1 Introduction

Most studies of intergenerational mobility use data on two generations to estimate children's earnings as a function of fathers' earnings[1]. Solon (1999) presents a theoretical derivation for this relationship adding the assumption of identical children's endowment or that the child's endowment is not related to the parent's endowment to a variant of the Becker and Tomes (1979) model. (Examples of child's endowment are family connections, ability, race, skills, genetic characteristics and family culture.)

The Becker and Tomes model establishes additional relationships between the child's earnings and family background. For instance, the model predicts a negative relationship between the child's earnings and the grandparent's earnings when controlling for the parent's earnings. The model in this study yields more intuitive predictions by modifying some assumptions of the original Becker and Tomes model. For instance, the assumption that the child's endowment depends on the parent's endowment is replaced by the more relaxed assumption that it depends on the parent's and grandparent's endowment.

I use data on two and three generations from *Pesquisa Nacional por Amostra de Domicilios* (PNAD)–a nationally representative household survey for Brazil. A three-generations data set allows me to test the predictions established by both models–the original and the modified version–relaxing the assumption of identical children's endowment. In addition, a three-generations data set allows me to estimate a lowerbound for the variation in earnings explained by differences in endowments across families, and an upperbound for the variation in earnings explained by differences in human capital.

The mobility supplement of the 1996 PNAD provides data on 36,705 father-son pairs. I build a three-generations data set consisting of 5,125 grandfather-father-son triplets by restricting the sample to households with adult sons present. In Brazil individuals live with their parents until they

[1] Solon (2002), Table 1, Solon (1990), tables 3 and 4, and Corak (2006), Appendix tables, list articles that estimate the elasticity of child's earnings with respect to parental earnings. For Brazil, see Ferreira and Veloso (2006), Dunn (2007) and Marchon (2008).

marry, and quit school and begin working at an early age. There are many households with adult sons who are not at the very beginning of their working careers. A number of statistics support this argument. In the 1996 PNAD sample, at age 25 about 50.9% of the males lived in their parents' house. About 47.46% of the 25-year-old males were married. Among the 25-year-old males living with their parents only 3.74% were married. On average, males aged 20 to 35 have 6.6 years of education and began working at age 13.4.

To address potential sample selection problems, I apply the Heckman (1979) estimation procedure.

The estimation results are consistent with the predictions of the modified model, but contradict some of the predictions of the Becker and Tomes model.

My estimate of the intergenerational elasticity in earnings in Brazil is about 0.84. Dunn's (2007) estimate of the intergenerational elasticity in earnings in Brazil is 0.85 and Ferreira and Veloso's (2006) estimate is 0.58. The estimates for the U.S. range between 0.4 and 0.5.

Estimating the main equation yielded by the modified model, I find that family background explains 34.9% of the variation in earnings among males between the ages of 16 and 27 who live with their parents. If it were possible to eliminate the differences in investment in the children's human capital, the variation in earnings would fall by no more than 21.1%. Additionally, if there were no differences in endowments among children, the variation in earnings would fall by no less than 26%.

This article is organized as follows. The first part of section 2 presents a variant of the Becker and Tomes model according to Solon (1999). The second part of section 2 presents my modification of the Becker and Tomes model. Section 3 describes the data and provides details of the estimation procedure. Section 4 presents the estimation results for the relationships established by both the original and the modified Becker Tomes model. Section 5 concludes.

2 Theoretical Model

The Becker and Tomes Model

In this section I present a variant of the Becker and Tomes (1979) model presented by Gary Solon (1999) in the "Handbook of Labor Economics". I have added some obvious derivations because they are relevant to this paper.

Consider the following assumptions and notation. A family consists of one child and one parent. Let y_t and c_t represent the parent's lifetime earnings and consumption, and let I_{t+1} represent investment at time t in the child's earning capacity at time $t+1$. The parent's budget constraint is given by $y_t = c_t + I_{t+1}$. The child's earnings depend on the parent's investment according to the equation $y_{t+1} = (1+r)I_{t+1} + A_{t+1}$, where r is the return to human capital investment and A_{t+1} represents all other determinants of the child's earnings besides human capital investment. Also $A_{t+1} = a_{t+1} + u_{t+1}$, where a_{t+1} represents the child's endowment of earning capacity, for instance, family connections, ability, race, skills, genetic heritage and family culture. The second component, u_{t+1}, is the child's market luck. u_t is uncorrelated with y_t and a_{t+1}. The child's endowment, a_{t+1}, is positively correlated with the parent's endowment, a_t. Becker and Tomes assume an AR(1) process where $a_{t+1} = \delta a_t + w_{t+1}$, $\delta \in [0,1)$ and w_{t+1} is white noise. Parents know their children's endowment and market luck at time t. Parents maximize the utility function $u(y_{t+1}, c_t) = (1-\beta)\ln(c_t) + \beta\ln(y_{t+1})$. Consider all variables in deviation from mean form.

At time t, parents solve the maximization problem

$$\begin{aligned} &Max\{(1-\beta)\ln(c_t) + \beta ln(y_{t+1})\} \\ &\text{subject to} \quad c_t = y_t - I_{t+1} \\ &\qquad\qquad\quad\; y_{t+1} = (1+r)I_{t+1} + A_{t+1}. \end{aligned}$$

The solution to this optimization problem is

(1) $$I_{t+1} = \beta y_t - \tfrac{(1-\beta)}{1+r}A_{t+1}.$$

Substitute the above equation in the child's earnings equation to obtain,

(2) $$y_{t+1} = (1+r)\beta y_t + \beta A_{t+1}.$$

Substitute A_{t+1} by $a_{t+1} + u_{t+1}$ to obtain

(3) $$y_{t+1} = (1+r)\beta y_t + \beta a_{t+1} + \beta u_{t+1}.$$

Substitute a_{t+1} by $\delta a_t + w_{t+1}$ in equation (3) to obtain the child's earnings as a function of family background (y_t, a_t) and child's market and endowment luck (w_{t+1}, u_{t+1}),

(4) $$y_{t+1} = (1+r)\beta y_t + \beta\delta a_t + \beta w_{t+1} + \beta u_{t+1}.$$

From the above derivation it is clear that, in this model, family background drives child's earnings through two channels. First, the child's earnings depend on human capital investment that depends on parent's earnings. Second, the child's earnings depend on child's endowment that depends on parent's endowment.

Since a_t and a_{t+1} are usually unobservable variables, the focus from now on will be on relationships predicted by the model that do not include those two variables. The term A_{t+1} in equation (2) can be eliminated solving equation (1) for A_{t+1} and plugging it in equation (2),

(5) $$y_{t+1} = \frac{(1+r)\beta}{(1-\beta)} y_t - \frac{(1+r)\beta}{(1-\beta)} I_{t+1}.$$

Note that, $\frac{\partial y_{t+1}}{\partial I_{t+1}} < 0$ in the above equation. Also, the coefficient on y_t is equal to the coefficient on I_{t+1}, except for the latter's negative sign.

The model also yields a relation between child and grandparent's earnings that does not include endowment of earnings capacity. First, consider equation (4) one period ahead,

$$y_{t+2} = (1+r)\beta y_{t+1} + \beta\delta a_{t+1} + \beta w_{t+2} + \beta u_{t+2}.$$

Then, solve equation (4) for a_{t+1} and substitute in the above equation to obtain

(6) $$y_{t+2} = [(1+r)\beta + \delta] y_{t+1} - (1+r)\beta\delta y_t + \phi_{t+2},$$

where $\phi_{t+2} = \beta(w_{t+2} + u_{t+2} - \delta u_{t+1})$.

Note that, in the above equation, parents' market luck (u_{t+1}) is related to parents' earnings (y_{t+1}). Becker and Tomes (1993) considered the above equation in a model with no variance in market luck. Under this assumption the model predicts a negative relation between y_{t+2} and y_t.

Assuming $\boldsymbol{\delta = 0}$ ***or*** $\mathbf{Var(a_{t+1}) = 0}$

Ignoring the possibility that children inherit part of the parent's endowment ($\delta = 0$) or assuming that there is no variance in endowments ($Var(a_{t+1}) = 0$), equation (4) simplifies to

$$(7) \qquad y_{t+1} = (1+r)\beta y_t + \epsilon_{t+1},$$

where $\epsilon_{t+1} = \beta w_{t+1} + \beta u_{t+1}$ if $\delta = 0$ and, $\epsilon_{t+1} = \beta u_{t+1}$ if $Var(a_{t+1}) = 0$[2].

Several papers estimate equation (7) to assess intergenerational mobility. An interesting aspect of the above specification is that the coefficient on parent's earnings can be interpreted as the intergenerational correlation between parent's earnings and child's earnings if the variances of earnings are the same for the parent and child generations. Yet, the assumption that all individuals have the same endowment or the assumption that the child's endowment is not related to parent's endowment can be considered strong assumptions for an intergenerational mobility model[3].

A Modification of the Becker and Tomes Model

The variant of Becker and Tomes model previously presented yields some counterintuitive predictions. For instance, the model predicts that the child's earnings are negatively related to grandparent's earnings when controlling for parent's earnings (see equation (5)). Moreover, the child's earnings are negatively related to the investment in the child's human cap-

---

[2]Equation (i) can also be obtained from equation (3) considering father's earnings as a proxy for child' endowment.

[3]In the same study Solon obtained an equation equivalent to equation (6) considering an alternative model. He assumes that a sibling's lifetime earnings is determined by a family component (f_i) and a sibling specific component (b_{ij}) according to the equation $y_{ij} = f_i + b_{ij}$. The family component can be decomposed into parent i's long run earnings (X_i) and the combined effect of family background characteristics uncorrelated with parent's earnings (z_i), $f_i = \rho X_i + z_i$. So, the earnings of sibling j from family i can be written as $y_{ij} = \rho X_i + \epsilon_{ij}$, where $\epsilon_{ij} = z_i + b_{ij}$. By construction Cov($Xi, \epsilon_{ij}$)=0. But, note that the assumption that X_i is not correlated to z_i is equivalent to the assumption that $\delta = 0$ in the model presented in this subsection.

Assuming $\delta \neq 0$ and $Var(a_{t+1}) \neq 0$ equation (6) can be the obtained solving the parents' maximization problem if parents' preferences are represented by the utility function

$$u(c_t, E(y_{t+1})) = \begin{cases} 0 & \text{if} \quad E(y_{t+1}) < \beta y_{t+1} + \alpha \\ f(c_t) & \text{if} \quad E(y_{t+1}) > \beta y_{t+1} + \alpha, \end{cases}$$

where $f'(c_t) > 0$. These preferences imply that parents compensate perfectly with investments in children's human capital any difference between the children's endowment from their own endowment.

ital after controlling for the parent's earnings (see equation (4)). It is reasonable to imagine that those variables have a positive impact on the child's earnings. In the next two paragraphs, I explain in words how the model in the previous section yields those predictions.

Consider two sets of child-parent-grandparent. Suppose that the two parents have the same earnings but one grandparent is wealthier than the other. The wealthier grandparent had more money to invest in his child than the poor grandparent, and yet their children have the same earnings. In the model this happens because the child of the wealthy grandparent is less endowed than the child of the poor grandfather. In the next generation, the child of the grandparent is a parent. The parents of the second generation of children have the same earnings to invest in their children's human capital. But, the grandchild of the wealthy grandparent is expected to inherit at least part of his parent's relatively low endowment, yielding a negative relation between child's and grandparent's earnings when controlling for the parent's earnings. This happens because in the model the grandparent can only affect the grandchild though the parent, but common sense would suggest that wealthy grandparents may also have a positive direct impact on the grandchild's earnings.

The Becker and Tomes model yields a simple compensation scheme in which parents partially compensate for the child's endowment (or market luck) with investment in human capital. Since the child's earnings are positively related to the child's endowment (or market luck) this compensation yields a negative relationship between the child's earnings and investment in human capital when controlling for the parent's earnings. Different assumptions yields more complex compensation structures.

In fact, the data disagree with the predictions of the original Becker and Tomes model. In section 4, I show that after controlling for the parent's earnings, the impact of the grandparent's earnings on the grandchild's earnings is positive. The same is true for investment in human capital.

In this section I obtain more intuitive predictions by introducing a more general assumption and replacing some assumptions by others more in accordance with economic theory.

I modify three aspects of the previous model. First, I assume parents know their children's endowment but not their children's market luck. Second, I assume that parents' utility function depends on their children's consumption instead of their children's earnings. Third, I assume that the child's endowment depends on the parent's and grandparent's endowment according to the equation $a_{j+2} = \delta a_{j+1} + \theta a_j + w_{j+2}$ for all $j \in \mathbb{N}$.

The first assumption is reasonable if we expect that the realization of the child's market luck happens at the time the child reaches the job market and not before. Solving the parents' maximization problem under this assumption, requires an extra assumption about the parent's preference toward risk. Alone, the assumption that parents do not know the child's market luck combined with the assumption that parents are risk neutral will not change the basic results of the previous model.

Economist believe that consumption, not income, provides utility. That is the reason why I replace the assumption that parents' utility depends on their children's earnings[4] for the assumption that parents' utility depends on children's consumption. It is interesting to note that, since parents' utility depends on their children's consumption and their children' utility depends on their grandchildren's consumption, and so on, parents' utility indirectly depends on all their descendants consumption[5].

The assumption that the child's endowment also depends on the grandparent's endowment is more general. The assumption in the previous section can be reestablished at any point by setting $\theta = 0$. The advantage of a more general assumption is that it introduces a relation between children's and grandparents' earnings that does not work through parents' earnings or endowments. If the parameters of the model satisfy some conditions, this assumption introduces a positive relation between the child's and grandparent's earnings. For specifications that do not require proxies for the child's endowment, this assumption requires at least three generations in order to estimate the importance of family background in explaining variation in

[4]Becker and Tomes assume that parents' utility depends on children's quality or economic success that is measured by the children's wealth.

[5]One could argue that parents care about their children's earnings because they care about the investment in their grandchild. But it is the consumption that today's investment will make affordable in the future that improves utility, not the investment per se.

earnings.

At time t=0, the parent solves the maximization problem

$$\begin{aligned} & Max\{(1-\beta)\ln(c_0)+\beta\ln(c_1)\} \\ \text{subject to} \quad & y_0=(1+r)I_0+A_0 \\ & c_0=y_0-I_1 \\ & y_1=(1+r)I_1+A_1 \\ & c_1=y_1-I_2. \end{aligned}$$

As before, $A_t = a_t + u_t$ for all $t \in \mathbb{N}$ and the expected value of u_t is zero.

Suppose parents are risk neutral and maximize expected values. Let the superscript e_t refer to the parent's expectation at time t. So, the parent's maximization problem can be written as

(8) $$\underset{\{I_1\}}{Max}\ \{(1-\beta)\ln((1+r)I_0+A_0-I_1)+\beta\ln((1+r)I_1+a_1-I_2^{e_0})\}.$$

Assume that parents believe that their descendants' preferences are the same as their own preferences, and the future rates of return to human capital investment will be equal to r. Then, the solution to the maximization problem (8) is

(9) $$I_1=\beta y_0+\tfrac{\beta(1+r)-\delta-r}{r(1+r)}a_1-\tfrac{\theta}{r(1+r)}a_0.$$

Derivations appear in the Appendix.

The model predicts a relationship in which an individual's earnings depends on his/her family background that does not include endowment of earnings capacity as one of the dependent variables. First, in the grandchild's earnings equation, $y_2=(1+r)I_2+a_2+u_2$, substitute a_2 by $\delta a_1+\theta a_0+v_2$ to find

(10) $$y_2=(1+r)I_2+\delta a_1+\theta a_0+v_2+u_2.$$

It is possible to replace the terms a_1 and a_0 in the above equation with functions of I_1, I_2, y_0 and, y_1. Solve equation (9) for one period ahead to obtain

$$I_2=\beta y_1+\tfrac{\beta(1+r)-\delta-r}{r(1+r)}a_2-\tfrac{\theta}{r(1+r)}a_1.$$

Substitute a_2 by $\delta a_1+\theta a_0+v_2$ in the above equation to find

$$I_2=\beta y_1+\tfrac{[\beta(1+r)-\delta-r]\delta-\theta}{r(1+r)}a_1+\tfrac{\beta(1+r)-\delta-r}{r(1+r)}\theta a_0+\tfrac{\beta(1+r)-\delta-r}{r(1+r)}v_2.$$

Use the above equation and equation (9) to find a_1 and a_0 as functions of I_1, I_2, y_0 and, y_1. Substitute those equations in (10) to find

$$(11) \qquad y_2 = \frac{(1+r)}{[\beta(1+r)-\delta-r][\beta(1+r)-r]-\theta}\Big[[(\beta(1+r)-r)(\beta(1+r)-\delta)-\theta]I_2 - r(\beta(1+r)-r)\beta y_1 - r\theta(\beta y_0 - I_1) + \frac{\theta_0}{(1+r)}v_2\Big] + u_2.$$

Note that $\frac{\partial y_2}{\partial I_2} > 0$, $\frac{\partial y_2}{\partial y_1} > 0$, $\frac{\partial y_2}{\partial y_0} > 0$ and, $\frac{\partial y_2}{\partial I_1} < 0$ for $r < \beta(1+r) < \delta$.

In equation (10), the grandchild's earnings are a function of I_2, a_1 and a_0. In order to obtain equation (11) the terms a_1 and a_0 are replaced by functions of I_1, I_2, y_0 and, y_1. The terms I_1, y_0 and, y_1 are introduced for the first time in equation (10) to capture the impact of heritage of endowment on the grandchild earnings. Therefore, the share of variation in grandchildren's earnings explained by variations in I_1, y_0 and y_1 in equation (11) represents a lowerbound for the variation in grandchildren's earnings explained by variation in endowment capacity across families. Analogously, the share of variation in grandchildren's earnings explained by variation in I_2 in equation (11) represents an upperbound for the variation in grandchildren's earnings explained solely by variation in investment in the grandchild's human capital.

***Assuming* $\mathbf{Var(a_{t+1}) = 0}$**

Assuming that $Var(a_j) = 0$ for all $j \in \mathbb{N}$, equation (9) simplifies to $I_1 = \beta y_0$. Substitute it in the the child's earnings equation to obtain

$$(12) \qquad y_1 = (1+r)\beta y_0 + u_1.$$

This equation is equivalent to equation (6). But, again, the assumption that all individuals have the same endowment is strong for an intergenerational mobility model.

Assuming* $\boldsymbol{\theta = 0}$ *(no skipping generations effect)

Assuming that the child's endowment does not depend on the grandparent's endowment ($\theta = 0$), equation (9) simplifies to

$$(13) \qquad I_1 = \beta y_0 + \frac{\beta(1+r)-\delta-r}{r(1+r)}a_1.$$

Substitute the above equation in the child's earnings equation to obtain,

$$(14) \qquad y_1 = (1+r)\beta y_0 + \frac{\beta(1+r)-\delta}{r}a_1 + u_1.$$

The term a_1 in the above equation can be eliminated solving equation (13) for a_1 and plugging it in equation (14),

$$y_1 = -\frac{\beta(1+r)r}{\beta(1+r)-\delta-r}y_0 + \frac{(1+r)(\beta(1+r)-\delta)}{\beta(1+r)-\delta-r}I_1 + u_1. \quad (15)$$

Note that $\frac{\partial y_1}{\partial I_1} > 0$ and, $\frac{\partial y_1}{\partial y_0} > 0$ for $r < \beta(1+r) < \delta$.

A relationship between the child's and the grandparent's earnings can be obtained. First, substitute $a_1 = \delta a_0 + w_1$ in equation (14),

$$y_1 = (1+r)\beta y_0 + \frac{\beta(1+r)-\delta}{r}\delta a_0 + \frac{\beta(1+r)-\delta}{r}w_1 + u_1.$$

Then, solve the above equation for one period ahead,

$$y_2 = (1+r)\beta y_1 + \frac{\beta(1+r)-\delta}{r}\delta a_1 + \frac{\beta(1+r)-\delta}{r}w_2 + u_2.$$

Finally, solve equation (14) for a_1 and substitute in the above equation to obtain,

$$y_2 = [\beta(1+r)+\delta]y_1 - \beta(1+r)\delta y_0 - \delta u_1 + \frac{\beta(1+r)-\delta}{r}w_2 + u_2. \quad (16)$$

Note that $\frac{dy_2}{dy_1} > 0$ and, $\frac{dy_2}{dy_0} < 0$, that is, the grandchild's earnings are positively related to the parent's earnings but, negatively related to the grandparent's earnings.

Table 1 summarizes the assumptions and predictions of the models presented. In section 4, I estimate the four equations presented in Table 1.

3 Data

The *Pesquisa Nacional por Amostra de Domicilios* (PNAD) is a nationally representative household survey conducted almost every year in Brazil. Survey participants are asked about the age, education, occupation and earnings of all members in the household. Also, they are asked to specify the relationship of each member to the reference person in the household, allowing identification of the father, sons and brothers for most households. Table 2 and 3 presents some descriptive statistics for males in the 1996 PNAD.

In Table 2, note the low educational achievement (5.9 years on average) and the early age at which Brazilians start working (age 12.9 on average). In Table 3, note that the percentage of males living with parents is high even among older males. For instance, at age 25, 50.9% of the males are still living with their parents. The percentage of males living with parents falls for older ages and the percentage of married males raises with age, and the sum of those percentages yields a number close to one for all ages. Among the males living with their parents only a small percentage is married. The numbers suggest that most males live with their parents until they marry.

The 1996 PNAD includes a mobility supplement with information on both the parents of the reference person and his/her spouse. It provides information on the parents' education and the father's occupation when the son was 15 years old. That information allows estimation of equations (i) and (ii) in Table 1.

Estimation of equations (iii) and (iv) requires a data set that covers at least three generations within a family. I take advantage of two specific characteristics of Brazil to build a data set with three generations. First, typically, sons live with their parents until they marry. Second, individuals quit school and start working at an early age. The first characteristic implies that it is likely that there are many households with adult sons in the basic sample of PNAD. The second characteristic suggests that those adult sons are not at the very beginning of their working careers.

Therefore, it is possible to build a nearly representative data set spanning three generations for Brazil by restricting the sample to households with adult sons. For those households, information is available on the reference person of the household (father), the son of the reference person (son) and the father of the reference person (grandfather)[6].

I built the three-generations data by focusing on sons aged 16 to 27 years[7]. The reason for the age restriction is to exclude the extremely unusual sons. As shown in Table 3, it is somewhat common and normal for males in the selected age interval to be working, living with parents and single[8].

---

[6]For the households in which a woman is the reference person, her spouse, if any, is considered the father and, his father the grandfather. Only about 2% of the households have a married woman as the reference person.

[7]I discuss the results for different age intervals in section 4.

[8]Participants of the 1996 survey are not explicitly asked about their marital status. But, they are asked to specify the relationship of each member with the head of the family. Married couples are identified if one member in the family is the spouse of the head of the family. Everyone else is assumed to be single. The accuracy of this procedure is checked for the 1995 PNAD that explicitly ask about each member marital status. The procedure

Sample selection problems may arise when dealing with a subsample consisting of households with adult sons. I use Heckman's estimation procedure to deal with this problem. I discuss the selection equation later in this section.

Note that the mobility supplement in PNAD does not provide the earnings of the father of the reference person (or spouse) in the household. Instead, it provides information on the fathers' occupation and education when the son was 15 years old.

To estimate a father's earnings, first, I calculate the year the son was 15-year-old and use the PNAD of that year (or the closest available year) to estimate the earnings equation of males. I assume earnings depend on education, 46 occupation category, experience (up to a quartic term), interaction dummies between 8 more broad occupation categories and experience, race and, state of residency. Second, I apply the estimated equation to the corresponding set of characteristics of the father to estimate the father's earnings.

Ferreira and Veloso (2006) and Dunn (2007) applied similar procedure. In Ferreira and Veloso (2006), the earnings equation depend on education, 6 occupational categories, dummies for cohort and interactions for cohort-occupation and cohort-education. In Dunn (2007), earnings depends on education and age (up to a quadratic term).

The mobility supplement does not inform the race or state of residence of the father. I estimate the father's expected earnings conditional on son's race. For instance, consider the father of a white male. I calculate the percentage of fathers in each race group conditional on the son's race being white. Then, I use those percentages as weights to calculate the weighted average earnings. Because the incidence of interracial marriage may differ over time, I calculate those percentages for all the years that all the necessary information are available, the appropriate year is chosen according to the son's age.

The migration segment informs the number of year the individual have being living in the current state (up to ten years), which allows identification

identified 64,427 of the 64,766 married males.

of the state of residence at age 15 for some males. For the remaining males, I estimate the father's expected earnings conditional on son's state of birth. For instance, suppose the son was born in Rio de Janeiro. I calculate the percentage of males in the age interval 13-17 living in each Brazilian state conditional on being born in Rio de Janeiro. Then, I use those percentages as weights to calculate the weighted average earnings. Because migration flows may differ over time, I calculate those percentages for all the years that all the necessary information are available, the appropriate year is chosen according to the son's age[9].

Note that the father of the reference person (or spouse) corresponds to the grandfather in the data set for three generations.

Selection equation in the three-generations data set

I assume that the probability that a male lives in his parents' house depends on age, marital status, whether he is enrolled in school, and his father's earnings.

$P(\text{son live with parents}) = f(\text{age, marrital status, school enrollment, father's ernings})$

As shown in Table 3, older males are less likely to live with their parents and few males stay in their parents' house after getting married. I expect that sons enrolled in school are more likely to live with their parents and, I expect that better paid fathers can afford having their children at home for a longer time.

Note that the fathers' earnings in September of 1996 are available only for males living with their parents. For both groups, males living with parents and males that do not, the set of variables used to impute the fathers' earnings are available. The difference is that for a male living with his parents, I have data on the father's occupation and education in 1996 instead of the father's occupation and education when the son was 15 years old. Assuming that the father's occupation and education does not change while the son is aged between 15 and 27, I can impute the fathers' earnings

[9] I describe in details the procedure to impute the father's earnings in the paper "Intergenerational mobility in earnings in Brazil with data on three generations".

for both groups repeating the same procedure applied before.

The top panel of Figure 1 shows the percentage of males living with their parents by age and marital status: the bulk of the married sons do not live with their parents, regardless of their age. The bottom panel of Figure 1 shows there is a higher percentage of sons enrolled in school among the sons living with their parents than among the ones that do not. Figure 2 shows that the fathers of sons living with parents are better paid than the fathers of sons not living with parents, for all ages considered[10].

One may argue that the males that succeed in the job market are more likely to marry and leave their parents' house, while the males that stay with their parents have lower earnings and worse marriage perspectives. Figure 3 shows that conditional on age, the average earnings of sons living with parents is not lower than the average earnings of sons who have moved out. The same is true for the average earnings of single and married males. Conditional on age, the average earnings of married males is not higher than the average earnings of single males[11].

4 Estimation Results

Estimation of equations (i) and (ii) requires data on at least two generations. Estimation of equations (iii) and (iv) requires data on at least three generations.

The PNAD combined with the 1996 Mobility Supplement provides information on two generations: all male heads of household (or spouse) and his father. The subsample consisting of households with adult sons contains

[10]The general shape of Figure 1 and 2 does not change for the sample used in the regressions.

[11]Figure A1 in the appendix presents the equivalent figure for the sample used in the regressions. There is a difference in average years of education. For sons aged 16 to 27, average years of education is 6.41 if they live with their parents; it is 5.94 if they live by themselves. Considering only the sons who are not enrolled in school, the averages are basically the same, 5.62 and 5.59. The averages are also similar when considering only the sons who are enrolled in school, 7.57 and 7.46. Therefore, only enrollment in school is kept in the selection equation. Among the sons not living with parents the percentage of blacks or mixed race is higher than the percentage of whites but the coefficient in the selection equation is not significant when controlling for father's earnings.

information on a third generation[12].

I apply OLS when dealing with the sample of all heads of household (or spouse) and Heckman's estimation procedure when dealing with the subsample of households with adult sons.

I drop sons working fewer than 15 hours a week and those with no paid jobs[13].

Estimation of Equation (i): $y_{t+1} = \gamma y_t + \epsilon_{t+1}$

Sample of males head of household (or spouse) and their fathers

Consider first the sample of males head of household (or spouse) and their fathers. I restrict the sample to sons aged 16 to 64. The proxy for the son's lifetime earnings is his earnings in September of 1996. The fathers' earnings are not available. I impute their earnings at age 47 as described in the previous section.

A control for sons' age is necessary to avoid bias of the estimate of the intergenerational elasticity. To illustrate the point, consider a society in which every son has exactly the same lifetime earnings as his father. Because of the age difference among sons, at any given year they are in different points of their career paths. Estimation the intergenerational elasticity using the fathers' earnings at age 47 and the sons' earnings in 1996 without any control for sons' age would yield an estimate lower than one. To control for sons' age, the specification includes an indicator variables for every five-year age intervals.

Estimation results of equation (i), presented in Table 4, column (a), suggest a highly immobile society. The intergenerational elasticity in earnings

[12] Others papers have used a three-generations data set to study intergenerational mobility. They estimate mobility tables or the probability that an individual belongs to a certain socioeconomic stratum as a function of parents and grandparents' characteristics. Recent papers include Erola and Moisio (2007), Warren and Hauser (1997) and Biblarz et al. (1996). Peters (1992) uses data from the National Longitudinal Surveys to estimate the grandson's earnings as a function of parents' characteristics and grandfathers' education and finds no significant effect for the grandfathers' education.

[13] The percentage of males working part time (15-39 hours) in the three generations sample used in the regressions is not high among males living with parent (13.7%) or for males head of the household (7.9%).

is 0.84, meaning that a difference in earnings among fathers of 100% is perpetuated by a difference in earnings among sons of 84%[14]. The estimated intergenerational elasticity range from 0.4 to 0.5 for U.S..

Dunn (2007) and Ferreira and Veloso (2006) also use data from PNAD and find the intergenerational elasticity in earnings in Brazil of 0.85 and 0.58, respectively[15]. Behrman et al. (2001) and Ferreira and Veloso (2003) use data from PNAD and find that the correlation between parent's and child's education in Brazil is 0.7 and 0.79, respectively.

As shown in column (b) the intergenerational elasticity increases with age. There are two interpretations for the higher elasticity for older males. First, the intergenerational elasticity is decreasing over time. Second, the elasticity between father's and son's earnings increases over the son's life cycle. Mayer (2005) finds that for U.S. "the strength of the relationship between parental features and the wages of sons increases over the life cycle of the sons." He points out that the second interpretation is consistent with wage evolution models like Ben-Porath (1969) and Jovanovic (1979). Haider and Solon (2005) suggest that for U.S. the son's one-year-earnings is a better proxy for lifetime earnings when measured in the early thirties and mid forties because the attenuation bias associated to one-year-earnings is small in this age interval. According to Grawe (2006), the life cycle bias when estimating the intergenerational elasticity in lifetime earnings is likely to be reduced by observing both fathers and sons near midlife. Dunn (2007) estimate the intergenerational elasticity using sons' earnings at age 40 and find that the elasticity is lower for younger cohorts.

[14]Note that equation (i) does not have a constant term. A constant was included in the estimations to account for the fact that average earnings of fathers and sons are not necessarily the same. For instance, the constant is positive in a society that experienced per capita economic growth.

[15]In the paper "Intergenerational mobility in earnings in Brazil with data on three generations", I show that most of the difference between the estimate in this paper and in Ferreira and Veloso (2006) is explained by three differences in the estimation criteria. Fist, Ferreira and Veloso exclude the rural areas and the sons working between 15 and 39 hours a week. Second, they include indicator variables for race and region in the regressions. Third, they estimate the fathers' earnings ignoring the difference in earnings across race and state.

The remaining columns of Table 4 show that intergenerational elasticity is higher among the sons of males with earnings below the median than above the median, it is higher among the sons currently living in rural than urban residencies and, it is higher among the sons currently living in the southeast than the northeast–a poor region in Brazil. The elasticity is not significantly different for for blacks or individuals of mixed race than for other race groups[16].

How bad is the selection problem?

As mentioned earlier, I can impute fathers' earnings for both sons living with parents and sons that do not. The difference is that for sons living with parents I have information on the father's occupation and education in 1996 rather than at the time the son was 15 years old.

Using the imputed father's earnings, I can estimate the intergenerational elasticity in earnings among young sons (living with parents or not) and compare the result with the elasticity obtained using the subsample of young sons living with parents. The results are presented in Table 5.

Among the sons aged 16 to 21, the intergenerational elasticity in earnings is 0.389 for the full sample[17]. It is about the same for the sample of sons living with their parents: 0.393 (OLS), 0.398 (Heckman two step), and 0.398 (Heckman maximum likelihood).

Among the sons aged 22 to 27, the intergenerational elasticity is higher for the full sample (0.22) than for the subsample of sons living with parents (about 0.19). I obtain similar results applying OLS (0.185), Heckman two step (0.189) or Heckman maximum likelihood (0.188).

The Heckman estimation is justified by the high chi-squared value for a test of the Heckman model versus a model with no selection problem. The p-value is 0.0016.

As expected, married and older sons are less likely to live with their parents. Sons of better paid fathers and sons enrolled in school are more likely to live with their parents.

[16]The paper "Intergenerational mobility in earnings in Brazil with data on three generations" provides a detailed discussion about the estimation results for equation (i).

[17]Note that I am dealing with a young cohort and, as previously discussed, the intergenerational elasticity is lower for younger cohorts.

Estimation of Equation (ii): $y_{t+1} = \eta_1 y_t + \eta_2 I_{t+1} + \epsilon_{t+1}$

I use years of formal education as a proxy for investment in human capital[18].

Consider first the sample of males head of household (or spouse) and their fathers. Controlling for father's earnings, the impact of the son's education on his own earnings, η_2 in equation (ii), is positive (see Table 6), contrary to the negative sign predicted by the Becker and Tomes model.

The sign is also positive across cohorts, race groups, regions and rural versus urban areas. It is also positive for the sample of sons living in their parents' house.

I obtain equation (ii) in the (partially) modified model by assuming there is no skipping generations effect ($\theta = 0$). Under this assumption, the sign of the estimated coefficient agrees with the predictions of the model. But the assumption that $\theta = 0$ is contradicted in the estimation of equation (iii).

Estimation of Equation (iii): $y_{t+2} = \tau_1 y_{t+1} + \tau_2 y_t + \epsilon_{t+1}$

The rest of the estimations require a three-generations data set.

To account for the fact that two sons may be at different points in their careers in 1996 just because one is older than the other, I use an indicator variable for the older cohort between 16 and 21 years old.

Two fathers may have different earnings in 1996 just because one is older than the other. If all fathers were the same age when their sons were born, an indicator variable for the sons' age would control for the fact that not all father-son pairs are at the same stage of their careers in 1996. Although the father's and son's age are related, this is not a one to one relation. The coefficient in a regression of the sons' age on the father's age is 0.16 and the R^2 is 0.15. To correct for the fact that not all fathers are at the same stage of their careers in 1996 an indicator variable for son's cohort may not perfectly characterize the father's cohort, I use the father earnings minus

[18] I am assuming that parents who care about quality of education also care about years of education. So more years of formal education implies more investment in human capital because parents invest for a longer period of time, and because parents invested in more quality.

the average earnings for his age group. The age groups for the fathers are aged 45 and younger, aged 46 to 55, and older than 55.

Controlling for father's earnings, the estimated impact of the grandfather's earnings on the grandson's earnings is positive. It is positive for all three estimations: OLS, Heckman Two Step and Heckman Maximum Likelihood (see Table 7). It is positive for different cohorts, race groups, regions and rural versus urban areas.

The results suggest that grandfathers with high earnings have a positive impact on their grandsons' earnings through a channel other than the fathers. In other words, the grandfathers with high earnings have a direct positive impact on their grandsons' earnings. One possibility, examined in the modified model, is that the grandfather's endowment directly impacts the grandson' endowment or $\theta > 0$.

Assuming that there is no variation in market luck, the Becker and Tomes model predicts a negative coefficient for the grandfather's earnings. The positive sign can be evidence against the assumption of no variation in market luck.

Assuming that the grandson's endowment does not directly depend on the grandfather's endowment ($\theta = 0$), the (partially) modified model predicts a negative coefficient for the grandfather's earnings. I interpret the positive estimated coefficient as evidence in favor of a less restrictive assumption or $\theta > 0$.

Estimation of Equation (iv): $y_{t+2} = \pi_1 y_{t+1} + \pi_2 y_t + \pi_3 I_{t+2} + \pi_4 I_{t+1} + \epsilon_{t+1}$

The signs of the estimated coefficients agree with the predictions of the (fully) modified model. They are all positive except for the coefficient on father's education. See Table 8[19].

I examine whether the coefficients in equation (iv) differ across age or race groups, among the sons of (relatively) poor and wealthy fathers, among

[19]The plot of the grandfathers' earnings versus grandsons' earnings suggests a concave relationship between the two variables. Table A3 in the Appendix of Tables presents the results when including a quadratic term for grandfathers' earnings in the equation.

sons currently living in (relatively) poor versus wealthy regions or in rural versus urban areas.

At a 10% significance level, I cannot reject the hypothesis that the coefficients in equation (iv) are the same for males living in poor versus relatively wealthy regions[20].

At a 10% significance level, I reject the hypothesis that the coefficients in equation (iv) are the same for: (1) old and young cohorts, (2) Blacks and non Blacks, (3) the sons of poor and wealthy fathers and (4) for males living in urban versus rural areas.

For the same significance level, I cannot reject the hypothesis that the coefficients for father's and grandfather's education are the same for (1) Blacks and non Blacks, (2) the sons of poor and wealthy fathers and (3) males living in urban versus rural areas. Also, I cannot reject the hypothesis that the coefficients for son's and father's earnings are the same for (1) Blacks and non Blacks and (2) old and young cohorts.

I reject the hypothesis that the coefficients for son's and father's education are the same across age groups. Also, I reject the hypothesis that the coefficients for father's and grandfather's earnings are the same for (1) sons living in urban and rural areas and (2) among the sons of poor and wealthy fathers. Columns (d), (e) and (f) in Table 8 present the estimation results.

Replacing paternal grandfather's earnings by an average between maternal and paternal grandfather's earnings I obtain the results in column (g) in Table 8.

As explained in section 2, the variation in earnings explained by variations in y_{t+1}, y_t and I_{t+1} in equation (iv) represent a lowerbound for the variation in earnings explained by variation in endowment across families. Besides, the variation in earnings explained by variation in I_{t+2} represents an upperbound for the variation in earnings explained by variation in human capital.

Among the sons living with their parents, about 34.9% of the variation in earnings is explained by differences in family background. If it was possible to eliminate all variation in human capital within cohorts, the variation in

[20] I consider Northeast and North as poor region and the remaining regions as relatively wealthy regions.

the *predicted* earnings would fall by no more than 60.5%. This implies that if every son had the same human capital, the variation in earnings would drop in no more than 21.1%. If there was no variation in endowment across families within cohorts, the variation in the *predicted* earnings would fall by no less than 74.5%, implying that the variation in sons' earnings would be at least 26% lower.

Using the estimates in column (d), Table 8, I calculate the percentage of the variation in earnings explained by differences in family background for young and old cohorts. I repeat the same procedure for sons living in rural and urban areas using the results in column (f). The percentages are presented in Table 9.

Family background explains a higher share of the variation in earnings for older males and males living in rural areas. The results presented in Table 9 are consistent with the results for equation (i). Family background explains 44% of the variation in earnings among males aged 22 to 27 living with parents and 28.7% among males aged 16 to 21. It explains 35.6% of the variation in earnings among the males living in rural areas and 29.6% of the variation in earnings among males living in urban areas.

Replacing paternal grandfather's earnings by an average between maternal and paternal grandfather's earnings, I find that family background explains 40.4% of the of the variation in earnings among males aged 16 to 27 living with parents. Eliminating all the differences in endowment the variation in earnings would fall by no less than 30.9%. Eliminating variations in investment in human capital the variation in earnings among them would fall by no more than 24.0%

Different Age Groups

Table 10 presents the Heckman estimation of equation (iv) for two different age intervals: a broader interval from 12 to 40 years old, and a narrower interval from 18 to 25. For ease of comparison, the results for males aged 16 to 27 appear in the table as well.

The estimated coefficients are stable for the three age groups, except for the coefficient of the father's education which is no longer significant for the narrower age group (ages 18 to 25).

Among males aged 12 to 40 living with their parents, family background explains about 35.3% of the variation in earnings. The percentage is 34.9% for the males aged 16 to 27 and 33.4% for those aged 18 to 25. If it was possible to eliminate all variation in human capital within cohorts for every male aged 12 to 40, the variation in earnings would fall by no more than 21.7%. The percentage is 21.1% for those aged 16 to 27 and 20.1% for those aged 18 to 25. If it was possible to eliminate variation in endowments across families, the variation in earnings would fall by no less than 26.0% for those aged 12 to 40, and those aged 16 to 27, and 25.0% for those aged 18 to 25.

5 Conclusion

I estimate several relationships established by the Becker and Tomes (1979) model presented by Gary Solon (1999), and the modified version of the model proposed in this article. The relationships are listed below.

(i) $y_{t+1} = \gamma y_t + \epsilon_{t+1}$

(ii) $y_{t+1} = \eta_1 y_t + \eta_2 I_{t+1} + \epsilon_{t+1}$

(iii) $y_{t+2} = \tau_1 y_{t+1} + \tau_2 y_t + \epsilon_{t+1}$

(iv) $y_{t+2} = \pi_1 y_{t+1} + \pi_2 y_t + \pi_3 I_{t+2} + \pi_4 I_{t+1} + \epsilon_{t+1}$

where y represents earnings and I represents investment in human capital. The subscript t represents generation t, $t+1$, their sons and $t+2$, their grandsons.

Equation (i), commonly estimated in intergenerational mobility studies, can be obtained from the Becker and Tomes model (original and modified) adding the assumption of identical children's endowments. This is probably a strong assumption for an intergenerational mobility model since it is reasonable to think that the children of wealthy parents are in average wealthy not only because their parents could afford a high investment in their human capital but also because those children inherit characteristics from their parents that explain why their parents are wealthy in the first place. Those characteristics can be talent, good health, IQ, physical appearance, attitudes toward work, leisure activities that enhance their productivity at work, family connections and family culture in general.

The estimate of the intergenerational elasticity in earnings in Brazil is about 0.84, suggesting a highly immobile society. The elasticity in earnings is higher if the father earnings are below the the median than above median, among the sons currently living in the northeast than in the southeast and, among the sons currently living in rural than urban residencies.

Estimation results of equation (ii) contradict the predictions of the Becker and Tomes model presented by Solon (1999). Controlling for parent's earnings, the child's earnings are positively related to investment in the child's human capital contrary to the negative relationship predicted by the model.

Estimation results of equation (iii) contradict the prediction of the Becker and Tomes model and the partially modified model. Controlling for parent's earnings, the child's earnings are positively related to grandparent's earnings contrary to the negative relationship predicted by Becker and Tomes model and the partially modified model. The positive sign can be interpreted as evidence against the assumption of no variation in market luck across individuals in the Becker and Tomes model, and as evidence against the assumption of no skipping generation effect in the partially modified model.

Estimation results of equation (iv) are consistent with the predictions of the (fully) modified model and indicate that family background explains 34.9% of the variation in earnings among males between 16 and 27 years old living with parents. Eliminating all the differences in endowment the variation in earnings would fall by no less than 26%. If every male in Brazil living with his parents had the same investment in human capital the variation in earnings among them would fall by no more than 21.1%.

In this article the 1996's earnings of the father and sons are used as proxies for lifetime earnings although the transitory component of one year earnings may be quite large at young ages. Some males in the sample are still enrolled in school and their earnings at this stage may not be a good proxy for their lifetime earnings[21]. Another limitation is that differences in the quality of education are ignored assuming that parents who care about years of education also care about quality of education.

[21] See Haider and Solon (2005) for a discussion about the attenuation bias caused by using one-year-earnings as a proxy for lifetime earnings.

REFERENCES

Becker, Gary S., and Tomes, Nigel. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87, 1979.

Becker, Gary S., and Tomes, Nigel. Human capital and the rise and fall of families. *Human capital: a theoretical and empirical analysis with special reference to education*, 3th edition, 1993.

Behrman, Jere R., Gaviria, Alejandro, Szekely, Miguel. Intergenerational mobility in Latin America. Inter-American Development Bank, *working paper 452*, 2001.

Ben-Porath, Yoram. The production of human capital and the life cycle of earnings. *The Journal of Political Economy*, Vol.75, August, 1967.

Biblarz, Timothy J., Bengtson, Vern L., and Bucur, Alexander. Social mobility across three generations. *Journal of Marriage and the Family*, Vol.58, No.1, February, 1996.

Dunn, Christopher E.. The intergenerational transmission of lifetime earnings: evidence from Brazil. *The B.E. Journal of Economic Analysis & Policy*, Vol.7, Issue 2, 2007.

Corak, Miles. Do poor children become poor adults? Lesson from a cross country comparison of general earnings mobility. IZA *Discussion Paper* No.1993, March, 2006.

Erola, Jani, and Moisio, Pasi. Social mobility over three generations in Finland. *European Sociological Review*, Vol.23, No.2, 2007.

Ferreira, Sergio Guimares, Veloso, Fernando A.. Mobilidade intergeneracional de educacao no Brasil. *Pesquisa e Planejamento Econmico*, Rio de

Janeiro, Vol.33, No.3, December 2003.

Ferreira, Sergio Guimares, and Veloso, Fernando A.. Intergenerational mobility of wages in Brazil. *Brazilian Review of Econometrics.* Vol.26, No.2 , 2006.

Grawe, Nathan D.. Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics*, Vol.13, October, 2006.

Heckman, James J.. Sample selection bias as a specification error. *Econometrica*, Vol.47, No.1, January, 1979.

Haider, Steven, Solon Gary. Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, Vol.96, No.4, 2006.

Jovanovic, Boyan. Job Matching and the Theory of Turnover. *The Journal of Political Economy*, Vol.87, October, 1979.

Mayer, Adalbert. Intergenerational relationships over the lifecycle and the determination of wages. January, 2005.

Marchon, Cassia. Intergenerational mobility in earnings in Brazil with data on three generations. Working paper available at http://econweb.tamu.edu/cassiahm/papers.htm

Peters, H. Elizabeth. Patterns of intergenerational mobility in income and earnings. *The Review of Economics and Statistics*, Vol.74, August, 1992.

Solon, Gary. Intergenerational mobility in the labor market. *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card, 1999.

Solon, Gary. Cross-Country Differences in Intergenerational Earnings Mobility. *The Journal of Economic Perspectives*, Vol.16, No.3, Summer, 2002.

Warren, John Robert, and Hauser, Robert M.. Social stratification across three generations: new evidence from the Wisconsin Longitudinal Study. *American Sociological Review*, Vol.62, August, 1997.

TABLES

Table 1: Summary of Assumptions and Predictions of the Becker and Tomes and the Modified Model

Equation	Equation number in text	Becker and Tomes	Modified Model
(i) $y_{t+1} = \gamma y_t + \epsilon_{t+1}$	(7)&(12)	Assumption Added to the Model: $\delta = 0$ or $Var(a_j) = 0$	Assumption Added to the Model: $Var(a_j) = 0$
(ii) $y_{t+1} = \eta_1 y_t + \eta_2 I_{t+1} + \epsilon_{t+1}$	(5)&(15)	Predictions: $\frac{\partial y_{t+1}}{\partial I_{t+1}} < 0$ and $\frac{\partial y_{t+1}}{\partial y_t} = -\frac{\partial y_{t+1}}{\partial I_{t+1}}$	Assumption Added to the Model: $\theta = 0$. Predictions: $\frac{\partial y_{t+1}}{\partial I_{t+1}} > 0$ and $\frac{\partial y_{t+1}}{\partial y_t} > 0$
(iii) $y_{t+2} = \tau_1 y_{t+1} + \tau_2 y_t + \epsilon_{t+1}$	(6)&(16)	Assumption Added to the Model: $Var(u_{t+1}) = 0$. Prediction: $\frac{\partial y_{t+2}}{\partial y_t} < 0$	Assumption Added to the Model: $\theta = 0$. Prediction: $\frac{\partial y_{t+2}}{\partial y_t} < 0$
(iv) $y_{t+2} = \pi_1 y_{t+1} + \pi_2 y_t + \pi_3 I_{t+2} + \pi_4 I_{t+1} + \epsilon_{t+1}$	(11)	The equation is not predicted in the model.	for $r < \beta(1+r) < \delta$. Predictions: $\frac{\partial y_2}{\partial I_2} > 0$, $\frac{\partial y_2}{\partial y_1} > 0$, $\frac{\partial y_2}{\partial y_1} > 0$ and, $\frac{\partial y_2}{\partial I_1} < 0$

Table 2: Selected Characteristics of Males Aged 16 to 64

Number of Observations	96,788
	Average
Monthly Earnings (reais)	580.40
	(952.4)
Years of Education	5.9
	(4.3)
Age Started Working	12.9
	(3.8)
	Percentage
Black/Mixed Race	43.9
Living in Rural Areas	20.1
Region of Residence:	
Southeast	45.8
Northeast	27.0
South	15.8
Midwest	7.0
North	4.4
Aged 16 to 24	29.2
Aged 25 to 34	25.3
Aged 35 to 44	21.4
Aged 45 to 54	14.7
Aged 55 to 64	9.5

*S*ource: 1996 PNAD.

Note: Standard deviations in parenthesis.

Table 3: SELECTED CHARACTERISTICS OF MALES IN THE 1996 PNAD , BY AGE

		Percentages				
Age	Number of Observations	Living with Parents	Married	Working	Living with Parents and Married	Living with Parents and Working
12	3,305	99.97	0.03	4.36	0.00	4.36
13	3,496	99.94	0.06	7.18	0.00	7.15
14	3,519	99.94	0.03	12.45	0.03	12.39
15	3,492	99.71	0.14	21.05	0.06	20.85
16	3,482	99.25	0.78	31.68	0.26	31.05
17	3,226	98.17	1.77	40.67	0.65	39.03
18	3,048	95.44	3.97	47.34	1.08	43.47
19	2,883	91.47	9.02	58.20	2.43	50.78
20	2,706	85.37	14.34	64.63	2.92	51.66
21	2,495	81.48	18.48	66.77	3.37	50.06
22	2,530	71.07	28.74	74.55	4.11	47.94
23	2,404	65.43	33.69	77.04	3.74	45.26
24	2,396	56.30	41.90	79.34	3.09	39.44
25	2,244	50.89	47.46	82.26	3.74	36.50
26	2,388	45.48	53.43	82.79	3.64	32.29
27	2,308	36.83	60.10	84.10	2.38	25.52
28	2,212	30.65	67.04	87.66	2.22	22.65
29	2,164	27.45	68.76	86.92	2.22	19.45
30	2,513	24.11	71.55	86.87	1.59	16.39
31	2,265	19.91	75.89	88.52	1.59	14.04
32	2,350	19.40	77.19	88.21	1.66	13.28
33	2,358	16.96	78.29	89.31	1.06	11.87
34	2,214	14.91	81.21	90.24	0.99	10.75
35	2,152	13.75	81.83	89.41	1.12	9.53
36	2,222	12.83	82.49	90.37	0.72	8.51
37	2,115	11.11	83.88	90.54	0.57	7.85
38	2,111	9.95	84.37	89.82	0.47	6.40

*S*ource: 1996 PNAD.

Table 4: OLS ESTIMATION OF EQUATION (i): $y_{t+1} = \gamma y_t + \epsilon_{t+1}$

	(a)	(b)	(c)
Dependent Variable: Son's Earnings			
Father's Earnings	0.843***	0.596***	1.195***
	(0.01)	(0.03)	(0.03)
Father's Earnings*Son's Age 25-34		0.150***	
		(0.03)	
Father's Earnings*Son's Age 35-44		0.289***	
		(0.03)	
Father's Earnings*Son's Age 45-54		0.371***	
		(0.03)	
Father's Earnings*Son's Age 55-64		0.381***	
		(0.04)	
Father's Earnings*Father's Earn. Above the Median			–0.559***
			(0.03)
Constant	0.795***	2.117***	–1.040***
	(0.07)	(0.15)	(0.16)
Fathers Earnings Above the Median			3.239***
			(0.17)
Adjusted R^2	0.297	0.302	0.310
N. of Observations	36,705	36,705	36,705

Source: PNAD.

Note: Standard errors in parentheses. Omitted age: 16-24 years old. Coefficients of indicator variables for age are not presented.

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4 (continued): OLS ESTIMATION OF EQUATION (i): $y_{t+1} = \gamma y_t + \epsilon_{t+1}$

	(d)	(e)	(f)
Dependent Variable: Son's Earnings			
Father's Earnings	0.755***	0.740***	0.724***
	(0.01)	(0.01)	(0.01)
Father's Earnings*Black/Mixed	0.021		
	(0.02)		
Father's Earnings*North		–0.143***	
		(0.03)	
Father's Earnings*Northeast		0.194***	
		(0.02)	
Father's Earnings*South		0.004	
		(0.02)	
Father's Earnings*Midwest		–0.078***	
		(0.02)	
Father's Earnings*Rural			0.171***
			(0.02)
Constant	1.444***	1.527***	1.606***
	(0.07)	(0.09)	(0.07)
Black/Mixed	–0.427***		
	(0.09)		
North		0.711***	
		(0.18)	
Northeast		–1.423***	
		(0.10)	
South		–0.083	
		(0.12)	
Midwest		0.369***	
		(0.14)	
Rural			–1.540***
			(0.12)
Adjusted R^2	0.316	0.319	0.352
N. of Observations	36,705	36,705	36,705

Source: PNAD.

Note: Standard errors in parentheses. Omitted region: Southeast. Coefficients of indicator variables for age are not presented.

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 5: ESTIMATION OF EQUATION (i): $y_{t+1} = \gamma y_t + \epsilon_{t+1}$ FOR FATHERS AND GRANDSONS

(a) Sons living with parents - OLS
(b) Sons living with parents - Heckman Two Step Estimation
(c) Sons living with parents - Heckman Maximum Likelihood Estimation
(d) All Sons (living or not with parents) - OLS

	(a)	(b)	(c)	(d)
Dependent Variable: Grandson's Earnings				
Father's Earnings (Imputed)	0.393***	0.398***	0.398***	0.389***
	(0.018)	(0.016)	(0.018)	(0.016)
Father's Earnings (Imputed)*Older Coh.	0.185***	0.189***	0.188***	0.220***
	(0.028)	(0.025)	(0.028)	(0.021)
Constant	2.912***	2.869***	2.872***	3.007***
	(0.099)	(0.091)	(0.100)	(0.090)
Older Cohort	–0.788***	–0.843***	–0.840***	–0.877***
	(0.162)	(0.146)	(0.161)	(0.119)
Selection Equation				
Married		–2.570***	–2.558***	
		(0.042)	(0.042)	
Age		–0.157***	–0.157***	
		(0.007)	(0.007)	
Enrolled in School		0.144**	0.152**	
		(0.054)	(0.055)	
Father's Earnings (Imputed)		0.190***	0.186***	
		(0.027)	(0.028)	
Constant		3.474***	3.484***	
		(0.184)	(0.186)	
R^2	0.286			0.309
LR Test of Indep. Equations			12.657	
N. of Observations Censored		5,125	5,125	
N. of Observations	5,125	10,176	10,176	10,176

Source: 1996 PNAD.

Note: Standard errors in parentheses. Cluster: family (except for the two step regression).

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 6: OLS Estimation of Equation (ii): $y_{t+1} = \eta_1 y_t + \eta_2 I_{t+1} + \epsilon_{t+1}$

	(a)	(b)
Dependent Variable: Son's Earnings		
Father's Earnings	0.334***	0.375***
	(0.008)	(0.028)
Son's Years of Education	0.120***	0.073***
	(0.001)	(0.005)
Father's Earnings*Son's Age 25-34		–0.048
		(0.031)
Father's Earnings*Son's Age 35-44		–0.036
		(0.031)
Father's Earnings*Son's Age 45-54		–0.037
		(0.032)
Father's Earnings*Son's Age 55-64		0.004
		(0.038)
Son's Education*Son's Age 25-34		0.036***
		(0.006)
Son's Education*Son's Age 35-44		0.049***
		(0.006)
Son's Education*Son's Age 45-54		0.057***
		(0.006)
Son's Education*Son's Age 55-64		0.058***
		(0.006)
Constant	2.942***	2.949***
	(0.059)	(0.142)
Adjusted R^2	0.478	0.481
N. of Observations	36,705	36,705

Source: PNAD.

Note: Standard errors in parentheses. Omitted age: 17-24 years old. Coefficients of indicator variables for age are not presented. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7: Estimation of Equation (iii): $y_{t+2} = \tau_1 y_{t+1} + \tau_2 y_t + \epsilon_{t+1}$

(a) OLS				
(b) Heckman Two Step Estimation				
(c)-(d) Heckman Maximum Likelihood Estimation				
	(a)	(b)	(c)	(d)
Dependent Variable: Grandson's Earnings				
Grandfather's Earnings	0.195***	0.200***	0.200***	0.174***
	(0.024)	(0.018)	(0.024)	(0.027)
Father's Earnings	0.331***	0.333***	0.333***	0.304***
	(0.013)	(0.010)	(0.013)	(0.015)
Grand. Earnings*Older Coh.				0.062
				(0.046)
Father's Earnings*Older Coh.				0.074***
				(0.025)
Constant	4.084***	4.042***	4.045***	4.182***
	(0.133)	(0.102)	(0.134)	(0.149)
Older Cohort	0.368***	0.338***	0.339***	–0.008
	(0.020)	(0.020)	(0.021)	(0.257)
Selection Equation				
Married		–2.570***	–2.561***	–2.560***
		(0.042)	(0.035)	(0.035)
Age		–0.157***	–0.157***	–0.157***
		(0.007)	(0.004)	(0.004)
Enrolled in School		0.144***	0.155***	0.156***
		(0.054)	(0.030)	(0.030)
Father's Earnings (Imputed)		0.190***	0.174***	0.175***
		(0.027)	(0.018)	(0.018)
Constant		3.474***	3.544***	3.545***
		(0.184)	(0.111)	(0.111)
R^2	0.334			
LR Test of Indep. Equations			11.729	14.476
N. of Observations Uncensored		5,125	5,125	5,125
N. of Observations	5,125	10,176	10,176	10,176

Source: PNAD.

Note: Standard errors in parentheses. Cluster: family (except for the two step regression). Omitted age: 16-21 years old.

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 8: ESTIMATION OF EQUATION (iv): $y_{t+2} = \pi_1 y_{t+1} + \pi_2 y_t + \pi_3 I_{t+2} + \pi_4 I_{t+1} + \epsilon_{t+1}$

(a) OLS
(b) Heckman Two Step Estimation
(c)-(f) Heckman Maximum Likelihood Estimation

	(a)	(b)	(c)	(d)
Dependent Variable: Grandson's Earnings				
Grandfather's Earnings	0.146***	0.149***	0.149***	0.147***
	(0.023)	(0.019)	(0.024)	(0.023)
Father's Earnings	0.244***	0.244***	0.244***	0.246***
	(0.015)	(0.011)	(0.015)	(0.015)
Son's Years of Education	0.061***	0.061***	0.061***	0.050***
	(0.003)	(0.003)	(0.003)	(0.004)
Father's Years of Education	–0.008**	–0.007**	–0.007**	–0.010***
	(0.003)	(0.003)	(0.003)	(0.004)
Son's Education*Older Coh.				0.023***
				(0.006)
Father's Education*Older Coh.				0.009
				(0.006)
Constant	4.000***	3.956***	3.958***	4.052***
	(0.127)	(0.103)	(0.127)	(0.126)
Older Cohort	0.306***	0.268***	0.269***	0.065*
	(0.019)	(0.020)	(0.020)	(0.039)
Selection Equation				
Married		–2.570***	–2.555***	–2.552***
		(0.042)	(0.035)	(0.035)
Age		–0.157***	–0.156***	–0.157***
		(0.007)	(0.004)	(0.004)
Enrolled in School		0.144***	0.177***	0.180***
		(0.054)	(0.031)	(0.031)
Father's Earnings (Imputed)		0.190***	0.173***	0.176***
		(0.027)	(0.018)	(0.018)
Constant		3.474***	3.531***	3.528***
		(0.184)	(0.111)	(0.110)
R^2	0.386			
LR Test of Indep. Equations			19.195	23.424
N. of Observations Uncensored		5,125	5,125	5,125
N. of Observations	5,125	10,176	10,176	10,176

Source: PNAD.

Note: Standard errors in parentheses. Cluster: family (except for the two step regression). Omitted age: 16-21 years old.

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 8 (continued): ESTIMATION OF EQUATION (iv): $y_{t+2} = \pi_1 y_{t+1} + \pi_2 y_t + \pi_3 I_{t+2} + \pi_4 I_{t+1} + \epsilon_{t+1}$

	(e)	(f)	(g)
Dependent Variable: Grandson's Earnings			
Grandfather's Earnings	0.229***	0.132***	
	(0.037)	(0.024)	
Father's Earnings	0.343***	0.233***	0.258***
	(0.028)	(0.016)	(0.023)
Son's Years of Education	0.059***	0.058***	0.063***
	(0.003)	(0.003)	(0.005)
Father's Years of Education	–0.004	–0.006*	–0.010*
	(0.004)	(0.004)	(0.005)
Grandfather's Earn.*Above the Median Earn.	–0.112**		
	(0.046)		
Father's Earn.*Above the Median Earn.	–0.145***		
	(0.037)		
Grand. Earnings*Rural		0.189**	
		(0.087)	
Father's Earnings*Rural		0.025	
		(0.035)	
Average Between Maternal and			0.178***
Paternal Grandfather's Earnings			(0.043)
Constant	3.613***	4.089***	3.797***
	(0.200)	(0.132)	(0.233)
Older Cohort	0.276***	0.275***	0.184***
	(0.021)	(0.020)	(0.036)
Father's Earnings Above the Median	0.556**		
	(0.253)		
Rural		–1.094**	
		(0.474)	
Selection Equation			
Married	–2.555***	–2.555***	–2.446***
	(0.035)	(0.035)	(0.053)
Age	–0.156***	–0.156***	–0.184***
	(0.004)	(0.004)	(0.006)
Enrolled in School	0.177***	0.178***	0.193***
	(0.031)	(0.031)	(0.044)
Father's Earnings (Imputed)	0.174***	0.174***	0.193***
	(0.018)	(0.018)	(0.025)
Constant	3.527***	3.527***	3.486***
	(0.111)	(0.111)	(0.157)
LR Test of Indep. Equations	19.359	19.529	11.661
N. of Observations Uncensored	5,125	5,125	2,025
N. of Observations	10,176	10,176	7,076

Source: PNAD.

Note: Standard errors in parentheses. Cluster: family. Omitted age: 16-21 years old.
* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 9: Drop In The Variation In Earnings Among Males Aged 16 to 27 Living With Parents After Eliminating Differences In: (a) Family Background, (b) Human Capital, And (c) Endowment (Percentages)

	Family Background	Human Capital (upperbound)	Endowment (lowerbound)
Males Between 16 And 27 Years Old	34.9	21.1	26.0
Males Between 16 And 21 Years Old	28.7	14.9	23.0
Males Between 22 And 27 Years Old	44.0	29.8	30.7
Males Living In Rural Areas	35.6	18.5	27.7
Males Living In Urban Areas	29.6	17.4	22.1
Using Average Between Maternal Paternal Grandfather's Earnings - Males Between 16 And 27 Years Old	40.4	24.0	30.9

Source: PNAD.

Table 10: Heckman Maximum Likelihood Estimation of Equation (iv): $y_{t+2} = \pi_1 y_{t+1} + \pi_2 y_t + \pi_3 I_{t+2} + \pi_4 I_{t+1} + \epsilon_{t+1}$

(a) Sons living with their parents from 13 until 38 years old
(b) Sons living with their parents from 16 until 27 years old
(c) Sons living with their parents from 18 until 25 years old

	(a)	(b)	(c)
Dependent Variable: Grandson's Earnings			
Grandfather's Earnings	0.141***	0.149***	0.120***
	(0.022)	(0.024)	(0.026)
Father's Earnings	0.251***	0.244***	0.237***
	(0.014)	(0.015)	(0.017)
Son's Years of Education	0.063***	0.061***	0.058***
	(0.003)	(0.003)	(0.004)
Father's Years of Education	–0.007**	–0.007**	–0.005
	(0.003)	(0.003)	(0.004)
Constant	3.784***	3.958***	4.232***
	(0.117)	(0.127)	(0.143)
Selection Equation			
Married	–2.490***	–2.555***	–2.577***
	(0.030)	(0.035)	(0.039)
Age	–0.132***	–0.156***	–0.162***
	(0.003)	(0.004)	(0.007)
Enrolled in School	0.208***	0.177***	0.129***
	(0.027)	(0.031)	(0.036)
Father's Earnings (Imputed)	0.099***	0.173***	0.221***
	(0.016)	(0.018)	(0.021)
Constant	3.400***	3.531***	3.392***
	(0.095)	(0.111)	(0.152)
LR Test of Indep. Equations	9.665	19.195	1.813
N. of Observations Uncensored	6,111	5,125	3,696
N. of Observations	23,180	10,176	6,901

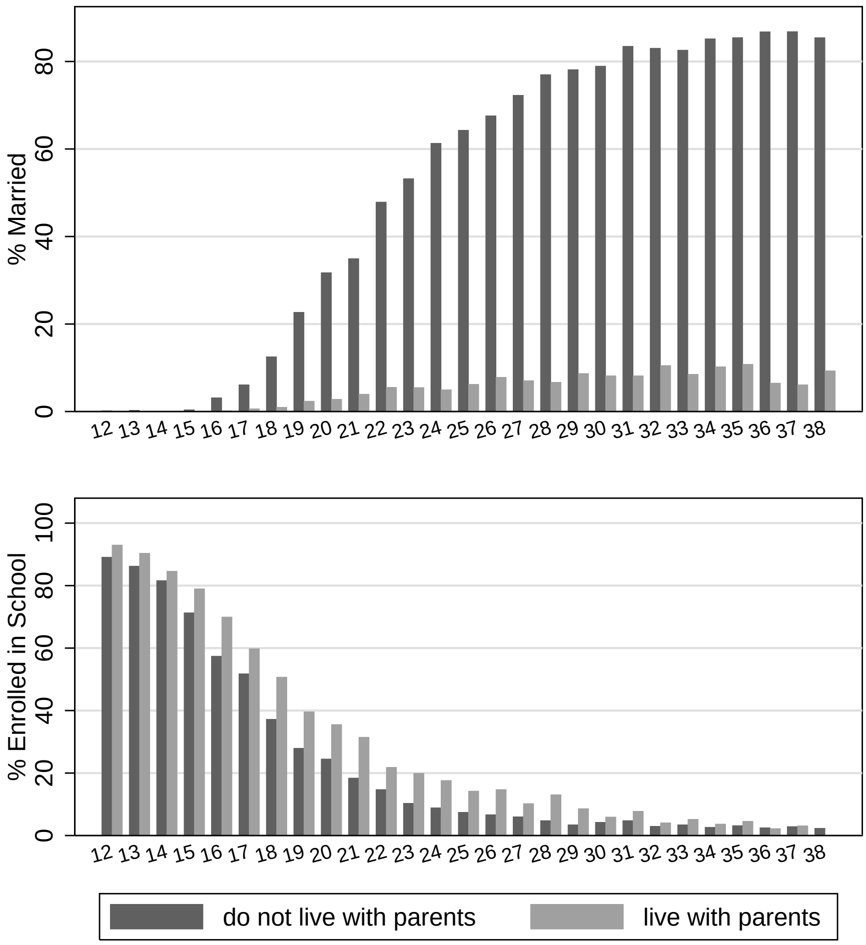
Source: PNAD.

Note: Standard errors in parentheses. Cluster: family. Coefficients of indicator variables for cohort of sons are not presented. In (a) the estimation includes level dummies for the age intervals 18-22, 23-27, 28-32 and 33-38. In (b) it is 22-27. In (c) it is 22-25.

* significant at 10%, ** significant at 5%, *** significant at 1%.

Figures

FIGURE 1: Living Arrangements of Sons, by Age, Marrital status and School Enrollment Status



Source : 1996 PNAD

FIGURE 2: Average Earnings of Fathers by Living Arrangements of Sons, and Sons' Age

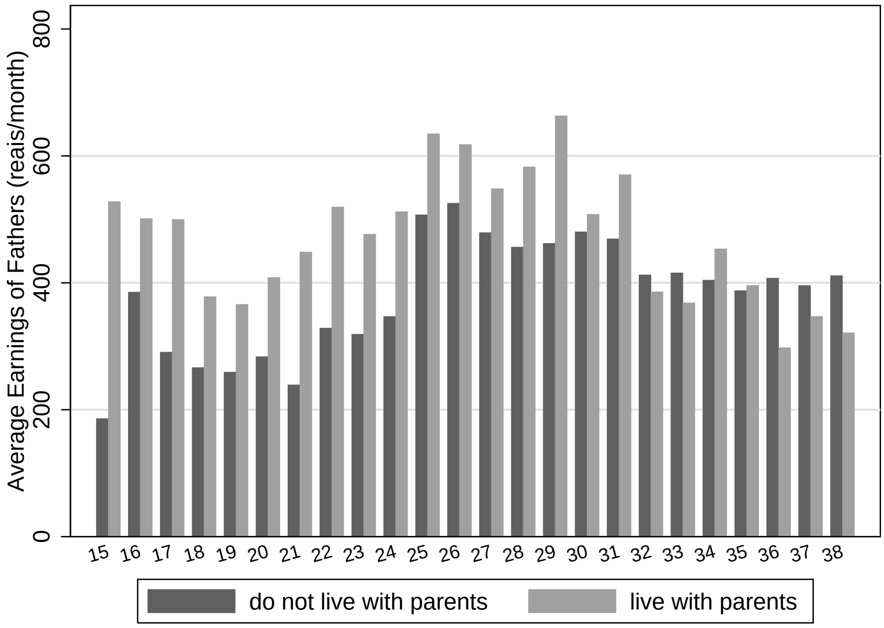
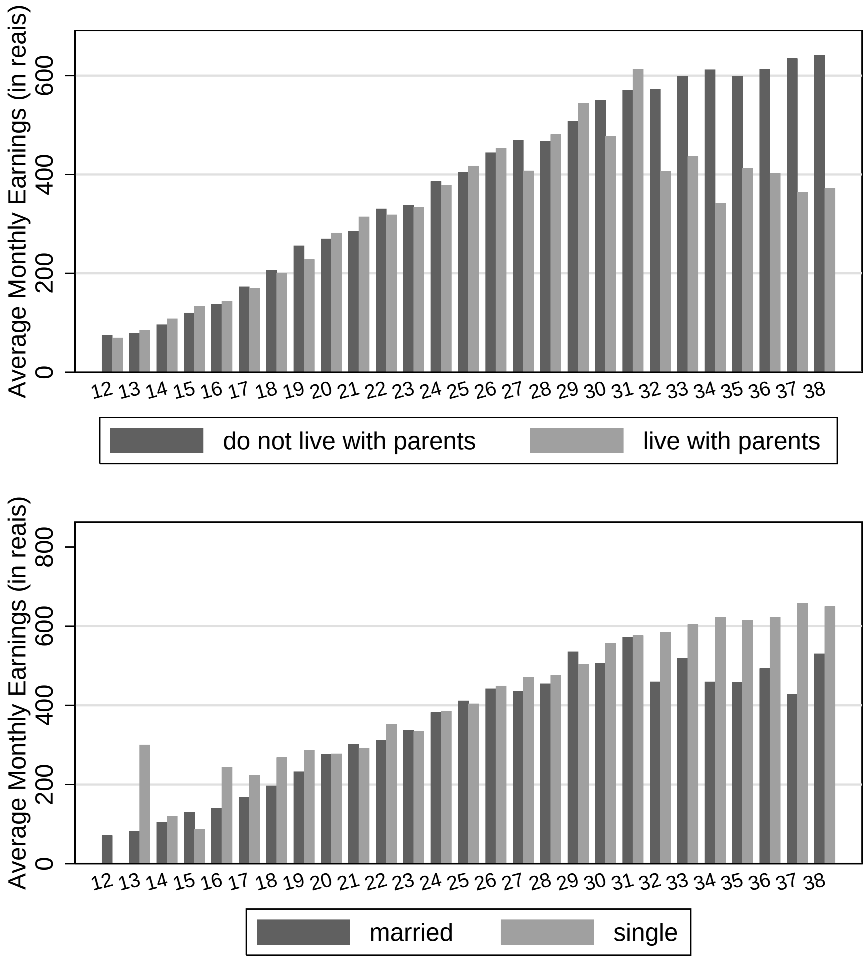


FIGURE 3: Average Earnings of Sons by Living Arrangements, Marital Status and Age



Source : 1996 PNAD

Appendix

Mathematical Derivations

The child at time t=0 will be a parent at time t=1 and will solve the same maximization problem, with the proper adjustments for time index. At t=1, the optimal choice of I_2 will depend on I_1. Let $I_2^*(I_1)$ be the parent's best response function at t=1. For every investment made in the parent when he/she was a child at t=0, $I_2^*(I_1)$ gives the optimal investment the parent should make in the child at t=1. Given that I_2 is a function of I_1, a more convenient way of writing the equation (8) would be

$$\text{(A.1)}\quad \underset{\{I_1\}}{Max}\ \{(1-\beta)\ln((1+r)I_0+A_0-I_1)+\beta\ln((1+r)I_1+a_1-I_2^{*e_0}(I_1))\}.$$

At t=0, parent solves the maximization problem for t=1

$$\underset{\{I_2\}}{Max}\ \{(1-\beta)\ln((1+r)I_1+a_1-I_2)+\beta\ln((1+r)I_2+a_2^{e_0}-I_3^*(I_2))\}.$$

At t=0, parent solves the maximization problem for t=2

$$\underset{\{I_3\}}{Max}\ \{(1-\beta)\ln((1+r)I_2+a_2^{e_0}-I_3)+\beta\ln((1+r)I_3+a_3^{e_0}-I_4^*(I_3))\}.$$

In general, at t=0 parent will solve the maximization problem for any time $t \in \mathbb{N}$,

$$\underset{\{I_{t+2}\}}{Max}\ \{(1-\beta)\ln((1+r)I_{t+1}+a_{t+1}^{e_0}-I_{t+2})+\beta\ln((1+r)I_{t+2}+a_{t+2}^{e_0} -I_{t+3}^*(I_{t+2}))\}.$$

The first order condition of this problem is

$$\frac{(1-\beta)}{(1+r)I_{t+1}+a_{t+1}^{e_0}-I_{t+2}} = \frac{\beta(1+r-I_{t+3}^{*\prime}(I_{t+2}))}{(1+r)I_{t+2}+a_{t+2}^{e_0}-I_{t+3}^*(I_{t+2})}.$$

For every $t \in \mathbb{N}$, the same maximization problem will be solved. So, for all t, the solution should be the same. Consider the following guess for the solution of the above equation, $I_{j+2} = bI_{j+1} + c$ for all $j \in \mathbb{N}$. In particular,

$I_{t+2} = bI_{t+1} + c$ and $I_{t+3} = bI_{t+2} + c$. Substitute the former two equations in the above equation to obtain

$$\frac{(1-\beta)}{(1+r)I_{t+1}+a^{e_0}_{t+1}-bI_{t+1}-c} = \frac{\beta(1+r-b)}{(1+r)(bI_{t+1}+c)+a^{e_0}_{t+2}-bI_{t+2}-c}.$$

Solve for I_{t+2} as a function of I_{t+1} to find

$$I_{t+2} = \frac{[(1-\beta)(1+r)b-\beta(1+r-b)^2]}{(1-\beta)b} I_{t+1}+ \frac{(1-\beta)rc+(1-\beta)a^{e_0}_{t+2}-\beta(1+r-b)(a^{e_0}_{t+1}-c)}{(1-\beta)b}.$$

According to the guess, the coefficient on I_{t+1} should be equal to b and the second term on the right hand side should be equal to c,

$$b = \frac{[(1-\beta)(1+r)b-\beta(1+r-b)^2]}{(1-\beta)b}$$

$$c = \frac{(1-\beta)rc+(1-\beta)a^{e_0}_{t+2}-\beta(1+r-b)(a^{e_0}_{t+1}-c)}{(1-\beta)b}.$$

Solve for b and c, to find two set of solutions,

$$b = 1+r \qquad \text{and} \qquad c = a^{e_0}_{t+2}$$

$$b = \beta(1+\mathrm{r}) \qquad \text{and} \qquad c = \frac{\beta(1+r)}{r}a^{e_0}_{t+1} - \frac{1}{r}a^{e_0}_{t+2}.$$

So, the guess is correct.

The first set of solutions is not an interesting case because it would imply that average endowed parents with average market luck and average endowed children invest all their earnings in their children.

$I_{t+2} = bI_{t+1} + c$ for the second set of solution is

$$I_{t+2} = \beta(1+r)I_{t+1} + \frac{\beta(1+r)}{r}a^{e_0}_{t+1} - \frac{1}{r}a^{e_0}_{t+2}.$$

For t=0,

$$I_2 = \beta(1+r)I_1 + \frac{\beta(1+r)}{r}a_1 - \frac{1}{r}a^{e_0}_2.$$

Substitute $a^{e_0}_2$ by $\delta a_1 + \theta a_0$ in the above equation to find

$$I_2 = \beta(1+r)I_1 - \frac{\delta-\beta(1+r)}{r}a_1 - \frac{\theta}{r}a_0.$$

At t=0, parent maximizes equation (A.1). Substitute the above equation in equation (A.1) to obtain

$$\underset{\{I_1\}}{Max}\ \{(1-\beta)\ln((1+r)I_0 + A_0 - I_1) + \beta\ln((1+r)(1-\beta)I_1 + \frac{\delta+r-\beta(1+r)}{r}a_1 + \frac{\theta}{r}a_0)\}.$$

The first order condition of this problem is

$$\frac{(1-\beta)}{(1+r)I_0+A_0-I_1} = \frac{\beta(1+r)(1-\beta)}{(1+r)(1-\beta)I_1+\frac{\delta+r-\beta(1+r)}{r}a_1+\frac{\theta}{r}a_0}.$$

Solve for I_1 to find

$$I_1 = \beta((1+r)I_0 + A_0) + \frac{\beta(1+r)-\delta-r}{r(1+r)}a_1 + \frac{\theta}{r(1+r)}a_0.$$

Equivalently,

$$I_1 = \beta y_0 + \frac{\beta(1+r)-\delta-r}{r(1+r)}a_1 + \frac{\theta}{r(1+r)}a_0.$$

Table A1: Estimation of Equation (iv) Allowing for Non-linearities

(a) OLS

(b) Heckman Two Step Estimation

(c)-(d) Heckman Maximum Likelihood Estimation

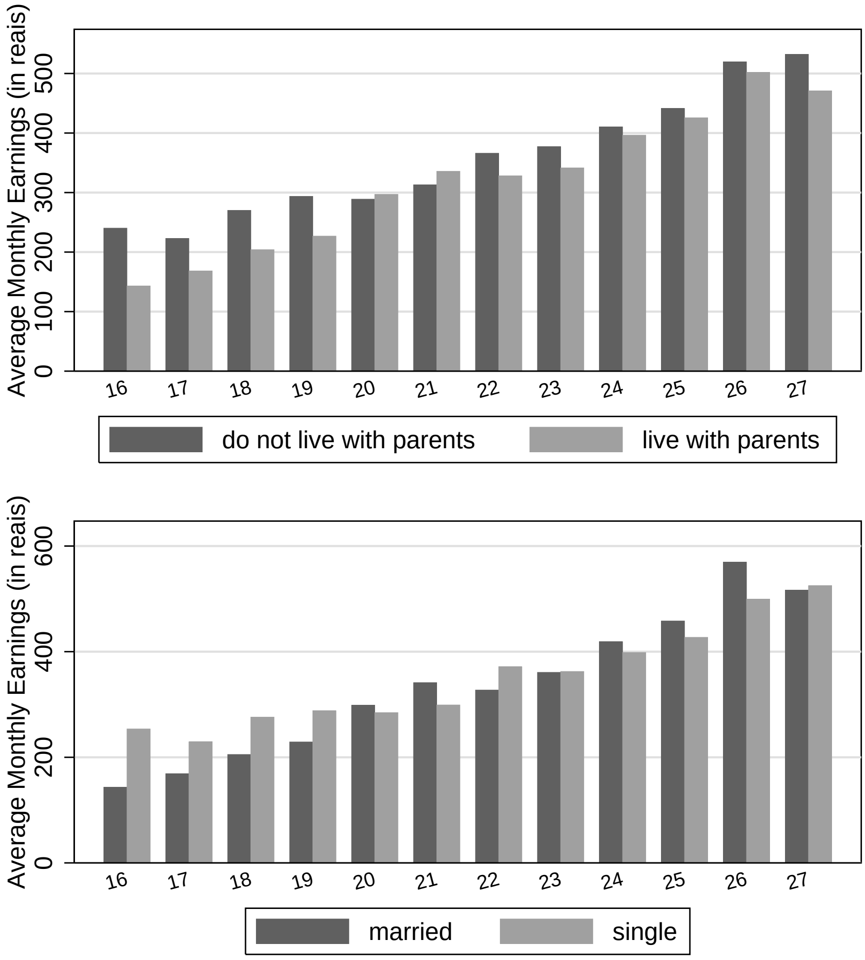
	(a)	(b)	(c)
Dependent Variable: Grandson's Earnings			
Grandfather's Earnings	1.802***	1.834***	1.836***
	(0.254)	(0.191)	(0.255)
Grandfather's Earnings Squared	–0.137***	–0.139***	–0.140***
	(0.021)	(0.016)	(0.021)
Father's Earnings	0.235***	0.235***	0.235***
	(0.015)	(0.011)	(0.015)
Son's Years of Education	0.059***	0.060***	0.060***
	(0.003)	(0.003)	(0.003)
Father's Years of Education	–0.007**	–0.006**	–0.006*
	(0.003)	(0.003)	(0.003)
Constant	–0.919	–1.047*	–1.051
	(0.762)	(0.574)	(0.765)
Older Cohort	0.313***	0.273***	0.274***
	(0.019)	(0.020)	(0.020)
Selection Equation			
Married		–2.570***	–2.553***
		(0.042)	(0.035)
Age		–0.157***	–0.156***
		(0.007)	(0.004)
Enrolled in School		0.144***	0.177***
		(0.054)	(0.031)
Father's Earnings (Imputed)		0.190***	0.175***
		(0.027)	(0.018)
Constant		3.474***	3.527***
		(0.184)	(0.111)
R^2	0.395		
LR Test of Indep. Equations			21.667
N. of Observations Uncensored		5,125	5,125
N. of Observations	5,125	10,176	10,176

Source: PNAD.

Note: Standard errors in parentheses. Cluster: family (except for the two step regression). Omitted age: 16-21 years old.

* significant at 10%, ** significant at 5%, *** significant at 1%.

FIGURE A1: Average Earnings of Sons by Living Arrangements, Marital Status and Age



Source : 1996 PNAD