uniform access to Internet, one would assume that the effect of one's immediate surroundings on schooling attainment is eliminated or diminished. On the other hand, Internet penetration is not uniform across socio-economic strata and it is not inconceivable that Internet access is more limited in poor neighborhoods. Consequently, if Internet penetration is high and uniform above a particular demographic but sparse below that threshold then we would see a greater division in information access and therefore, more stratification in educational attainment. There has been research that supports this information disparity. For instance, Hoxby and Avery (2013) find that the vast majority high-achieving low-income students don't apply to selective colleges despite the fact that, due to financial aid, they would cost them less than the schools they do apply to. This also makes the Polarization Simulation assumption that social isolation ends once the student gets into college less plausible, since these students are not being surrounded by the higher-quality peers they would have at more selective schools.

   Secondly, the past 2 decades have seen increasing de facto segregation in communities which essentially translates into the minority poor being more isolated than in 1980s. Consequently, ceteris paribus, role models may have a greater influence on the educational attainment of kids stemming from these isolated neighborhoods than the relationship depicted in the simulations.

3. Finally, the model omits numerous essential factors in schooling, such as peer effects, thereby potentially overstating the effect of role models on student school achievement.For example, consider a plausible scenario where a student does observe high-earning role models but is surrounded by peers who don't. Consequently, through peer pressure or expectations, she does not go on to achieving the level of schooling otherwise predicted by her access to high earning role models. Of course, the opposite case can be made in which the peers counteract the effect of absent high earning role models. However, such a setup would substantially complicate this simple model and so the omissions make modeling sense.

8 Extensions

8.1 Our Extension Ideas

It would be interesting to see the effects of not observing lower levels of education on educational attainment, in other words the antithesis of social isolation in underserved neighborhoods. Such an illustration may be used in depicting the perception of the poor from the perspective of those who have never access to information on the poor. Furthermore, by combining the Polarization Model with our censorship of low income for the rich, it seems possible to model the ongoing income stratification.

First, lets model the truncation from below, as opposed to the censorship done in Streufert's model. Let "top 50%" denote the half of the population with the more negative schooling disutility. Let the "bottom" be the converse. Consider a student in the top 50% percentile of income distribution, which 1989 was above$41.5k[4]. Given a 1990 figure for a rough estimate, this was about the median level of income for someone with a Backelor's Degree[5]which translates to 16 years of education.

Consequently, for the simulation we censor every F_s distribution from below at $\beta = 41,000/year$ for 50% of the population with the most negative θ or schooling disutility. This will suggest that a person with respective θ never observes incomes below $41,000$. We also keep to the perfect regression assumption under which the student is not fooled into an ability to earn $50k on a high school degree through a lack of information (an assumption that does not truly pertain to a truncation from above). For our first case, the other half of θ parameters do not truncate their income distribution.

For our latter simulation, let the above setup remain for the top 50% of θ parameters. Let those θs that fall into the 25%-50% quartiles have their $\{F_\sigma\}$ unchanged. Now, for the bottom 25%, bound the $\{F_\sigma\}_S$ with α=$15,000, as done by Streufert.

Unfortunately, due to our inability to recreate Streufert's simulation code written in 1989 we are unable to compare the results of this setup to those of Streufert's paper.

8.2 Extensions in Literature

1. Accordingly to Google Scholar, the Streufert paper has 85 citations, most of which seem to revolve around urban economics. More specifically, community structure, location, and

[4]Page 33 in http://www.federalreserve.gov/pubs/feds/2009/200913/200913pap.pdf
[5]http://www.infoplease.com/ipa/A0883617.html

education seem to be the predominant themes that cite Streufert's paper suggesting that his approach or analysis did not venture far beyond the constraints of his subdiscipline. Nevertheless, it is often cited in papers attempting to build models that capture interaction between society and the individual. Specifically, there is a class of models that heavily implement statistical mechanics and distributions as a means of capturing individual and social behaviour concurrently [Brock 2001] Furthermore, through our survey of articles that cite Streufer's paper, the ostensible repeating theme is increasing stratification or polarization in income and geographical isolation between the rich and the poor.

2. Another paper by Steven N. Durlauf (1996) explores an extension of Streufert's paper, mentioned by us, which captures the reaction of the rich to social isolation. Namely, wealthy families have an incentive to create homogenous communities because, just as spillover community effects heavily influence the residents of poor neighborhoods, the opposite effects also influence the behaviour of the wealthy, thereby inducing them to homogenize or leave heterogenous neighborhoods. Although this paper eschews simulations in favour of direct probabilistic derivations, the approach seems to fit the one mentioned above where probability distributions are implemented to fuse aggregate and individual behaviour.

3. A paper by Melanie O'Gorman [O'Gorman 2010] actually applies a more elaborate simulation in vein of Streufert's simulations in order to model the persistence of the black-white wage gap. She then appends her analysis to show that conditional cash transfer programs and policy efforts do have a role in arresting the perpetuation of the income disparity but the effort has to be above a critical size in order to be effective.

9 Appendix and Proofs

9.1 Proof of Theorem 1:

Assume σ is weakly preferred to $\sigma - 1$

$$\sum_{a=\sigma+1}^{59} \delta^{a-1} E[F_\sigma] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta \geq \sum_{a=\sigma}^{59} \delta^{a-1} E[F_{\sigma-1}] - \sum_{a=1}^{\sigma-1} \delta^{a-1}\theta \tag{1}$$

$$\sum_{a=\sigma+1}^{59} \delta^{a-1} E[F_\sigma] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta \geq \sum_{a=\sigma}^{59} \delta^{a-1} E[F_{\sigma-1}] - \delta^{a-1} E[F_{\sigma-1}] \tag{2}$$

$$\sum_{a=\sigma+1}^{59} \sigma^(a-1)(E[F_\sigma] - E[F_{\sigma-1}]) - \delta^{\sigma-1} E[F_{\sigma-1}] \geq \delta^{a-1}\theta \tag{3}$$

Two side Dived by $\delta^{\sigma-1}$

$$\sum_{a=\sigma+1}^{59} \sigma^{(a-1)(\sigma-1)}(E[F_\sigma] - E[F_{\sigma-1}]) - E[F_{\sigma-1}] \geq \theta \tag{4}$$

$$\sum_{q=1}^{59-\sigma} \sigma^q (E[F_\sigma] - E[F_{\sigma-1}]) - E[F_{\sigma-1}] \geq \theta \tag{5}$$

$$\frac{\delta(E[F_\sigma] - E[F_{\sigma-1}](1-\delta^{59-\sigma})}{(1-\delta) - E[F_{\sigma-1}]} \geq \theta \tag{6}$$

$$B_\sigma(< E[F_S] >_S) \geq \theta \tag{7}$$

Define $\sigma^* = S^\theta(< E[F_S] >_S)$ where:

$$\sigma^* \in \{7\} \cup \{\sigma \geq 8 | B_\sigma(< E[F_S] >_S) \geq \theta\} \tag{8}$$

$$(\forall \sigma > \sigma^*) B_\sigma(< E[F_S] >_S) < \theta \tag{9}$$

Equation (8) and (9) imply that $\sigma^* \in \{7\} \cup \{\sigma \geq 8 | B_\sigma(< E[F_S] >_S) \geq \theta\}$

9.2 Proof of Theorem2:

This paragraph shows that

$$(\forall \beta' \geq \beta)(\forall \sigma^+ > \sigma) \tag{10}$$

$$E[F_{\sigma^+,\beta'}] - E[F_{\sigma,\beta'}] \leq E[F_{\sigma^+,\beta}] - E[F_{\sigma,\beta}] \tag{11}$$

$$(\forall \beta' \leq \beta)(\forall \sigma^+ > \sigma) \tag{12}$$

$$E[F_{\sigma^+,\beta'}] - E[F_{\sigma,\beta'}] \tag{13}$$

$$= \{\beta' + \int_{-\beta'}^{\infty} (1 - F_{\sigma^+}(y))dy\} \tag{14}$$

$$= \beta' - \beta - \int_{\beta}^{\beta'} (1 - F_{\sigma^+})dy \tag{15}$$

$$= \int_{\beta}^{\beta'} F_{\sigma^+}(y)dy \leq \int_{\beta}^{\beta'} F_\sigma(y)dy \tag{16}$$

$$= \beta' - \beta - \int_{\beta}^{\beta'} (1 - F_\sigma(y))dy \tag{17}$$

$$= \{\beta' + \int_{\beta'}^{+\infty} (1 - F_\sigma(y))dy\} - \{\beta + \int_{\beta}^{+\infty} (1 - F_\sigma(y)dy)\} \tag{18}$$

$$= E[F_{\sigma,\beta^+}] - E[F_{\sigma,\beta}] \tag{19}$$

Now define $\sigma = S^\theta(< E[F_{s,\beta}] >_s)$ Then

$$(\forall \beta' \leq \beta)(\forall \sigma^+ > \sigma) \tag{20}$$

$$\{\sum_{a=\sigma^++1}^{59} \delta^{a-1} E[F_{\sigma^+,\beta}] - \sum_{a=1}^{\sigma^+} \delta^{a-1}\theta\} - \{\sum_{a=\sigma^++1}^{59} \delta^{a-1} E[F_{\sigma,\beta}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta\} \tag{21}$$

$$= \sum_{a=\sigma^++1}^{59} \delta^{a-1}\{E[F_{\sigma^+,\beta'}] - E[F_{\sigma,\beta'}]\} - \sum_{a=1}^{\sigma^+} \delta^{a-1}\theta - \sum_{a=\sigma+1}^{\sigma^+} \delta^{a-1} E[F_{\sigma,\beta'}] + \sum_{a=1}^{\sigma} \delta^{a-1}\theta \tag{22}$$

For β'

$$\leq \sum_{a=\sigma^++1}^{59} \delta^{a-1}\{E[F_{\sigma^+,\beta}] - E[F_{\sigma,\beta}]\} - \sum_{a=1}^{\sigma^+} \delta^{a-1}\theta - \sum_{a=\sigma+1}^{\sigma^+} \delta^{a-1} E[F_{\sigma,\beta'}] + \sum_{a=1}^{\sigma} \delta^{a-1}\theta \tag{23}$$

$$\leq \sum_{a=\sigma^+ +1}^{59} \delta^{a-1}\{E[F_{\sigma^+,\beta}] - E[F_{\sigma,\beta}]\} - \sum_{a=1}^{\sigma^+} \delta^{a-1}\theta - \sum_{a=\sigma+1}^{\sigma^+} \delta^{a-1}E[F_{\sigma,\beta}] + \sum_{a=1}^{\sigma} \delta^{a-1}\theta \tag{24}$$

$$= \{ \sum_{a=\sigma^+ +1}^{59} \delta^{a-1}\{E[F_{\sigma^+,\beta}] - E[F_{\sigma,\beta}]\} - \sum_{a=1}^{\sigma^+} \delta^{a-1}\theta\} - \{ \sum_{a=\sigma+1}^{\sigma^+} \delta^{a-1}E[F_{\sigma,\beta}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta\} < 0 \tag{25}$$

9.3 Proof of Theorem 3:

$$(\sum_{a=\sigma+1}^{59} \delta^{a-1}E[F_{\sigma,\beta}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta) \tag{26}$$

$$-(\sum_{a=\sigma^*+1}^{59} \delta^{a-1}E[F_{\sigma^*}] - \sum_{a=1}^{\sigma^*} \delta^{a-1}\theta) \tag{27}$$

$$- \sum_{a=\sigma+1}^{59} \delta^{a-1}(E[F_{\sigma,\beta}] - E[F_{\sigma}]) \tag{28}$$

$$= (\sum_{a=\sigma+1}^{59} \delta^{a-1}E[F_{\sigma,\beta}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta) \tag{29}$$

$$-(\sum_{a=\sigma^*+1}^{59} \delta^{a-1}E[F_{\sigma^*}] - \sum_{a=1}^{\sigma^*} \delta^{a-1}\theta) \tag{30}$$

$$-(\sum_{a=\sigma+1}^{59} \delta^{a-1}E[F_{\sigma,\beta}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta) \tag{31}$$

$$+(\sum_{a=\sigma+1}^{59} \delta^{a-1}E[F_{\sigma}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta) \tag{32}$$

$$= (\sum_{a=\sigma+1}^{59} \delta^{a-1}E[F_{\sigma}] - \sum_{a=1}^{\sigma} \delta^{a-1}\theta) - (\sum_{a=\sigma^*+1}^{59} \delta^{a-1}E[F_{\sigma^*}] - \sum_{a=1}^{\sigma^*} \delta^{a-1}\theta) \leq 0 \tag{33}$$

References

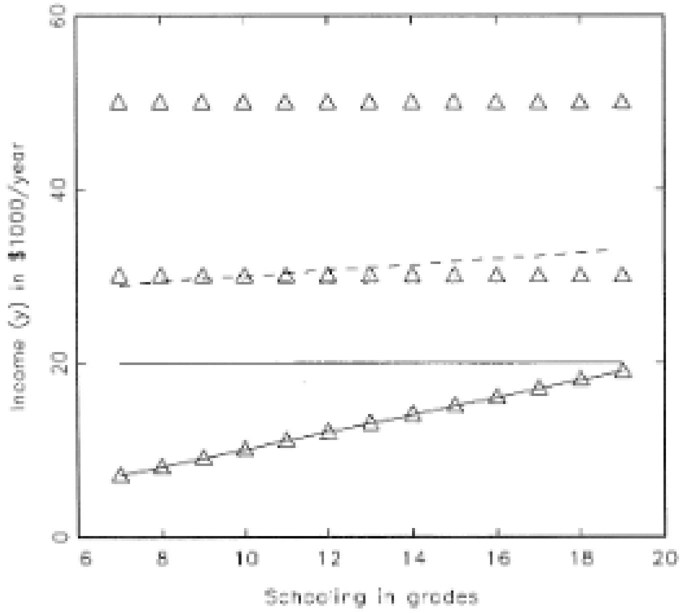
[Brock 2001] Brock, A. William; Durlauf, N. Steven. *Interaction-Based Models.* Chapter 54. Handbook of Economics, Volume 5, 2001, pg 3297-3380, online

[Durlauf 1996] Durlauf, N. Steven. *A Theory of Persistent Income Inequality.* Journal of Economic Growth. I. pg 75-93. March 1996. Online

[Hoxby 2013] Hoxby, Caroline; Avery, Christopher. *The Missing One-Offs: The Hidden Supply of High-Achieving, Low-Income Studnets.* Brookings Papers on Economic Activity. Spring 2013.

[O'Gorman 2010] O'Gorman, Melanie. *Educational disparity and the persistence of the black-white wage gap in the U.S..* Economics of Education Review. Volume 29, issue 4. pg 526-542. August 2010. Online

Figure 2.2: The Counter Example to Effect of Truncation



a

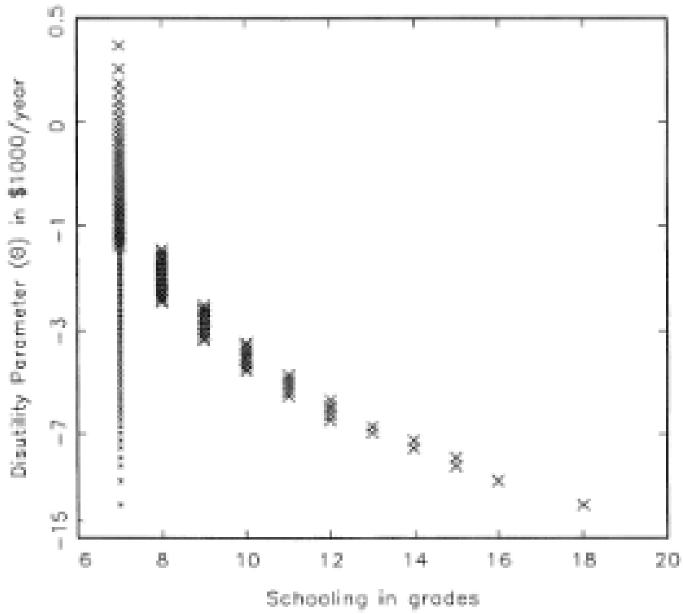


Figure 3.1: Results from Polarization Simulation

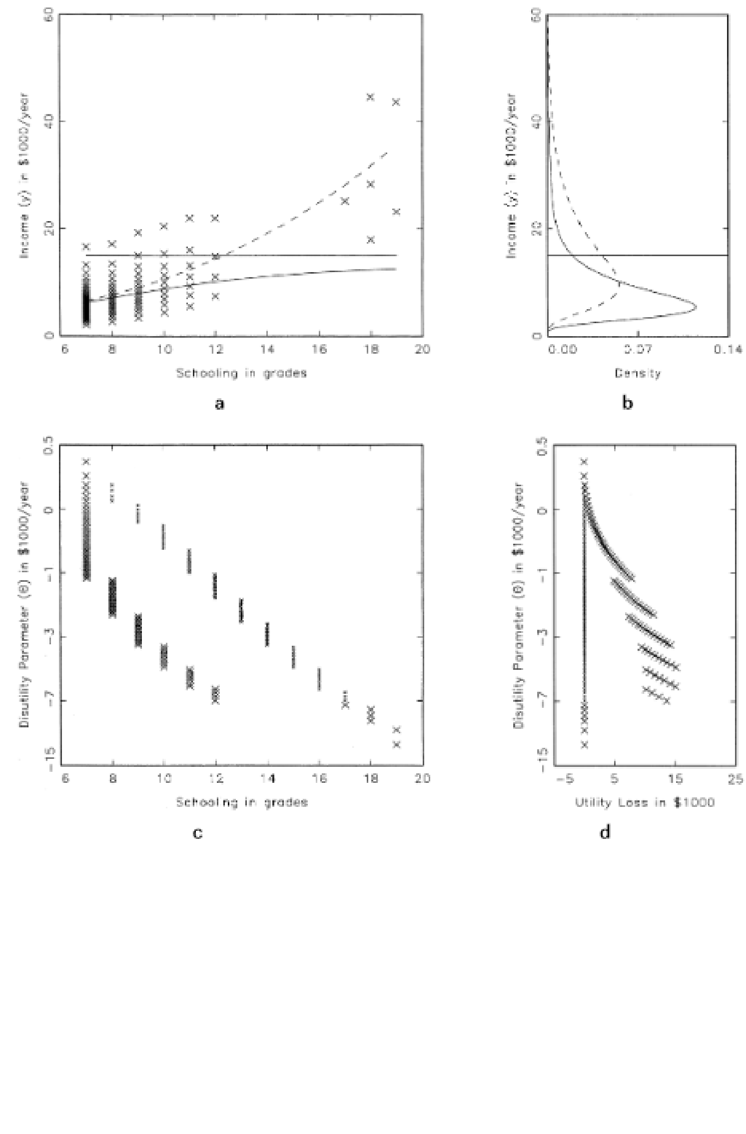


Figure 4.1: Income Support Simulation Results

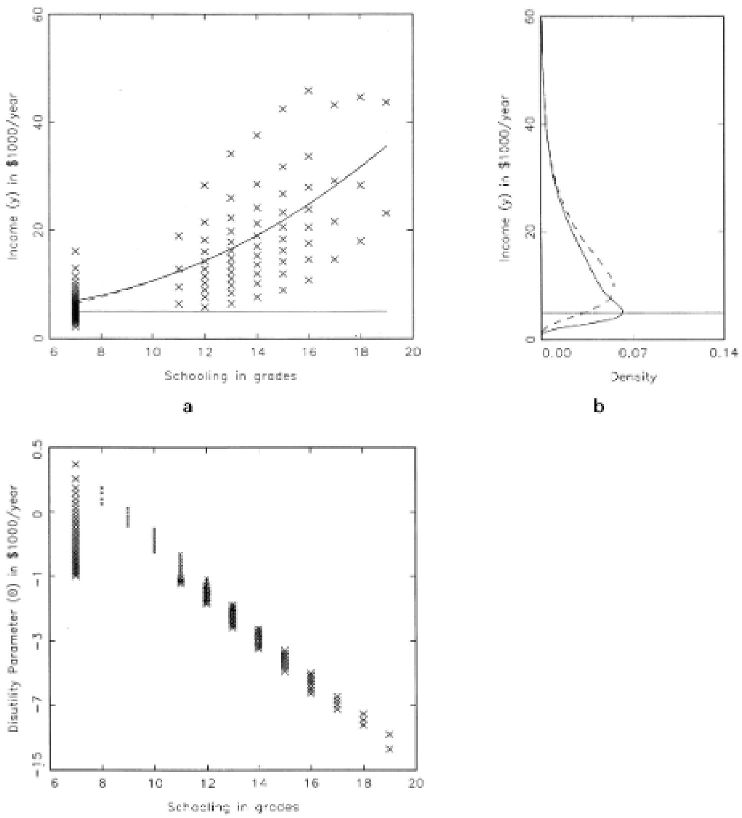


Figure 5.1: Combination Simulation Results