

2015

## Loss Aversion, Education, and Intergenerational Mobility

Liam C. Malloy

*University of Rhode Island, lmalloy@uri.edu*

Follow this and additional works at: [https://digitalcommons.uri.edu/ecn\\_facpubs](https://digitalcommons.uri.edu/ecn_facpubs)

The University of Rhode Island Faculty have made this article openly available.  
Please let us know how Open Access to this research benefits you.

### Terms of Use

This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our [Terms of Use](#).

---

### Citation/Publisher Attribution

Malloy, L. C. (2015). Loss aversion, education, and intergenerational mobility. *Education Economics*, 23(3), 318-337.

Available at: <https://doi.org/10.1080/09645292.2013.823909>

This Article is brought to you for free and open access by the Economics at DigitalCommons@URI. It has been accepted for inclusion in Economics Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons-group@uri.edu](mailto:digitalcommons-group@uri.edu).

---

## Loss Aversion, Education, and Intergenerational Mobility

The University of Rhode Island Faculty have made this article openly available.  
Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

### Terms of Use

This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our [Terms of Use](#).

# Loss Aversion, Education, and Intergenerational Mobility

Liam C. Malloy\*

July 3, 2013

Department of Economics

University of Rhode Island

Kingston, RI 02881, USA

This draft: May 2013, First draft: March 2009

Contact info:

Telephone: 401-874-4117

Fax: 401-874-2858

E-mail: lmalloy@mail.uri.edu

*JEL Classification:* E24, D03, I22

Date of Receipt:

Date of Acceptance of Revised Manuscript:

---

\*I would like to thank Allan Drazen, Pablo D'Erasmus, and especially John Shea for very helpful comments and suggestions. I am also grateful to seminar participants at the University of Maryland and the NBER Summer Institute on Income Distribution and Macroeconomics (especially Oded Galor) for thoughtful comments that led to improvements in the paper. I am also in debt to two anonymous referees who suggested the J-curve term and the semi-parametric test. Any errors or omissions are solely mine.

*Keywords:* Intergenerational Income Mobility, Loss aversion, Education

**Abstract**

Existing empirical work looking at the effects of parental income on IQ, schooling, wealth, race, and personality is only able to explain about half of the observed intergenerational income elasticity. This paper provides a possible behavioral explanation for this elasticity in which heterogeneous agents in sequential generations choose their education levels in the face of loss averse preferences and weak borrowing constraints. These borrowing-constrained agents make education investment choices in part to avoid consumption losses rather than to maximize lifetime resources. The model generates a positive intergenerational income elasticity even when there are functioning capital markets to finance education investments. I find empirical support for the J-shape education decision rule generated by the model and show that it is mostly successful in matching the asymmetric intergenerational transition rates between income quintiles of white families.

# 1 Introduction

A free market system can create a large amount of income inequality. The top 1% of earners in the United States earned 12.4% of the total wage income in the United States in 2007 (Piketty and Saez 2009). Acceptance of this system, especially in the United States, is generally predicated on the idea that each person has a good chance of success, no matter his starting point. However, recent research (e.g. Mazumder 2005, Mulligan 1997) shows that the elasticity of income between parents and children is as high as 0.6. As Mazumder (2005) notes, at this level it will take six generations for the descendants of someone at 25% or 200% of the mean to be within five percent of the average.<sup>1</sup> Furthermore, Hertz (2005) shows that there is a significant affluence net, in which children of parents with income in the top decile have a high chance of remaining in at least the top quintile, while those with parents in the bottom decile are very likely to be poor themselves, caught in the well-documented poverty trap. Both Hertz (2005) and Mazumder (2008) show that white and black families experience quite different, and asymmetric, transition rates which would not be implied by a linear model with an elasticity of 0.6 (or any other value).

In the empirical literature, there is some evidence (Blanden and Gregg, 2004, Corak et al 2004) that children from families with higher income do receive more and higher-quality education. Shea (2000), on the other hand, finds only a minimal relationship between family income and child education. In this paper, I present my own findings that there is a non-linear relationship between family income and education which may help explain these differences.

Theoretically, there are a number of reasons why parents' and children's income (and education) may be correlated across generations. Those parents with more education will

---

<sup>1</sup>Becker and Tomes (1986) review a number of studies that find an intergenerational elasticity of less than 0.2. At this level it would take only two generations for the descendants of someone at 200% of the mean to be within 5% of the average. Unfortunately, these studies suffered from a lack of data so that estimates of lifetime income were much noisier than in recent studies. Mazumder (2005) is able to use up to 16 years of father's income, averaging out the noise of transitory income, and finds an intergenerational elasticity of 0.6. Other studies (such as Mulligan 1997) use instrumental variable methods and get similar estimates while the Pew Economic Mobility study group settled on a number between 0.45-0.5 (Isaacs et al, 2008).

also tend to have more income, out of which they can pay for more (and higher quality) education for their children. Parents with fewer financial resources (who will also tend to have lower education) will be less able to finance the education of their children. The literature has explored the idea, beginning with Galor and Zeira (1993) and continued in a number of other papers such as Owen and Weil (1998), that capital markets for education are not perfect and that lower-income families may face more severe borrowing constraints than higher income families, getting caught in a poverty trap and unable to invest in the optimal amount of education.

The problem with this explanation is that it assumes imperfect, or non-existent, capital markets in financing education. Given the high correlation between education and income, markets should (and do) exist to finance education expenses despite the lack of collateral. According to data from the National Center for Education Statistics (NCES 2008), many (if not most) students are able to cover much of their higher education expenses at public universities through a combination of grants and loans, casting doubt on the idea that market failure in capital markets is to blame.

This paper proposes a novel transmission mechanism between parents and children that focuses on the differences in expectations, or reference consumption, among children from different income levels. If children form a reference level of consumption during childhood (based on family consumption) and have loss-averse utility functions, so that utility is concave in gains but convex in losses, then children from higher-income families will choose more education than an otherwise identical person from a poorer family<sup>2</sup>. Those at the very bottom of the income distribution will invest more in education than those immediately above them as they find it easier to replace reference consumption (and avoid losses), generating a J-shaped education investment decision and asymmetric intergenerational transition rates

---

<sup>2</sup>The model assumes that increased time spent investing in education reduces consumption in the first period. While there is very little data available to test this assumption, there is some qualitative evidence in the sociological literature (see, eg, Connor 2001 for the UK, Roderick et al 2008 for Chicago, Brown et al 2009 for Appalachia, Hahn and Price 2008 for a white vs. black comparison, and McGrath et al 2001 for Iowa farm families) that shows many who are eligible forgo higher education in order to work and increase consumption levels.

between income quintiles.

Previous research has assumed that reference consumption depends either on a person's level of consumption as a child (e.g. de la Croix, 2000) (supported by Waldkirch et al, 2004) or as coming from those around them or above them in the income distribution as in Genicot and Ray (2009)<sup>3</sup>. I develop a model with sequential generations in which heterogeneous agents are loss averse, inherit their reference level of consumption from their parents, and then must choose their education investment. The baseline model with loss aversion and weak borrowing constraints is able to generate an intergenerational elasticity of wages of 0.14 even when earning ability is completely independent between generations, suggesting that loss aversion is one possible explanation for the unexplained income elasticity between generations identified in Bowles et al (2005). The model generates asymmetric intergenerational transition rates between income quintiles and a J-shaped decision rule in education with family consumption as the explanatory variable. I test this implication using the NLSY97 data set and find some support for the J-shaped decision rule in education using both OLS and semi-parametric regressions.

The next section describes the model with particular attention to the assumption of loss-averse preferences while section three reviews relevant data moments from the intergenerational persistence literature. The numerical solution to the model and simulated results from the calibrated model are presented in section four. Section five presents the empirical results which support the J-shaped education decision rule implied by the theoretical model. Section six concludes.

---

<sup>3</sup>The key assumption that drives results different than those presented here appears to be that aspirations (against which utility is measured) are formed only by looking at those above you in the income distribution despite their use of an S-shaped utility function as in this paper.

## 2 Model

### 2.1 Basic Setup

The model presented here is a partial-equilibrium model with sequential generations. Agents live for three periods. The first period, labeled 0, can be thought of as childhood, in which agents do not have to make any decisions. Children learn their earning ability level,  $\psi$ , their reference consumption level,  $c_0$ , and the bequest,  $T$ , that they will receive at the beginning of period one. Both the reference level of consumption and the bequest come directly from their parents. Earning ability differs across agents and is normally distributed:

$$\begin{aligned}\psi_i &= \rho\psi_{i-1} + (1 - \rho)\bar{\psi} + \chi_i \\ \chi_i &\sim N(0, (1 - \rho^2)\sigma_\psi^2)\end{aligned}$$

I assume, in the baseline model, that earning ability is independent across generations. That is, a child's earning ability is completely independent of the earning ability of the parent so that  $\rho = 0$ . Thus each agent draws a random earning ability from a normal distribution with mean  $\bar{\psi}$  and standard deviation  $\sigma_\psi$ .

The second period of life, labeled 1, is the education period. Agents, knowing their earning ability, reference consumption, and bequest, invest in education to maximize lifetime utility. Because there is no uncertainty in this model, agents are also able to choose consumption in periods 1 and 2, saving (or borrowing) between periods, and the bequest level they will leave to the next generation. The value function is described by equation 1.



$$V(c_0, T, \psi) = \max_{e, c_1, c_2, T'} U(z_1) + \beta[U(z_2) + B(T')] \quad (1a)$$

*s.t.*

$$s \geq \Omega \quad (1b)$$

$$\varepsilon(e) \leq T + \Omega \quad (1c)$$

$$c_1 + s + \varepsilon(e) = (1 - e)\omega(e, \psi) + T \quad (1d)$$

$$c_2 + T' = \omega(e, \psi) + s(1 + r) \quad (1e)$$

$$c_1, c_2, T' \geq 0 \quad (1f)$$

Where  $z_t$  is the percent change from the reference level of consumption:  $z_t = \frac{c_t - c_0}{c_0}$ . Note that this gives a natural minimum to  $z_t$  of -1, which is important for loss aversion, as discussed below. The level of education is  $e$ , and the function  $B(\cdot)$  measures the warm glow utility derived from giving a bequest,  $T'$ , to the next generation. The wage is determined by  $\omega(\cdot, \cdot)$  which I assume to have the form:

$$\ln[\omega(e, \psi)] = \alpha_1 e + \alpha_2 e^2 + \zeta e \psi + \kappa \psi \quad (2)$$

Where I assume that  $\alpha_1, \alpha_2, \zeta, \kappa \geq 0$ . This is a fairly standard semi-log wage function that is quadratic in education. The major difference between equation (2) and a more standard wage function from the labor literature is that earning ability,  $\psi$ , is observable. The parameter  $\zeta$  represents the idea captured in quantile wage regressions (e.g. Lemieux 2006) that the return to education is higher for those with higher earning ability. The cost function for education is given by:

$$\varepsilon(e) = \phi_1 e + \phi_2 e^2 \quad (3)$$

With  $\phi_1, \phi_2 \geq 0$  so that education costs are at least weakly convex. The wage function

and the education function assure that there is some level of education,  $e^*$ , which may be at a corner, that uniquely maximizes lifetime resources for each level of earning ability,  $\psi$ .

The first two constraints in equation (1) reflect potential borrowing constraints, where borrowing is less than some level  $\Omega$  (discussed below) and where the agent may only use her bequest and her borrowing in order to finance the education expense,  $\varepsilon(e)$ . The two constraints (1d) and (1e) represent the per period budget constraints, where agents are not allowed to die in debt and the final constraint ensures that consumption and bequests are non-negative. Savings are assumed to grow at a risk-free rate  $r$ . Agents face both a direct cost of education,  $\varepsilon(e)$ , and an opportunity cost of education so that they work only  $(1 - e)$  in the first period. The opportunity cost is higher for those with higher earning ability as wages are increasing in  $\psi$ . The wage per unit of time worked is the same in both periods and depends on the chosen level of education,  $e$ , and innate earning ability,  $\psi$ . Because the wage is increasing in the chosen level of education, this opportunity cost is also convex. However, because loss-averse utility is not everywhere concave, the first order conditions do not provide a solution to the maximization problem as the second order test fails. Therefore I use numerical methods to solve the maximization problem as described in this section below.

At the beginning of period 2, agents reproduce so that the next generation's period 0 will match up with period 2 of the previous generation. The consumption level chosen by the parent generation in period 2 becomes the reference level of consumption for the child generation. Further, the bequest level chosen by the parent is given to the child at the beginning of period 1. The intergenerational persistence in the model comes from the fact that children inherit the reference consumption directly from their parents.

## 2.2 Loss Aversion

The utility function,  $U(z_t)$ , will be allowed to take two main forms. In the version of the model with reference consumption (but no loss aversion), the per period utility will be given by equation (5).

$$U(z_t) = \frac{1 - e^{-\mu z_t}}{\mu} \quad (5)$$

Equation (5) is a standard constant absolute risk aversion (CARA) utility function which is increasing and concave for all levels of  $z_t$ . In this case, agents will try to smooth consumption across the two periods as utility is concave everywhere so that there is no substantial difference between consuming just above the reference level or just below.

The more interesting case (at least in this model) is when the utility function represents loss averse preferences<sup>4</sup>. The key insight represented by loss aversion is that agents have concave utility when it comes to gains (as in traditional expected utility theory) but have convex utility when it comes to losses, even in riskless choice. In addition, the marginal utility of a loss is around twice as large as the marginal utility of a comparably sized gain. In this model there is no risk per se (as there is no uncertainty for the agent), so a loss averse utility function means that agents will avoid losses as much as possible. We can represent loss averse preferences by equation (6) as suggested by Kobberling and Wakker (2004).

$$U(z_t) = \begin{cases} \frac{1 - e^{-\mu z_t}}{\mu} & \text{for } z_t \geq 0 \\ \lambda \frac{e^{\nu z_t} - 1}{\nu} & \text{for } z_t < 0 \end{cases} \quad (6)$$

---

<sup>4</sup> While loss aversion was first introduced by Kahneman and Tversky (1979) as part of their prospect theory about making decisions under uncertainty, in this model there is no uncertainty and I use loss aversion as in Tversky and Kahneman (1991).

The loss aversion parameter  $\lambda$  was estimated by Kahneman and Tversky to be around 2.25 with other estimates usually between 1.5 and 3 (Camerer 2005). A value larger than one means that marginal utility is larger for losses than for comparably sized gains. Note also that utility is concave in gains but convex in losses. That is, marginal utility is decreasing away from zero in both directions, so that just as the marginal utility of a gain decreases as the gain increases, so too does the marginal disutility decrease as the loss increases.

Neilson (2002) has suggested a nomenclature for loss aversion which divides potential loss averse utility functions into those exhibiting weak and strong loss aversion. With strong loss aversion, marginal utility is everywhere higher for a loss than for a gain. That is, a loss averse function  $U(\cdot)$  is strongly loss averse if  $U(0) = 0$  and  $U'(y) \leq U'(z) \quad \forall z < 0 < y$ . In this case, since  $z_t$  is measured in percentage terms (and so has a minimum at -1) we can write  $U'(-1) \geq U'(0)^+$  where  $U'(0)^+$  is the marginal utility as  $z_t$  approaches 0 from the right. A strongly loss averse utility function is represented in Figure 1(a) and marginal utility in figure 1(b).

[Figure 1 about here.]

In order to get tractable results, I assume strong loss aversion as in Bowman et al (1999), requiring that  $\nu < \ln(\lambda)$  in equation (6). For a given investment in education, agents will either face a gain in both periods, a loss in both periods, or a loss in one period and a gain in the other period. Because utility is concave in gains, agents will smooth gains across the two periods (as much as possible, given the borrowing constraint). If an agent faces two losses, utility will be maximized by putting as much of the loss into one period as possible and minimizing the loss in the next period. The larger loss will, by necessity, be in period 1, the education period, due to the borrowing constraint. Finally, if the agent faces a loss in period 1 and a gain in period 2, she will choose to minimize the loss as much as possible by reducing the gain in period 2. Eliminating the loss in period 1 may not be possible if the agent faces a borrowing constraint. More details of the solution are given in section 2.5.

## 2.3 Borrowing Constraint

Borrowing constraints, represented by  $\Omega$  in equation (1), interact with loss aversion in important ways. Easier credit, or less severe borrowing constraints, will allow agents to avoid losses. Tighter borrowing constraints, on the other hand, make a loss in period one much more likely even if the agent will have a large gain in period two.

I explore three different degrees of borrowing constraint in this paper. First, there may be perfect capital markets. In this case there is no period one borrowing constraint but only a lifetime solvency constraint and we can write:

$$\Omega_{lt} = -\left[\frac{\omega(e, \psi)}{(1+r)}\right] \quad (\text{weakest, lifetime budget constraint only}) \quad (7a)$$

Second, I assume more realistically that agents can borrow up to their direct cost of education (assuming this is less than the lifetime borrowing constraint). Students can borrow the direct cost of education, but capital markets are imperfect as they cannot borrow to bring consumption from the second period forward to the first. In this case the borrowing constraint becomes:

$$\Omega_{ed} = -\varepsilon(e) \quad (\text{weak, education cost borrowing constraint}) \quad (7b)$$

Winter (2009) shows that as many as 18% of households may face borrowing constraints when financing education and some education loans require parents to co-sign the loan<sup>5</sup>. The final borrowing constraint I explore in this paper is a function of parental income:

$$\Omega_p = f[\omega(e_{-1}, \psi_{-1})] \quad (\text{strong, borrowing constraint based on parent income}) \quad (7c)$$

---

<sup>5</sup>According to the Department of Education (<http://federalstudentaid.ed.gov/federalaidfirst/>), while Stafford loans only require a student signature, so-called parent PLUS loans require a credit check and co-signer as do most private (i.e. not federally subsidized) loans.

Where  $\omega(e_{-1}, \psi_{-1})$  is the income of the previous generation. In this case, agents from poorer families will face a more severe borrowing constraint than agents from richer families. This could play a significant role in keeping those who are born poor from rising substantially in the income distribution.

## 2.4 Bequest Motive

The function  $B(T')$  in equation (1) gives the utility gained from leaving a bequest of size  $T'$ . This type of bequest, used in Galor and Zeria (1993), is known as a warm glow bequest because the person giving the bequest does not take the *utility* of the person receiving the bequest directly into consideration. This is required to avoid perverse incentives for parents in setting the level of a child's reference consumption. I use the same CARA function for  $B(\cdot)$  as is used for gains in utility:

$$B(T') = \frac{1 - e^{-\theta(T'/c_0)}}{\theta} \quad (8)$$

Because  $z_t$  is measured in percentage terms, it is useful to scale bequests by the level of reference consumption. If  $\theta = \mu$  then the agent sets  $\frac{T'}{c_0} = z_2$  when  $z_2 \geq 0$  or  $T' = c_2 - c_0$ . In many cases, however, agents will optimally have zero gain in the second period, in order to minimize losses in period one. An agent consuming at (or below) the reference level of consumption in period 2 will leave no bequest to the next generation.

## 2.5 Solution Methodology

Because the model makes the assumption of strong loss aversion, and given that agents choose consumption for only two periods, we can still solve for how they will divide a certain level of lifetime resources between periods (as in Bowman et al (1999)) despite the lack of concavity in the loss-averse utility function.

For a given level of education we can calculate both total lifetime resources and the

maximum consumption in the first period and minimum consumption in the second period, which will vary depending on the given borrowing constraint. For simplicity, I assume that  $\beta = 1 + r = 1$ . For any given level of education, I let  $R^{lt}(e, \psi)$  denote lifetime resources,  $R_{max}^1(e, \psi)$  denote maximum period one resources and  $R_{min}^2(e, \psi, T')$  denote minimum period two consumption when period one consumption is maximized.  $R_{max}^1(e, \psi)$  gives the maximum consumption when the borrowing constraint binds while  $R_{min}^2(e, \psi, T')$  gives the lowest level of period two consumption when all available resources are consumed in period 1. We can write:

$$R^{lt}(e, \psi) = (1 - e)\omega(e, \psi) + T + -\varepsilon(e) + \omega(e, \psi) - T' \quad (9a)$$

$$R_{max}^1(e, \psi) = (1 - e)\omega(e, \psi) + T - \varepsilon(e) - \Omega \quad (9b)$$

$$R_{min}^2(e, \psi, T') = \omega(e, \psi) + \Omega(1 + r) - T' \quad (9c)$$

The consumption and saving decisions for an agent for all possible levels of these three values are given in Table 1. If the agent can, he will spread gains across both periods. If he has a binding borrowing constraint, this will reduce period one consumption. However, if faced with a loss, he will maximize the loss in period 1 in order to minimize the loss in period 2. This is the result of strong loss aversion.

[Table 1 about here.]

We can write the indirect lifetime utility function associated with each level of education, based on the consumption and bequest pattern solved for above:

$$W(e, c_0, T, \psi) = \max_{c_1, c_2, T'} U(z_1) + \beta[U(z_2) + B(T')]$$

The optimal level of education,  $\hat{e}$ , maximizes this indirect utility function:

$$\hat{e} = \operatorname{argmax}_e W(e, c_0, T, \psi) \quad (10)$$

Subject to the constraints in equation (1).

In my numerical solution in section four, I assume the agent has a discrete number of education choices available, designed to mimic a high school dropout, a high school graduate, some college/Associate's degree, Bachelor's degree, Master's degree, and a PhD/professional degree. Given the discrete choice set it is easy to find the utility-maximizing education level by simply comparing total lifetime utility for each education level.

In addition, we can calculate  $e^*$ , the level of education that maximizes lifetime resources, in the absence of borrowing constraints:

$$e^* = \operatorname{argmax}_e (1 - e)\omega(e, \psi) - \varepsilon(e) + \omega(e, \psi) \quad (11)$$

The education level  $e^*$  equates the marginal cost (both direct and opportunity) of education with the marginal benefit in terms of wages. No agent will ever choose an education level higher than  $e^*$  as adding an extra unit of education above  $e^*$  will reduce lifetime resources and either increase a loss or reduce a gain in one or both periods, and therefore must reduce utility.

However, it is possible that the utility-maximizing level of education is less than the resource-maximizing level ( $\hat{e} \leq e^*$ ) because of the presence of borrowing constraints and loss aversion. Reducing the level of education from  $e^*$  will increase resources available in period one by reducing the direct and opportunity cost of education. These extra resources can then be used to reduce or eliminate a loss in the first period even if it means reducing a gain in the second period.

I use the consumption/savings levels shown in Table 1 to solve the model, making sure that it converges to a stable distribution<sup>6</sup>. I can then calibrate and simulate the model and compare the non-calibrated moments in the data to the simulation results. These simulation results are reported in section four, but first I review the data that the model is attempting

---

<sup>6</sup>For wages, for example, I add up the wages in each decile and then take the norm of the difference between generations. For convergence, I require this to be less than some number  $\epsilon$ .



to match.

### 3 Target Intergenerational Moments

Measuring the elasticity of income between generations is conceptually straightforward but empirically difficult. Most longitudinal data sets do not span a long enough time frame to include income from more than one generation. And those that do have a long enough time span, such as the Panel Study of Income Dynamics (PSID), usually don't have many years of income for both generations. Studies that used only a limited number of years for either or both generations are prone to measurement error because they are likely to be measuring transitory income rather than permanent income (see Bowles et al (2005) for a discussion of the literature).

Using Social Security Administration data in the U.S., matched with the 1984 Survey of Income and Program Participation, Mazumder (2005) finds that the estimated intergenerational income elasticity increases as the number of years of income for the father increases. The elasticity is less than 0.3 using only two years of father's income but climbs as high as 0.6 when using 16 years of father's income. Mulligan (1997) uses the PSID to measure the intergenerational elasticity of a number of relevant factors using an instrumental variable approach to overcome measurement error. A regression of child wages on parent's wages yields an estimated coefficient of between 0.32 (OLS) and 0.53 (IV). Family income has a coefficient between generations of between 0.47 (OLS) and 0.71 (IV), while family consumption has a coefficient of between 0.54 (OLS) and 0.77 (IV)<sup>7</sup>.

Hertz (2005) constructs non-linear transition matrices from parental income deciles to child's deciles using the PSID. Adjusting income for age, he finds that a child with a parent in the bottom income decile has a 31.5% chance of staying in that bottom decile and over a 50

---

<sup>7</sup>In Mulligan's instrumental variable regressions he uses family income as an instrument for family consumption which is measured with a lot of noise, as well as using instruments such as occupation and school categories as instruments for lifetime parental income.

percent chance of being in the bottom quintile. On the other extreme, a child with a parent in the top income decile has a 29.6% chance of staying in that top decile and a 43.3% chance of remaining in the top quintile. In this paper, I will focus on four key intergenerational income transition rates. The poverty trap measures how likely it is for children born in the bottom quintile of the income distribution to stay there as adults while the affluence net measures how likely children born in the top quintile will stay on top. The jump up rate measures how often children born in the bottom quintile are able to jump up to the top quintile of the income distribution, while the fall down rate measures how many children from the affluent top quintile fall down to the bottom quintile. Because of significant racial differences in the transition rates, in my numerical calibrations I focus on matching the distribution of white families.

## 4 Numerical Solution and Calibration Results

### 4.1 Parameter Values and Calibration Targets

This section presents a numerical solution and results from simulations of the model, comparing the model with loss aversion to that with reference consumption levels but with utility that is concave everywhere. Table 2 presents parameter values and their sources that are common to both versions of the model<sup>8</sup>.

[Table 2 about here.]

I calibrate wage parameters,  $\alpha_1, \alpha_2, \zeta$ , and  $\kappa$  from equation (2) to match median income by education and the education distribution for white full-time workers. I divide education

---

<sup>8</sup>In his review of the loss aversion literature, Camerer (2005) finds empirical support for a loss aversion parameter between 1.5 and 3.0. In this model, a lower loss aversion parameter does reduce the elasticity numbers reported below, mainly due to the assumption of strong loss aversion which requires  $\nu < \ln(\lambda)$  in equation (6).

into six possible levels ranging evenly from 0 to 1.<sup>9</sup> Because behavior is significantly different for loss averse agents as opposed to those with concave preferences, I calibrate the model separately for these two utility functions. For the baseline calibration, I allow agents to borrow the entire direct cost of their education so that  $\Omega = -\varepsilon(e)$ . The observed targets and calibrated moments are presented in Table 3.

[Table 3 about here.]

Despite the fact that the model is under identified, both models do a fairly good job of matching the targeted moments. They suffer at the extremes as neither is able to generate a large enough percentage of the population investing in the top level of education despite wages that are higher than in the data. This is likely due to the simplifying assumption that investing in a professional/PhD in the education period consumes all of the agent's time so that they are unable to earn anything in the first period<sup>10</sup>.

I report the calibrated parameters from equation (2) in Table 4. In order to determine whether or not these are reasonable values, I also present the minimum, average, and maximum return to the equivalent of one year of education and the ratio of the maximum wage to the minimum. The return to education averages 8.4% in the loss-averse model and 7.3% in the concave utility model. The labor literature generally finds a value between 5 and 10 percent, so these values seem reasonable.

[Table 4 about here.]

## 4.2 Education Decision Rules Solution

Based on (11), the resource maximizing level of education,  $e^*$  (as opposed to the utility maximizing level,  $\hat{e}$ ), will depend only on an agent's earning ability. An individual's wage

---

<sup>9</sup>Cooley (1997) and Cooley & Prescott (1995) give a good overview of the process and goal of calibrations. To find my calibrated values, I use a simple Euclidean norm of the difference between target and calibrated moments of the 12 target moments and search over a grid that varies by values of 0.005.

<sup>10</sup>While I'm sure most of my readers would prefer the wage distribution in Table 3, in the model only those who are already facing a first-period loss will invest in the highest level of education despite the high (calibrated wages).

is completely determined by his level of education and earning ability, while the cost of education (both direct and opportunity) depends only on the level of education chosen. This resource-maximizing level of education,  $e^*$  is a monotonically increasing function of earning ability as those with more earning ability invest more in education. Because of the differences in the calibrated parameter values for the wage function shown in Table 4, the level of  $e^*$  will be slightly different for the two models.

The education level chosen to maximize lifetime utility,  $\hat{e}$ , may be less than or equal to  $e^*$ . Figure 2 graphs the decision rule for  $\hat{e}$ , (as a function of reference consumption) for a loss-averse agent with an ability level approximately one standard deviation above the mean, for different levels of bequest. For this individual, the level of education that would maximize lifetime resources is  $e^* = 1$ , the maximum level of education. However, those with an intermediate reference consumption level will optimally choose a lower, and sometimes substantially lower, level of education. Intuitively, agents with a moderate reference level of consumption can more easily replace it with a lower level of education, thus avoiding a loss in period 1. Choosing a level of education less than  $e^*$  allows them to minimize or eliminate the loss in period one. Overall lifetime utility will be increased, even if choosing  $\hat{e} < 1$  means a lower (or non-existent) gain in period two.

As reference consumption increases, the likelihood of a loss in the first period increases (until it is equal to 1). At very high levels of reference consumption, agents cannot avoid losses in period 1 and can only hope to eliminate or reduce losses in period 2. In order to decrease or eliminate the loss in period two, they will maximize second period income by investing in the resource-maximizing level of education,  $e^*$ , so that those from higher-income families will themselves have higher income.

Note that those with the lowest reference consumption will invest in higher levels of education than those directly to their right in the reference consumption distribution. This is because with very low levels of reference consumption it is easier to avoid a loss in the first period, even when getting more education. The dashed line at  $c_0 = 2$  is the equilibrium

minimum reference consumption for agents in the model simulation. To the left of this line agents would invest in more education, but agents never visit this region of the state space in equilibrium.

[Figure 2 about here.]

The education investment decision under concave preferences is quite different. When agents can borrow to finance the direct cost of education,  $\varepsilon(e)$ , as in the benchmark model, the decision rule is simply a horizontal line at  $e^*$ . If there is a more severe borrowing constraint, education will be upward sloping in  $c_0$  until the constraint no longer binds and the agent can invest in the optimal level of education. The J-shaped education investment decision for loss averse agents and the contrasting flat or upward sloping investment-decision rule for concave utility agents is the main econometrically testable implication of the model:

*H1: The education investment decision for children will depend on childhood family income and will initially decrease and then increase as family consumption increases.*

Note that this implies that students can borrow the direct cost of education. If those at the bottom face a borrowing constraint, then even with loss aversion we would only expect an upward sloping education decision rule (the J would not turn up at the bottom).

### 4.3 Matching Elasticities Across Generations

As discussed above, loss aversion contributes to a positive intergenerational elasticity of income for two reasons. First, those with moderately low parental income (and thus reference consumption) have an incentive to invest in less education in order to avoid losses in the first period of life even though this will reduce gains in the second period due to their lower wage. Second, those with higher reference consumption are likely to be faced with a loss in period one no matter what level of education they choose. In order to avoid or minimize a loss in period two, they will invest in more education in period one and raise their wage and

consumption in period two. Thus, the same factors that contribute to a positive elasticity of earnings across generations are also likely to contribute to a positive elasticity of education.

Table 5 lists intergenerational elasticities for wages, education, consumption, and bequests, allowing us to compare the simulated results to the data. The data for the intergenerational elasticity of income comes from Mazumder (2005) while the elasticities of education and consumption comes from Mulligan (1997). I present results for both types of utility functions as well as three different borrowing constraints and with  $\rho = 0$  so that earning ability is independent across generations<sup>11</sup>. When there is only a lifetime borrowing constraint as in equation (7a) (columns (1) and (4)), all agents choose the resource-maximizing education level for both models, and because earning ability is independent across generations, there is a zero elasticity of both income and education. With a borrowing constraint in which agents are able to borrow the full direct cost of education, as in equation (7b) (columns (2) and (5)), the loss averse model is able to generate an elasticity of income of 0.14 and an elasticity of education of 0.24, while the concave utility model still exhibits roughly zero elasticity of both income and education. Finally, I present results when the borrowing constraint is a function of reference consumption,  $\Omega = f(c_0)$  as in (7c). For simplicity, I assume that those with reference consumption in the bottom twenty percent do not have access to financial markets and so cannot borrow at all while everybody else can borrow the direct cost of education. Only in this case is the model with concave utility able to generate a positive elasticity of income.

[Table 5 about here.]

Consumption has a positive elasticity across generations in both versions of the model albeit in different ways. Because utility is measured in reference to the consumption level of the individual's parents, both models generate a a positive elasticity of consumption. However, in the model with concave utility, because utility is everywhere concave so that

---

<sup>11</sup>When  $\rho > 0$ , as it almost surely is, the pattern of the results in Tables 5 and 6 stays the same, but there is a higher built-in intergenerational elasticity.

marginal utility is only slightly higher for a loss than for a gain, agents try to smooth consumption as much as possible between period 1 and period 2. This leads to an elasticity of consumption that is approximately the same in period two as in total (period 1 plus period 2). In the model with loss aversion, however, an agent facing an unavoidable loss will lump the loss in period 1 so as to reduce or eliminate the loss in period 2. This leads to a much higher elasticity of consumption in period 2 than in total (or in period 1).

The model with loss aversion generates a positive intergenerational elasticity of income primarily due to an affluence net. Table 6 presents the transition probabilities in the data and in the model under various assumptions of preferences and borrowing constraints. When there are no borrowing constraints (columns (1) and (3)), most transition probabilities are close to 20%, indicating no poverty trap or affluence net. The model with loss aversion and the mild borrowing constraint (column 2) has an affluence net, while none of the versions of the model generate a significant poverty trap in which the children of the poor stay poor, similar to the data for white families in the United States.

[Table 6 about here.]

## 5 Empirical Tests

As noted in *H1*, one important difference between the model with loss aversion preferences and the model with concave preferences is the shape of the education decision rule. With loss aversion, the decision rule is J-shaped in reference (or family) consumption. In the model with concave utility, education is either flat or upward sloping in reference consumption, depending on the borrowing constraints. I test these implications using the National Longitudinal Survey of Youth (NLSY), specifically the NLSY97, which began in 1997 and first surveyed the young respondents between the ages of 12 and 17. The NLSY97 collects data on family income, education levels of parents, and the highest level of education completed of the respondent.

I use the NLSY rather than the PSID because the former includes a measure of ability, the Peabody Individual Achievement Test (PIAT). Controlling for ability is important for estimating the education decision rule. Most took the PIAT in 1997 when they were between the ages of 12 and 17. The PIAT consists of six subtests, covering such areas as reading recognition and comprehension and mathematics, designed to help diagnose learning disabilities. PIAT results are reported in percentile scores, ranging from 0 to 99. Table 7 reports summary statistics on test scores and other variables in the NLSY97.

[Table 7 about here.]

Because the NLSY lacks a measure of consumption, I am forced to use total parental income as a proxy for reference consumption. While this does not seem unreasonable, it will bias the estimated decision rule to the extent that family consumption is different than income. I use the average family income for the respondent before his eighteenth birthday. The NLSY97 has at least two years of family income data for all respondents (those who were 17 at the start) and a maximum of seven years for those who entered at age 12. I convert nominal income in the data to real using the CPI deflator from the BLS<sup>12</sup>.

Regression results are reported in Table 8. The dependent variable is the highest grade completed, measured in total years of education. The main explanatory variables are the natural log of family income and its square. I use robust standard errors to correct for possible heteroskedasticity as the model calls for using family consumption as opposed to the available family income. Custom sample weights provided by the NLSY were used in all regressions. Column (1) provides support for a quadratic education decision rule in the NLSY97 data set when controlling only for race and sex. The linear term in average family

---

<sup>12</sup> The main weakness of the NLSY97 data set is that respondents were only between the ages of 24 and 29 during the last available wave of interviews in 2009 and so may not be done with their educational investment.



income is negative while the quadratic term is positive, significant at the 1% level. The theoretical model implies that one should control for ability as well. The results presented in column (2) control for ability, using the PIAT score. As expected, the coefficient on ability is positive (not shown) and significant at 1%. The coefficients on both linear and quadratic family income continue to have the sign implied by the loss aversion model and are still significant at 1%.

While the model does not imply the need to control for parental education levels, there could be a number of exogenous reasons to expect that parental education may have a positive correlation with child education. Columns (3) and (4), show significant coefficients on both family income and its square, even when controlling for parent education and ability level. The difference between columns (3) and (4) is that column (3) controls for the education of the parents the child lives with (for example a mother and step-father) while column (4) controls for the respondent's biological parents.

[Table 8 about here.]

The main weakness of the results in Table 8 is that the left side of the parabola implied by the coefficients is supported only by approximately 7% of observations, or those with average family income below \$7,600. One possibility is that these represent mismeasured observations. However, households with under \$10,000 per year in income occur with a similar frequency in both the Current Population Survey (7.9% of households with dependent children) and the Consumer Expenditure Survey (8.4% of households), with about half of those having income less than \$5,000. This begs the question of how families can survive in the United States on such low income with the answer probably including consumption transfers from the government and/or family and friends which may not be reported as income. To approximate a more realistic value of family consumption, I construct a measure of consumption such that  $c_i = \max(c_{min}, income_i)$ . I repeat the regression in column (3) of Table 8 and use three values of  $c_{min} = \{\$10,000, \$15,000, \$20,000\}$  with the results

reported in Table 9. The results are strongest for  $c_{min} = \$10,000$  with both coefficients strongly significant and the minimum of the parabola at \$14,000 with approximately 13% of the observations on the left side of the J. Given this result, it is not surprising that the regression coefficients lose their significance when we set a minimum consumption level above this level.

[Table 9 about here.]

Another way to estimate the influence of family income on education investment decisions is to use a semi-parametric regression in which the regression does not place a linear or quadratic restriction on the family income variable as in equation (12):

$$ed_i = f(inc_i) + x_i\beta + \varepsilon_i \tag{12}$$

Where  $x_i$  are the controls for ability, parental education, and demographics and the function  $f(\cdot)$  is assumed to be smooth with a bounded first derivative (see Lokshin, 2006). Figure 3(a) provides the estimated function,  $f(inc_i)$  for the original reported values of family income while 3(b) provides the estimated function using  $c_{min} = \$10,000$  as in Table 9. The J-figure is evident in both, although certainly more prominent in the original data. 3(a) implies a minimum of approximately \$8,000 while 3(b) implies a minimum around \$24,000.

[Figure 3 about here.]

## 6 Conclusion

A simple model with loss-averse preferences and education investment can generate a positive and significant elasticity of intergenerational income, even when agents are able to borrow the entire (direct) cost of their education and earning ability is completely independent across generations. This expands on earlier work that required a much stronger

(and less realistic) borrowing constraint to generate a positive elasticity of income between generations. Previous empirical work (see Bowles et al, 2005) has shown that observable factors such as IQ, schooling, and personality, can only explain approximately half of the currently accepted value of the intergenerational elasticity of income of between 0.45 and 0.6. Loss averse preferences can explain about half the unexplained elasticity (0.14). A model with concave preferences can only generate similar levels of the elasticity of intergenerational income when there are severe borrowing constraints. While some students, especially in historically disadvantaged groups, may face difficulty in borrowing for post-secondary schooling, this does not seem to describe the experience of most students in the U.S.

The baseline model with loss aversion and weak borrowing constraints generates a positive elasticity of income mainly by generating an affluence net. Children born at the top of the income distribution tend to get more education in order to avoid losses in the second period of life. The loss averse model is able to generate an affluence net and a percentage of children who jump up from the bottom quintile to the top that match those rates found by Hertz (2005) and Mazumder (2008) for white families.

While precise welfare statements are not possible in this partial equilibrium model, it seems likely that loss aversion may generate behavior that is socially inefficient, in that a sizable portion of the population invests in less education than is socially optimal because of borrowing constraints and loss-averse preferences. If that's the case, there may be a role for public policy. Many industrialized countries have eliminated the direct cost of post-secondary education for those who qualify. While this would help reduce the inefficiency, additional policies, such as increasing grants and loans to students, may also improve welfare. The United Kingdom conducted a pilot experiment in 1999 in which it paid students to stay in school after the age of 16. Dearden et al (2007) found that these students did stay in school longer than a control group.

Loss aversion has been found to help explain phenomena as diverse as trade policy (Tovar, 2009), asset pricing (Yogo 2008), physician behavior (Rizzo and Zeckhauser, 2003), and

even the behavior of capuchin monkeys (Silberberg et al, 2008). Page et al (2007) conduct an experiment in which people are asked to make a costly education-like investment in order to increase future earnings and find that people do have loss-averse preferences when making these types of investments. In this paper, I show that when agents have loss-averse preferences, there is a J-shaped education investment decision in reference consumption, which is not the case if preferences are concave everywhere. Using the NLSY97 data set I find support for this J-shaped decision rule, with those at the lower end of the parental income distribution investing more in education than those immediately above them.

## References

- Becker, Gary S. and Nigel Tomes (1986) “Human Capital and the Rise and Fall of Families,” *Journal of Labor Economics*, Vol. 4, No. 3, pp. S1–S39.
- Bowles, Samuel, Herbert Gintis, and Melissa Osborne Groves (2005) *Unequal Chances: Family Background and Economic Success*: Princeton University Press.
- Bowman, David, Deborah Minehart, and Matthew Rabin (1999) “Loss Aversion in a Consumption-Savings Model,” *Journal of Economic Behavior and Organization*, Vol. 38, pp. 155–178.
- Brown, Ryan, William E. Copeland, E. Jane Costello, Alaattin Erkanli, and Carol M. Worthman (2009) “Family and Community Influences on Educational Outcomes Among Appalachian Youth,” *Journal of Community Psychology*, Vol. 37, No. 7, pp. 795–808.
- Camerer, Colin (2005) “Three Cheers—Psychological, Theoretical, Empirical—For Loss Aversion,” *Journal of Marketing Research*, Vol. 42, No. 2, pp. 129–133, May.
- Connor, Helen (2001) “Deciding For or Against Participation in Higher Education: the Views of Young People From Lower Social Class Background,” *Higher Education Quarterly*, Vol. 55, No. 2, pp. 204–224.
- Cooley, Thomas F. (1997) “Calibrated Models,” *Oxford Review of Economic Policy*, Vol. 13, No. 3, pp. 55–69.
- Cooley, Thomas F. and Edward C Prescott (1995) *Frontiers of Business Cycle Research*, Chap. Economic Growth and Business Cycles, pp. 1–38: Princeton University Press.
- De la Croix, David (2001) “Growth Dynamics and Education Spending: The Role of Inherited Tastes and Abilities,” *European Economic Review*, Vol. 45, pp. 1415–1438.

- Dearden, Lorraine, Carl Emmerson, Christine Frayne, and Costas Meghir (2007) “Education Subsidies and School Drop-out Rates,” *The Institute of Fiscal Studies*, Vol. WP05/11.
- Galor, Oded (2009) *Inequality and Economic Development: The Modern Perspective*, Chap. Inequality and the Development Process: An Overview, pp. ix–xxiii: Edward Elgar Pub.
- Galor, Oded and Joseph Zeira (1993) “Income Distribution and Macroeconomics,” *The Review of Economic Studies*, Vol. 60, No. 1, pp. 35–52.
- Genicot, Garance and Debraj Ray (2009) “Aspiration, Inequality, Investment, and Mobility.” Working Paper.
- Hahn, Ryan D. and Derek Price (2008) “Promise Lost: College-Qualified Students Who Don’t Enroll in College,” *A Report by Institute for Higher Education Policy*, Vol. November.
- Hertz, Tom (2005) *Unequal Chances: Family Background and Economic Success*, Chap. Rags, Riches, and Race: The Intergenerational Economic Mobility of Black and White Families in the United States, pp. 165–191: Princeton University Press.
- Isaacs, Julia B., Isabel V. Sawhill, and Ron Haskins (2008) “Getting Ahead of Losing Ground: Economic Mobility in America,” Technical report, The Brookings Institute and Pew Center’s Economic Mobility Project.
- Kahneman, Daniel and Amos Tversky (1979) “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica*, Vol. 47, pp. 263–291.
- Kobberling, Veronica and Peter P. Wakker (2005) “An Index of Loss Aversion,” *Journal of Economic Theory*, Vol. 122, pp. 119–131.
- Lemieux, Thomas (2006) “Post-Secondary Education and Increasing Wage Inequality,” *The American Economic Review*, Vol. 96, No. 2, pp. 1–23.
- Lokshin, Michael (2006) “Difference-Based Semiparametric Estimation of Partial Linear Regression Models,” *The Stata Journal*, Vol. 6, No. 3, pp. 377–383.
- Mazumder, Bhashkar (2005) “Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data,” *The Review of Economics and Statistics*, Vol. 87, No. 2, pp. 235–255.
- (2008) “Upward Intergenerational Economic Mobility in the United States,” economic mobility project, Pew Charitable Trusts.
- McGrath, Daniel J., Raymond R. Swisher, Glen H. Jr. Elder, and Rand D. Conger (2001) “Breaking New Ground: Diverse Routes to College in Rural America,” *Rural Sociology*, Vol. 66, No. 2, pp. 244–267.
- Mulligan, Casey (1997) *Parental Priorities and Economic Inequality*: University of Chicago Press.

- National Center for, Education Statistics (2008) “2006-2007 and 2007-2008 Integrated Post-secondary Education Data System (IPEDS),” *U.S. Department of Education*.
- Neilson, William S. (2002) “Comparative Risk Sensitivity with Reference-Dependent Preferences,” *The Journal of Risk and Uncertainty*, Vol. 24, No. 2, pp. 131–142.
- Owen, Ann L. and David N. Weil (1998) “Intergenerational earnings mobility, inequality and growth,” *Journal of Monetary Economics*, Vol. 41, No. 1, pp. 71–104.
- Page, Lionel, Louis Levy Garboua, and Claude Montmarquette (2007) “Aspiration levels and educational choices: An experimental study,” *Economics of Education Review*, Vol. 26, pp. 748–758.
- Piketty, Thomas and Emmanuel Saez (2003) “Income Inequality in the United States, 1913-1998,” *Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 1–39. Updated online through 2007 at <http://elsa.berkeley.edu/saez>.
- Rizzo, John A. and Richard J. Zeckhauser (2003) “Reference Income, Loss Aversion, and Physician Behavior,” *The Review of Economics and Statistics*, Vol. 85, No. 4, pp. 909–922.
- Roderick, Melissa, Jenny Nagaoka, Vanessa Coca, and Eliza Moeller (2008) “From High School to the Future: Potholes on the Road to College,” Technical report, Consortium on Chicago School Research at the University of Chicago.
- Shea, John (2000) “Does Parents’ Money Matter?” *Journal of Public Economics*, Vol. 77, No. 2, pp. 155–184.
- Silberberg, Alan, Peter G. Roma, Mary E. Huntsberry, Frederick R. Warren-Boulton, Takayuki Sakagami, Angela M. Ruggiero, and Stephen J. Suomi (2008) “On Loss Aversion in Capuchin Monkeys,” *Journal of the Experimental Analysis of Behavior*, Vol. 89, No. 2, pp. 145–155.
- Tovar, Patricia (2009) “The Effects of Loss Aversion on Trade Policy: Theory and Evidence,” *Journal of International Economics*, Vol. 78, No. 1, pp. 145–167.
- Tversky, Amos and Daniel Kahneman (1991) “Loss Aversion in Riskless Choice: A Reference-Dependent Model,” *The Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 1039–1061, November.
- (1992) “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty*, Vol. 5, pp. 297–323.
- Waldkirch, Andreas, Serena Ng, and Donald Cox (2004) “Intergenerational Linkages in Consumption Behavior,” *The Journal of Human Resources*, Vol. 39, No. 2, pp. 355–381.
- Winter, Christopher (2009) “Accounting for the Changing Role of Family Income in Determining College Entry.” Institute for Empirical Research in Economics, University of Zurich Working Paper No. 402.

Yogo, Motohiro (2008) “Asset Prices Under Habit Formation and Reference-Dependent Preferences,” *Journal of Business & Economic Statistics*, Vol. 26, No. 2, pp. 131–143.

Table 1: Consumption and Bequest Patterns for a Given Level of Resources

$R^{lt}$	$R_{max}^1$	$c_1$	$c_2$	$T'$	Description
$\geq 2c_0$	$\geq c_0$	$\frac{R^{lt}+c_0}{3}$	$\frac{R^{lt}+c_0}{3}$	$\frac{R^{lt}-c_0}{3}$	Spread gains
$c_0 \leq R^{lt} < 2c_0$	$< c_0$	$R_{max}^1$	$\frac{R_{min}^2+c_0}{2}$	$\frac{R_{min}^2-c_0}{2}$	Borrowing constrained
		$0 \leq c_1 \leq c_0$	$c_0$	0	Concentrate loss in period 1
	$< c_0$	0	$R^{lt}$	0	Concentrate loss in period 1



Table 2: Parameter Values

Parameter	Value	Description	Source/Explanation
$\lambda$	2.25	Loss aversion parameter	Tversky and Kahneman (1992)
$\beta = R = (1 + r)$	1	Rate of time preference & interest rate	Simplicity
$\phi_1$	0.0	Linear term in cost of education	Estimated from Dept. of Education,
$\phi_2$	0.3	Quadratic term in cost of education	National Ctr for Education Statistics
$e, \psi$	[0,1]	Range for education and earning ability	Normalization
$\bar{\psi}$	0.5	Average ability level	
$\sigma_\psi$	0.15	Standard deviation of earning ability, $\psi$	3 st. dev. above & below the mean

Table 3: Target and Calibrated Moments

Moment	Target	Loss Aversion	Concave Utility
Education Distribution			
High School Dropouts	7.9%	8.8%	8.7%
High School Grads	28.5%	23.8%	29.1%
Some College/Assoc. Degree	27.8%	27.5%	24.4%
Bachelor's Degree	23.1%	31.0%	25.3%
Master's Degree	9.2%	7.2%	11.1%
Prof. Degrees/PhD	3.5%	1.9%	1.4%
Median Income by Education Level (HS Grad = 1)			
High School Dropout Wage	0.71	0.78	0.77
High School Grad Wage	1.00	1.00	1.00
Some College/Assoc. Degree Wage	1.19	1.57	1.37
Bachelor's Degree Wage	1.60	1.65	1.85
Master's Degree Wage	1.80	2.10	2.75
Prof. Degrees/PhD Wage	2.71	4.04	4.57

Table 4: Calibrated Parameter Values

Parameter	Loss Aversion	Concave Utility
$\alpha_1$ : linear term on education	0.285	0.250
$\alpha_2$ : quadratic term on education	0.190	0.000
$\zeta$ : interaction term (education*ability)	0.655	0.900
$\kappa$ : linear term on ability	1.450	1.200
Minimum marginal return to education	3.1%	2.5%
Average marginal return to education	8.4%	7.3%
Maximum marginal return to education	13.9%	12.2%
Max-Min Wage Ratio	13.2	10.5

Table 5: Intergenerational Elasticities

	Data	Loss Aversion			Concave Utility		
		(1) No BC	(2) $\Omega_{ed} = \varepsilon(e)$	(3) $\Omega_p = f(c_0)$	(4) No BC	(5) $\Omega_{ed} = \varepsilon(e)$	(6) $\Omega_p = f(c_0)$
Income	0.60	0.00	0.14	0.17	0.00	-0.00	0.14
Education	0.29	0.00	0.24	0.33	0.00	-0.01	0.26
Consumption (total)	0.68	0.09	0.16	0.11	0.32	0.36	0.36
Consumption (period 2)		0.79	0.76	0.80	0.33	0.32	0.39
Bequests	NA	-0.01	-0.05	-0.02	-0.12	-0.18	-0.10

Table 6: Intergenerational Mobility

	<b>Data*</b>	<b>Simulations</b>					
		<i>Loss Aversion</i>			<i>Concave Utility</i>		
		(1)	(2)	(3)	(4)	(5)	(6)
White Families	No BC	$\Omega_{ed} = \varepsilon(e)$	$\Omega_p = f(c_0)$	No BC	$\Omega_{ed} = \varepsilon(e)$	$\Omega_p = f(c_0)$	
Elasticity of Income		0.001	0.142	0.171	0.000	-0.001	0.144
Poverty Trap	24.9%	23.4%	22.6%	23.5%	22.8%	23.4%	23.6%
Affluence Net	38.9%	23.4%	46.4%	42.0%	23.2%	23.2%	30.2%
Jump Up	10.6%	23.3%	12.7%	11.8%	23.2%	23.7%	5.8%
Fall Down	10.4%	22.9%	23.6%	23.2%	23.4%	23.7%	22.8%

\* Source: Mazumder (2008)

Table 7: Summary Statistics for NLSY79 and NLSY97 Data Set

Variable	NLSY97				
	# Obs	Mean	St. Dev.	Min	Max
R's Highest Grade	8,853	12.98	2.62	7	20
Family Income (log)	7,474	9.89	1.077	1.14	12.47
PIAT	6,044	70.32	17.39	0	100
Mother's Educ (bio)	8,255	12.48	2.83	3	20
Father's Educ (bio)	7,100	12.59	3.17	3	20
Mother's Educ (res)	7,974	12.57	2.88	3	20
Father's Educ (res)	5,691	12.90	3.26	3	20
Black		25.99%			
Hispanic		21.16%			

Table 8: Regression Results for Highest Grade Completed

Variable	NLSY97			
	(1)	(2)	(3)	(4)
Family Income (log)	-2.562*** (0.261)	-2.243*** (0.330)	-2.047*** (0.395)	-1.172*** (0.440)
Family Income (log) Squared	0.175*** (0.013)	0.150*** (0.016)	0.115*** (0.020)	0.082*** (0.021)
<b>Controls</b>				
AFQT (79) / PIAT (97)		✓***	✓***	✓***
Parent's Education			✓(res)***	✓(bio)**
Sample Size	7,401	5,093	3,039	4,014
$R^2$	0.160	0.260	0.299	0.304

Significant at: \*=10%, \*\*=5%, \*\*\*=1%

Robust standard errors in ( )

All regressions control for race and sex and use custom sample weights provided by the NLSY.

Table 9: Regression Results with Minimum Consumption Level

Variable	NLSY97		
	(1)	(2)	(3)
$C_{min}$	\$10,000	\$15,000	\$20,000
Est. Family Consumption (log)	-3.124** (1.263)	-2.393 (1.498)	-1.817 (1.724)
Est. Family Consumption (log) Squared	0.164*** (0.058)	0.131* (0.067)	0.106 (0.077)
Controls			
PIAT(97)	✓***	✓***	✓***
Parent's Education (res)	✓***	✓***	✓***
Sample Size	3,039	3,039	3,039
$R^2$	0.299	0.299	0.299

Significant at: \*=10%, \*\*=5%, \*\*\*=1%

Robust standard errors in ( )

All regressions control for race and sex and use custom sample weights provided by the NLSY.



Figure 1: Utility and Marginal Utility for Strongly Loss Averse Preferences

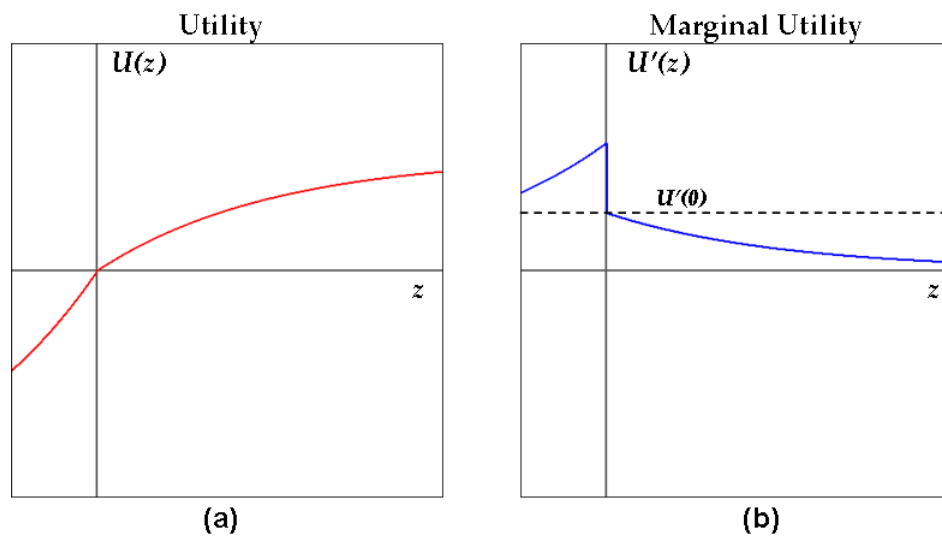


Figure 2: Utility Maximizing Education by Reference Consumption and Bequest Received for Loss-Averse Agent

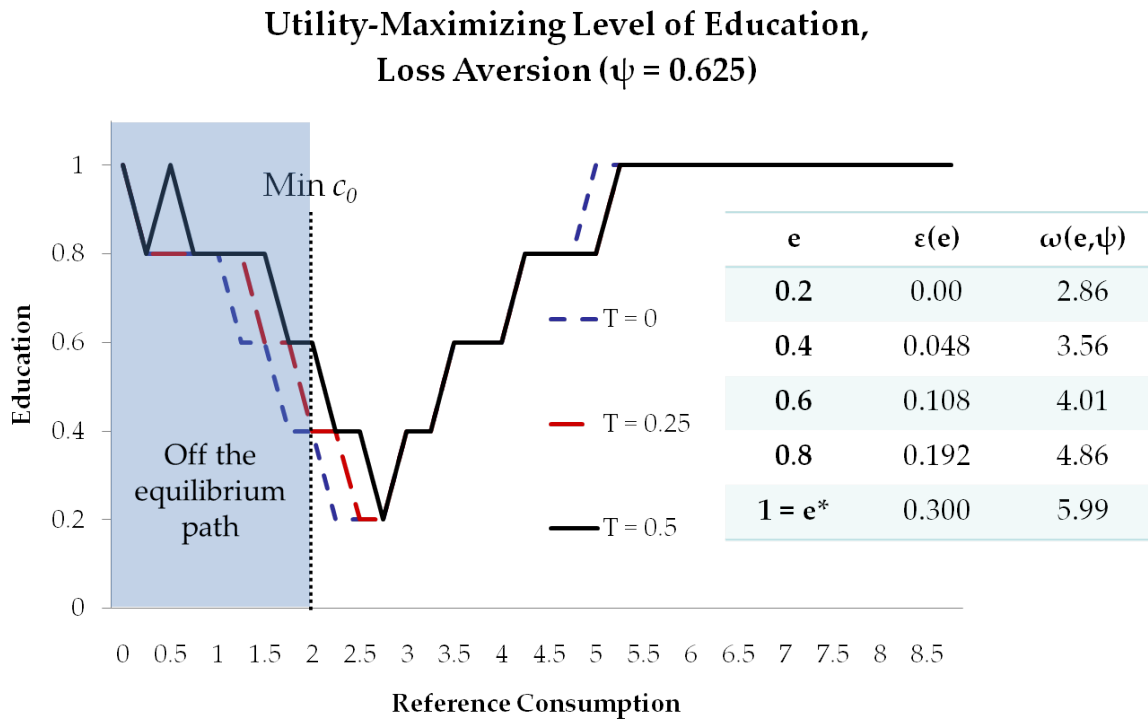


Figure 3: Semi-Parametric Regression Results

