Equilibrium Effects of Education Policies: A Quantitative Evaluation

Giovanni Gallipoli
University of British Columbia
gallipol@interchange.ubc.ca

Costas Meghir
University College London, and IFS
c.meghir@ucl.ac.uk

Giovanni L. Violante
New York University, CEPR, and NBER
glv2@nyu.edu

Preliminary – First Draft: April 2002. This draft: September 2010

Abstract
This paper compares the partial and general equilibrium effects of alternative education policies on the distribution of education and earnings. We build a life-cycle model with endogenous labor supply, consumption/saving and education choices, allowing for agents' heterogeneity in several dimensions and for incomplete insurance markets. The model internalizes the dynamic life-cycle effects of access to family resources by allowing altruistic parents to make a voluntary inter-vivos transfer to their child at the beginning of the child's independent life. PSID and NLSY data are used to estimate education-specific dynamic earnings processes, and the distribution of ability in the population as well as the size of transfers in different population groups. Using NLSY data, we estimate a transition process for ability which concisely characterizes the intergenerational transfer of skills. Through numerical simulations, we compare the effects of alternative policy interventions on optimal education decisions, inequality, and output. In particular, we experiment with conditional tuition subsidies. While in partial equilibrium such policies can be very effective in increasing education levels and reducing inequality, in general equilibrium the results are starkly different: the main effect of a subsidy is to increase the supply of human capital as one would expect. However, it is the more able but liquidity constrained individuals who take up extra education, while the education levels of the less able can actually decrease (they are crowded out). Thus the subsidy strongly acts on the composition of those in education. We find that subsidies made conditional on financial resources are generally preferable to those conditional on ability and large equilibrium effects can be induced by relatively small changes in marginal returns.

1 Introduction

This paper examines policies designed to alter the equilibrium distribution of education and their wider economic consequences. It also looks at the nature of education decisions and the role that such decisions play in shaping life cycle earnings and wealth profiles. Individual choices are analyzed in the context of a general equilibrium model with separate, education-specific spot markets for jobs. The unit price of (efficiency-weighted) labor differs by education group and equals marginal product.

We are interested in the equilibrium, long-term effects of policy interventions targeting the wider population rather than limited groups, with relative labor prices endogenously adjusting to changes in the aggregate supply of educated people. We examine traditional policies, such as tuition transfers and loan subsidies, but we also devise and evaluate alternative forms of policy intervention.[1] The policy experiments are carried out through numerical simulations, with some of the model's parameters directly estimated from PSID, NLSY and CPS data and others calibrated to match specific long-term features of the US economy. By simulating and comparing equilibrium outcomes we aim to explore the quantitative aspects of the relationship among schooling decisions, wages inequality and education policy. The impact of diverse education policies on equilibrium measures of productivity, consumption and welfare is also considered.

Research linking human capital (HC) investment to life cycle earnings dates back to original work by Mincer (1958), Becker (1964) and Ben-Porath (1967). The first studies ignored the important issue of self selection into education, as described by Rosen (1977) and Willis and Rosen (1979). Both permanent and persistent individual characteristics are now acknowledged as important determinants of education choices and have become a standard feature of HC models. Empirical evidence supporting the plausibility of a link between human capital accumulation and economic inequality has been provided, among others, by Mincer (1994).

In work relating education policies and individual preferences Fernandez and Rogerson (1995) originally point out that heterogeneity among individuals, whether in terms of income, ability or locality, can generate conflicting preferences as to the kind of policies

[1]Standard education policy is just one of the possible types of human capital policy. For example, changes in proportional income taxation affect the life-cycle returns on human capital and the opportunity costs of education, altering human capital investment decisions.

that are most desirable.[2]

Studies on the evaluation of policy interventions in equilibrium are more recent. Heckman, Lochner, and Taber (1998b, 1998c) have led the way in advocating an approach to policy evaluation which does not overlook equilibrium effects induced by the policy.[3] In fact, statements regarding the effects of policy interventions which ignore price changes induced by such interventions can be misleading. Fernandez and Rogerson (1998) provide an interesting application of general equilibrium (G.E.) modelling to the evaluation of education-finance reform in the US. Later work by Cunha, Heckman, and Navarro (2005) reinforces the view that models that are able to construct equilibrium counterfactuals are essential to understanding the wider consequences of policy interventions.

In the empirical literature on education policy, early work by Keane and Wolpin (1997) focuses on the partial equilibrium effect of a tuition subsidy on young males' college participation. A valuable generalization of their approach within a dynamic GE framework is due to Lee (2005) and Lee and Wolpin (2006). Also Abraham (2001) examines wage inequality and education policy in a GE model of skill biased technological change. All these studies restrict labor supply to be fixed, although earlier theoretical research has uncovered interesting aspects of the joint determination of life cycle labor supply and HC investment, among others Blinder and Weiss (1976).

Our model incorporates several important extensions with respect to earlier work: first, optimal individual labor supplies are an essential part of the lifetime earnings mechanism; second, agents' heterogeneity has different dimensions, including a permanent (ability) component and uninsurable efficiency shocks; third, ability is transmitted across generations; fourth, inter-vivos transfers from parents to offsprings are permitted to ease liquidity constraints in the education decision.

Recent empirical evidence in Hyslop (2001) indicates that labor supply explains over 20 % of the rise in (both permanent and transitory) family inequality during the period of rising wage inequality in the early 1980's. Moreover, even if individual labor supplies do not deviate much from the average levels of their demographic group, it is the case that average levels differ substantially between groups.

[2] Fernandez and Rogerson (1995) consider ex-ante identical individuals who differ only in income

[3] Heckman, Lochner, and Taber estimate and simulate a dynamic general equilibrium model of education accumulation, assets accumulation and labor earnings with skill-biased technological change.

The other second extension in our model is the introduction of individual uncertainty over the returns to HC in the form of idiosyncratic multiplicative shocks to labor efficiency. As Levhari and Weiss (1974) originally emphasized, uncertainty is of exceptional importance in human capital investment decisions as the risk associated to such decisions is usually not insurable nor diversifiable. Using a multiplicative form of earnings risk, Eaton and Rosen (1980) show how earnings taxation has an ambiguous effect on investment in human capital because it impinges on two important parameters of the decision problem: for one, taxation reduces the riskiness of returns to human capital investment.[4] In addition, taxation induces an income effect that can influence the agents' willingness to bear risk. Thus, ignoring the riskiness of education decisions can partly sway the results in the analysis of the effects of earnings taxation and education policies.

Gale and Scholz (1994) show that inter vivos transfer for education are sizeable, thus they should be incorporated in the mechanism of a model of education acquisition, especially if one is interested in quantifying the role of credit constraints. Altonji, Hayashi, and Kotlikoff (1997) provide both empirical evidence and insights regarding the nature of inter-vivos transfers using a model with income uncertainty where transfers can be used as an insurance devise by potentially altruistic parents; De Nardi (2004) argues that voluntary bequests and inter-generational earnings' correlation are important determinants of the distribution of wealth.

We also calibrate the level of correlation between ability of parents and kids. Besides genetic transmission, this can be thought of as a way to incorporate the effect of parental background on ability formation, as extensively documented in the literature, see Heckman and Carneiro (2003) for a review.[5]

We model three levels of education obtained through formal schooling and corresponding to three types of HC which enter the production technology.[6] Education and employment are mutually exclusive in each period. Foregone earnings and tuition charges

[4]As the proportional tax rate increases, agents earn less from high realization of the shock but also lose less from the bad ones. Therefore the overall risk is decreased.

[5]The simulations reported in the current draft do not yet embed inter-vivos transfers and inter-generational correlation in ability.

[6]We distinguish among people with less than high school degrees (LTHS), high school graduates (HSG) and college graduates (CG). The distinction between LTHS and HSG is based on different earning and labor supply characteristics. Schooling is the only way to accumulate human capital (no children nurturing or on-the-job training). The possible effects of OJT are accounted for through an age-efficiency profile which is estimated for each education group and is maintained to be policy-invariant.

are the direct costs of schooling, and a utility cost comes in the form of reductions in leisure when studying.

In general, the model provides a way to look at endogenous equilibrium levels of aggregate human capital, with associated wages, as a function of agents' optimizing schooling choices and demographic factors. Through its policy functions, it provides a mapping from a set of initial conditions (that is, initial agents' distribution over states such as permanent and persistent idiosyncratic shocks and assets) into distributions over educational and economic attainments: this mapping turns out to be ideal to study the economic implications of alternative policy interventions.

2 Model

2.1 Demographics and the life cycle

Demographics: The economy is populated by J overlapping generations. Let $j = 1, ..., J$ denote age. The conditional survival rate at age j is ζ_j. We let $\zeta_j = 1$ as long as the individual is in school or at work, but $\zeta_j < 1$ during retirement, from $j = j^{RET} + 1$ to J. Conditional on reaching age J, death is certain in the next period ($\zeta_J = 0$). We set the size of the newborn cohort so that total population is normalized to 1.

Life cycle: The life cycle of an individual has three distinct stages. In the first stage, the individual goes to school and acquires education. There are three possible educational attainments denoted by $e \in \{LHS, HSG, COL\}$ standing for Less than High-School, High-School Graduate and College Graduate. Let j^e denote the last period of the school cycle e, with the convention that j^{LHS} is the last period of compulsory high-school education. We normalize this age to 0. Until that age, individuals are "children", live with their parents and depend financially from them.

Individuals start their first stage of economic life when when they begin making independent educational and financial decisions, i.e. at age $j = 1$. At this age, they immediately decide whether dropping out of school or continuing. Such decision, denoted by $d^{HSG} \in \{0, 1\}$, entails commitment to be in school until age j^{HSG}. Next, at age $j^{HSG} + 1$ the agent decides whether to continue education to achieve a college degree. Once again, this education decision which we denote by $d^{COL} \in \{0, 1\}$ requires committing to be in school until age j^{COL}. During education, students choose their level of consumption/saving, while leisure is exogenously fixed at $\bar{l}$.

At age $j^{COL}+1$, or at ages $j^{HSG}+1$ and 1 for those who do not opt for continuing schooling, individuals begin their second stage of the life cycle: work. During this stage, which lasts until mandatory retirement, agents choose labor supply and consumption/saving.

From age j^{RET}, the last stage of life begins: retirement. During retirement, individuals do not work $(l=1)$, receive a pension from the government and allocate consumption/saving over their uncertain remaining lifetime.

2.2 Preferences and intergenerational links

Preferences: The period utility of workers and retirees $u(c_j, l_j)$ is strictly increasing and concave in consumption $c \geq 0$ and leisure $l \in [0,1]$. During schooling, a student's leisure is fixed at $\bar{l}$ and her utility has an additional, separable, component $\kappa(\theta)$, a strictly decreasing function of fixed individual innate "ability" $\theta \in [\theta_{\min}, \theta_{\max}]$. The function $\kappa(\theta)$ reflects psychic costs of schooling in terms of effort or like/dislike of the education process (see Heckman, Lochner, and Todd (2005)).

Intergenerational links: Individuals are altruistic towards their off-springs and value the expected lifetime utility of their children with weight ω relative to their own lifetime utility. This one-sided altruism manifests itself as a monetary transfer only once in the lifetime. At age j^{TRA}, during the work-stage, each individual (now a parent) has the opportunity to choose a positive amount to transfer to her/his child who, next period, will enter her/his education stage with age $j=1$. The parental transfer determines entirely the child's initial asset level.

Individuals are also linked by the intergenerational transmission of ability from parents to children. An individual with ability θ has a probability of having a child with ability less than or equal to $\hat{\theta}$ determined by the c.d.f. $\Gamma_\theta\left(\hat{\theta}, \theta\right)$. At parent's age j^{TRA}, before the inter-vivos transfer from parent to child, the ability of the child is fully revealed to both.

2.3 Labor productivity and production technology

Labor productivity: Individual labor efficiency ε_j^e for individual of education e at age j is the sum of three components: in logs,

$$\ln \varepsilon_j^e = \lambda \ln \theta + \xi_j^e + z_j^e \tag{1}$$

where λ is a loading factor on (log-) ability mediating the effect of innate ability on productivity, ξ_j^e is an education-specific age profile for productivity, and z_j^e is a stochastic component drawn from the education-specific c.d.f. $\Gamma_z^e(z_{j+1}, z_j)$ describing the conditional cumulative probability of a realization less than or equal to z_{j+1} at age $j+1$ when the idiosyncratic stochastic component at age j was z_j. Let Γ_0^e denote the initial distribution of productivity upon entry in the labor market with educational level e. We assume that during schooling labor productivity is zero, i.e. schooling and work are mutually exclusive.

Technology: The representative firm operates a constant returns to scale (CRS) technology

$$F\left(K, \mathcal{H}\left(H_{LHS}, H_{HSG}, H_{COL}\right)\right)$$

which employs physical capital K and the three types of human capital bundled in the aggregator $\mathcal{H}$, also displaying CRS. Capital depreciates at rate δ. Each human capital stock H_e is the sum over all working-age individuals within each education group e, of individual hours worked times their respective efficiency units of labor ε_j^e.

2.4 Commodities and markets

There are three commodities: (i) the final good which can be used for private/public consumption, private expenditures in education services, and investment, (ii) efficiency units of labor, and (iii) intermediary services provided by the banking sector. They are all exchanged in competitive labor markets. We let the price of the final good act as the numeraire.

There is only one financial asset: a risk-free bond (claim on physical capital) traded in a competitive market and used by households as self-insurance device. The bond is exchanged through the banking system. Households with positive savings receive by the banks an equilibrium interest rate r^+. Banks lend the funds to households with borrowing needs at the rate r^-. The wedge between the two interest rates is generated by the intermediation cost $\iota > 0$ per unit of consumption intermediated. In other words, savers must be indifferent between lending to firms or to banks with return r^+. Since banks pay the cost ι, in a competitive equilibrium where banks break even, borrowers are charged $r^- = r^+ + \iota$.

In what follows, we use the notation r for interest rates on saving/loans contracts with the private sector with the convention that if $a \geq 0$, then $r = r^+$, and $r = r^-$ otherwise.

Individuals face different private debt limits, depending on which phase of the life-cycle they are going through. During retirement, the debt limit is set at zero: retirees cannot borrow against their social security benefits. In the work-stage, agents can borrow in private markets up to $\underline{a}$. Conditional on being in college, individuals can borrow up to $\underline{a}^{COL}$, at the equilibrium rate r^-, from the private sector. High-school students cannot borrow.

2.5 Education and fiscal policies

Education policies: The government offers two types of help to students who are considering college education and face per-period tuition fees ϕ. First, per-period grants of size $g(a, \theta)$ whose size may depend on the wealth a of the applicant as well as on her ability θ. Second, education loans up to a limit $\underline{b}$. Government loans have a fixed repayment scheme: for n periods since employment, the individual repays an amount π every period until exhaustion of all the principal plus interests. Therefore, the last period of repayment in the individual life cycle is $j^{REP} = j^{COL} + n < j^{TRA}$.[7] Clearly, the size of π depends on the amount borrowed: the smaller is the government-sponsored loan, the smaller the associated repayment π.

The government subsidizes loans in two ways. It fully subsidizes interests payments during college, and sets a fixed interest rate $r^b < r^-$ on the principal, during the repayment period. In particular, if at the end of college the individual has borrowed an amount b from the government, π is determined by the actuarial formula

$$\pi = b \frac{r^b}{1 - [1 + r^b]^{-n}},$$

[7] The assumption that n is such that the individual must finish its repayment before the inter-vivos transfer is made for tractability. This restriction is not binding when the model is calibrated to US data on typical repayment periods of this type of debt contracts which is 10 years.

which shows that, given the policy parameter pair $\left(n, r^b\right)$, there is a one-to-one mapping between b and π.

Some comments are in order about government vs private borrowing. Government borrowing has two advantages: 1) lower rates, 2) foregone interest during education. However, it has a fixed repayment schedule. This is done mostly for simplicity, to avoid carrying around two financial assets. The implicit assumption is that the agent prefers goverment borrowing: he'll do that (all the way up to $\underline{b}$) before starting borrowing from banks. We can justify this assumption in two ways. First, the lender needs to screen the borrower before agreeing to the contract, and that is costly. That the government has lended to the agent is a signal, the agent is viable. In short, the private sector let the government pay this screening cost.

Second, we can assume that the private education loan has exactly the same fixed repayment schedule as the government loan. However, because of the two advantages discussed above, the individual will strictly prefer the government loan, up to its limit. Since when work starts, the individual can borrow (with free repayment schedule) up to $\underline{a}$, it is optimal to repay all the education loan through new private loans which carry the same rate, but a more flexible schedule. This is exactly equivalent to our model as well.

Finally, we note that in some cases, the agent may want to repay the government faster. We are not allowing this. However note that if $r^b \leq r^+$this decision is never optimal. We may want to choose such parameterization eventually.

Fiscal policies: The government levies proportional taxes at rate τ_c on consumption, τ_w on labor earnings, and τ_k on capital income. The tax τ_k is levied only on positive capital income, so with a slight abuse of notation, we use τ_k throughout, with the convention that if $a < 0$ (and $r = r^-$), then $\tau_k = 0$. Tax revenues are used to finance education policies, government consumption G, and a social security system that pays flat pension benefits p^e to all workers of type e.

2.6 The individual problem in recursive form

It is convenient to describe the individual problem going backward, from retirement to schooling.

Retirement stage: From age j^{RET} to age J, the individual solves:

$$\begin{aligned} \Omega_j(e, a_j) &= \max_{c_j} \; u(c_j, 1) + \beta\zeta_j\Omega_{j+1}(e, a_{j+1}) \\ & \quad s.t. \\ (1+\tau_c)c_j + a_{j+1} &= p^e + (\zeta_j)^{-1}\left[1 + r(1-\tau_k)\right]a_j \\ a_{j+1} &\geq 0 \end{aligned} \tag{2}$$

where p^e is a social security benefit conditional on the education level and explains why e remains a state variable of this problem besides wealth a_j. The term ζ_j in the budget constraint reflects the perfect annuity markets assumption. The retired agent does not work ($l_j = 1$) and cannot borrow.

Work stage after the inter-vivos transfer: From age $j^{TRA} + 1$ until retirement, the individual solves:

$$\begin{aligned} W_j(e, a_j, \theta, z_j) &= \max_{c_j, l_j} \; u(c_j, l_j) + \beta\mathbb{E}_z W_{j+1}(e, a_{j+1}, \theta, z_{j+1}) \\ & \quad s.t. \\ (1+\tau_c)c_j + a_{j+1} &= (1-\tau_w)w^e\varepsilon_j^e(\theta, z_j)(1-l_j) + \left[1 + r(1-\tau_k)\right]a_j \\ a_{j+1} &\geq -\underline{a} \\ z_{j+1} &\sim \Gamma_z^e(z_{j+1}, z_j) \end{aligned} \tag{3}$$

The individual states of this problem are the education level e, asset holdings a_j, ability θ, and the persistent productivity shock z_j. The variable w^e is the price of an efficiency unit ε_j^e of labor of type e. Workers can borrow up to an exogenously set debt limit $\underline{a}$ from private markets. In the last period before retirement $\left(j = j^{RET} - 1\right)$ the continuation value is replaced by $\beta\zeta_{j^{RET}}\Omega_{j^{RET}}\left(e, a_{j^{RET}}\right)$.

Work stage in the period of the inter-vivos transfer: At age j^{TRA}, the individual problem reads:

$$\begin{aligned} W_j\left(e, a_j, \theta, z_j, \hat{\theta}\right) &= \max_{c_j, l_j \hat{a}_1} \; u(c_j, l_j) + \beta\left[\mathbb{E}_z W_{j+1}(e, a_{j+1}, \theta, z_{j+1}) + \omega\mathbb{E}_{\hat{z}}V^*\left(\hat{a}_1, \hat{\theta}, \hat{z}_1\right)\right] \\ & \quad s.t. \\ (1+\tau_c)c_j + a_{j+1} + \hat{a}_1 &= (1-\tau_w)w^e\varepsilon_j^e(\theta, z_j)(1-l_j) + \left[1 + r(1-\tau_k)\right]a_j \\ a_{j+1} &\geq -\underline{a}, \quad \hat{a}_1 \geq 0 \\ z_{j+1} &\sim \Gamma_z^e(z_{j+1}, z_j), \quad \hat{z}_1 \sim \Gamma_0^{LHS} \end{aligned} \tag{4}$$

The parent puts weight ω on the discounted utility $V^*\left(\hat{a}_1, \hat{\theta}, \hat{z}_1\right)$ of her child. At this point, the parent knows his child ability $\hat{\theta}$ but still needs to form expectations about the child's productivity next period in order to choose the transfer $\hat{a}_1$. The transfer determines the initial asset position of the child in the period where she becomes an independent decision maker. The constraint $\hat{a}_1 \geq 0$ means that parents cannot force kids to transfer resources to them.[8]

Work stage after full repayment of government-sponsored loan & before the inter-vivos transfer: From age $j^{REP}+1$ until age $j^{TRA}-1$, the household's problem is exactly as in (3). The only difference being that, in the last period before the transfer $\left(j = j^{TRA}-1\right)$ the continuation value is replaced by $\beta W_{j^{TRA}}\left(e, a_{j^{TRA}}, \theta, z_{j^{TRA}}, \hat{\theta}\right)$.

Work stage before full repayment of government-sponsored loan: In this stage, the individual solves:

$$
\begin{aligned}
W_j\left(e, a_j, \theta, z_j, b\right) &= \max_{c_j, l_j} \; u\left(c_j, l_j\right) + \beta \mathbb{E}_z W_{j+1}\left(e, a_{j+1}, \theta, z_{j+1}, b\right) \\
& \quad s.t. \\
\left(1+\tau_c\right) c_j + a_{j+1} &= \left(1-\tau_w\right) w^e \varepsilon_j^e\left(\theta, z_j\right)\left(1-l_j\right) + \left[1 + r\left(1-\tau_k\right)\right] a_j - \pi \\
a_{j+1} &\geq -\underline{a} \\
z_{j+1} &\sim \Gamma_z^e\left(z_{j+1}, z_j\right) \\
\pi &= \left(b r^b\right) / \left[1 - \left(1+r^b\right)^{-n}\right]
\end{aligned} \tag{5}
$$

where the main difference with problem (3) is the presence of the additional state variable b, the size of the government-sponsored education loan. In the period just before the transfer (age $j = j^{TRA}-1$), the continuation value is replaced by $\beta \mathbb{E}_{z,\hat{\theta}} W_{j^{TRA}}\left(e, a_{j^{TRA}}, \theta, z_{j^{TRA}}, \hat{\theta}\right)$, defined above in equation (4).

Initial period of the work stage: In the first period of working life, the agent has the same value function as in (5), with:

$$
b = \begin{cases} 0 & \text{if } e \in \{LHS, HSG\}, \text{ or } e = COL \text{ and } \hat{a}_{j^{COL}+1} \geq 0 \\ \hat{a}_{j^{COL}+1} & \text{if } e = COL \text{ and } 0 \geq \hat{a}_{j^{COL}+1} > -\underline{b} \\ -\underline{b} & \text{if } e = COL \text{ and } a_{j^{COL}+1} \leq -\underline{b} \end{cases}
$$

[8] This constraint is here for clarity, but it is not necessary to restrict the solution to the optimization problem. Given that at age $j = 1$ (high-school) students cannot borrow, the parent cannot use the kid to loosen his borrowing constraint.

and with a_j determined as

$$a_j = \begin{cases} \hat{a}_j & \text{if } e = LHS \text{ and } j = 1 \\ \hat{a}_j & \text{if } e = HSG \text{ and } j = j^{HSG} + 1 \\ \hat{a}_j - b & \text{if } e = COL \text{ and } j = j^{COL} + 1 \end{cases}$$

Moreover, the individual faces a different set of constraints. First, the individual of education e draws an initial productivity level z_j from the education-specific distribution Γ_0^e,

$$z_j \sim \begin{cases} \Gamma_0^{LHS} & \text{if } e = LHS \text{ and } j = 1 \\ \Gamma_0^{HSG} & \text{if } e = HSG \text{ and } j = j^{HSG} + 1 \\ \Gamma_0^{COL} & \text{if } e = COL \text{ and } j = j^{COL} + 1 \end{cases}$$

This representation of the initial budget constraint allows to distinguish the subsidized debt which is repaid with fixed installments π to the government from the savings/loans contracts stipulated with the private sector. As explained above, π is determined by the actuarial formula $\pi = \left(br^b\right) / \left[1 - \left(1 + r^b\right)^{-n}\right]$.

Recall that since individuals with $e = LHS$ or $e = HSG$ cannot hold debt during their schooling period, they will always enter employment with nonnegative assets.

College education: College strudents, between ages $j^{HSG} + 1$ and j^{COL} solve

$$V_j\left(COL, \hat{a}_j, \theta\right) = \max_{c_j} \; u\left(c_j, \bar{l}\right) - \kappa\left(\theta\right) + \beta V_{j+1}\left(COL, \hat{a}_{j+1}, \theta\right) \tag{6}$$

$$s.t.$$

$$\left(1 + \tau_c\right) c_j + \hat{a}_{j+1} =$$

$$= \begin{cases} \left[1 + r\left(1 - \tau_k\right)\right] \hat{a}_j - \phi + g\left(\hat{a}_j, \theta\right) & \text{if } \hat{a}_j \geq 0 \\ \hat{a}_j - \phi + g\left(\hat{a}_j, \theta\right) & \text{if } 0 > \hat{a}_j > -\underline{b} \\ \left[1 + r\left(1 - \tau_k\right)\right]\left(\hat{a}_j + \underline{b}\right) - \underline{b} - \phi + g\left(\hat{a}_j, \theta\right) & \text{if } \hat{a}_j \leq -\underline{b} \end{cases}$$

$$\hat{a}_{j+1} \geq -\left(\underline{b} + \underline{a}^{PVT}\right)$$

Recall that ϕ are the per-period tuition fees and $g\left(a_j, \theta\right)$ is the means-tested government grant. The composite budget constraint above reflects the fact that if the individual is borrowing from the government, then she does not repay interests until after employment, while if she borrows from private markets, she starts repaying the market interest rate right away. Note that, since the student has no source of income during college and $\phi > g\left(a_j, \theta\right)$, his asset position will always deteriorate, unless her net capital income is large enough..[9]

Finally, note that the continuation value in the last period of college is replaced by $\beta \mathbb{E}_z W_{j+1}\left(COL, a_{j+1}, \theta, z_{j+1}, b\right)$ where $z_{j+1} \sim \Gamma_0^{COL}$ and where b is determined as explained above.

College decision: At age $j = j^{HSG} + 1$, the students draws $z_j \sim \Gamma_{z_0}^{HSG}$ and solves

$$V^{**}\left(\hat{a}_j, \theta, z_j\right) = \max \ \left\{V_j\left(COL, \hat{a}_j, \theta\right), W_j\left(HSG, \hat{a}_j, \theta, z_j\right)\right\} \tag{7}$$

The dummy variable $d^{COL} \in \{0,1\}$ reflects the college education decision.[10]

High-school education: A high-school student solves:

$$\begin{aligned} V_j\left(HSG, \hat{a}_j, \theta\right) &= \max_{c_j} \ u\left(c_j, \bar{l}\right) + \beta V_{j+1}\left(HSG, \hat{a}_{j+1}, \theta\right) \\ & \quad s.t. \\ \left(1+\tau_c\right) c_j + \hat{a}_{j+1} &= \left[1 + r\left(1-\tau_k\right)\right] \hat{a}_j \\ \hat{a}_{j+1} &\geq 0 \end{aligned} \tag{8}$$

No borrowing is allowed to high-school students. In the last period of high-school $\left(j^{HSG}-1\right)$, the continuation value is $\beta \mathbb{E}_z V^{**}\left(\hat{a}_{j+1}, \theta, z_{j+1}\right)$ where V^{**} is defined above.

High-school decision: At age $j = 1$, the students draws $z_1 \sim \Gamma_{z_0}^{LHS}$ and solves the following maximization problem

$$V^{*}\left(\hat{a}_1, \theta, z_1\right) = \max \ \left\{V_1\left(HSG, \hat{a}_1, \theta\right), W_1\left(LHS, \hat{a}_1, \theta, z_1\right)\right\} \tag{9}$$

where $\hat{a}_1$ is the transfer received by the parents. The dummy variable $d^{HSG} \in \{0,1\}$ reflects the high-school education decision .

[9] It follows immediately that, when making the college attendance decision (described below) the student will never borrow up to the constraint.

[10] The presence of discrete education choices introduces non-convexities in the budget sets. This implies that standard results on uniqueness and continuity of optimal policy functions cannot be applied to this problem. For a discussion of related issues and the numerical solution of this problem see Gallipoli and Nesheim (2007).

2.7 Equilibrium

It is useful to introduce some additional notation to simplify the description of the equilibrium. Let $\mathbf{s}_j \in S_j$ denote the age-specific state vector implicit in the recursive representation of the agent problems above. We also define $\mathbf{s}_j^e$ to be the state vector minus the education level, i.e. $\mathbf{s}_j^e \equiv \{\mathbf{s}_j \setminus e\} \in S_j^e$.

A *stationary recursive competitive equilibrium* for this economy is a collection of: (i) individual decision rules for consumption, leisure and wealth holdings $\{c_j(\mathbf{s}_j), l_j(\mathbf{s}_j), a_{j+1}(\mathbf{s}_j)\}$ inter-vivos transfers $\left\{\hat{a}_1\left(\mathbf{s}_{j^{TRA}}\right)\right\}$, and education choices $\left\{d^{HSG}(\mathbf{s}_1), d^{COL}\left(\mathbf{s}_{j^{HSG}}\right)\right\}$; (ii) value functions $\{V_j(\mathbf{s}_j), W_j(\mathbf{s}_j), \Omega_j(\mathbf{s}_j)\}$; (iii) aggregate capital and labor inputs $\{K, H_{LHS}, H_{HSG}, H_{COL}\}$; (iv) prices $\left\{r, w^{LHS}, w^{HSG}, w^{COL}\right\}$; (v) age and education specific measures $\left\{\mu_j^e\right\}$ such that:

1. Given prices $\left\{r, w^{LHS}, w^{HSG}, w^{COL}\right\}$, the individual decision rules $\left\{c_j(\mathbf{s}_j), l_j(\mathbf{s}_j), a_{j+1}(\mathbf{s}_j), \hat{a}\left(\mathbf{s}_{j^{TRA}}\right), d^{HSG}(\mathbf{s}_1), d^{COL}\left(\mathbf{s}_{j^{HSG}}\right)\right\}$ solve their respective individual problems (2), (3), (4), (5), (6), and (8). And $\{V_j(\mathbf{s}_j), W_j(\mathbf{s}_j), \Omega_j(\mathbf{s}_j)\}$ are the associated value functions.

2. Given prices $\{r, w_{LHS}, w_{HSG}, w_{COL}\}$, the representative firm chooses optimally factors of productions and prices are marginal productivities

$$\begin{aligned} r^+ + \delta &= F_K\left(K, \mathcal{H}\left(H_{LHS}, H_{HSG}, H_{COL}\right)\right) \\ w_e &= F_{H_e}\left(K, \mathcal{H}\left(H_{LHS}, H_{HSG}, H_{COL}\right)\right), \text{ for } e \in \{LHS, HDG, COL\}. \end{aligned}$$

3. The labor markets for each educational level clear

$$H_e = \sum_{j=j^e+1}^{j^{RET}-1} \int_{S_j^e} \varepsilon^e \left[1 - l\left(e, \mathbf{s}_j^e\right)\right] d\mu_j^e \text{ for } e \in \{LHS, HDG, COL\}$$

4. The intermediation market clears: $r^- = r^+ + \iota$

5. The asset market clears

$$K = \sum_{\substack{e=COL, j \geq j^{COL}+1 \\ e \neq COL, j \geq 1}} \int_{S_j^e} a_j\left(e, \mathbf{s}_j^e\right) d\mu_j^e + \sum_{e=COL, j \geq j^{HSG}+1}^{j^{COL}} \int_{S_j^e} \left[I_{\{a_j>0\}} a_j\left(e, \mathbf{s}_j^e\right) + I_{\{a_j<-\underline{b}\}}\left(a_j\left(e, \mathbf{s}_j^e\right) + \underline{b}\right)\right.$$

6. The goods market clears

$$\sum_{e,j}\int_{S_j^e} c_j\left(e,\mathbf{s}_j^e\right)d\mu_j^e+\delta K+G+\phi\sum_{j=j^{HSG}+1}^{j^{COL}}\int_{S_j^{COL}} d\mu_j^{COL}+\Upsilon=F\left(K,\mathcal{H}\left(H_{LHS},H_{HSG},H_{COL}\right)\right)$$

where the last term in the left-hand-side reflects the private expenditures in educational services by college students, and Υ is the output of the intermediation sector

$$\Upsilon=\iota\cdot\sum_{\substack{e=COL,j\geq j^{COL}+1\\ e\neq COL,j\geq 1}}\int_{S_j^e} I_{\{a_j<0\}}a_j\left(e,\mathbf{s}_j^e\right)d\mu_j^e+\iota\cdot\sum_{e=COL,j\geq j^{HSG}+1}^{j^{COL}} I_{\{a_j<-\underline{b}\}}\left(a_j\left(e,\mathbf{s}_j^e\right)+\underline{b}\right)d\mu_j^e$$

7. The government budget constraint holds

$$\begin{aligned} G+\sum_e p^e\sum_{j=j^{RET}}^{J}\int_{S_j^e} d\mu_j^e+E &= \tau_c\sum_{e,j}\int_{S_j^e} c_j\left(e,\mathbf{s}_j^e\right)d\mu_j^e \\ &\quad+\tau_w\sum_{e,j>j^e} w^e\varepsilon_j^e\int_{S_j^e}\left[1-l\left(e,\mathbf{s}_j^e\right)\right]d\mu_j^e+\tau_k r^+K \end{aligned}$$

where E are net government expenditures in education. Let $\hat{a}_j=\max\{a_j,-\underline{b}\}$. Then:

$$\begin{aligned} E &= -\sum_{j\geq j^{HSG}+1}^{j^{COL}}\left[\int_{S_j^{COL}}\hat{a}_{j+1}\left(\mathbf{s}_j^{COL}\right)I_{\{a_{j+1}<0\leq a_j\}}+\left(\hat{a}_{j+1}\left(\mathbf{s}_j^{COL}\right)-\hat{a}_j\left(\mathbf{s}_j^{COL}\right)\right)I_{\{a_{j+1}<a_j<0\}}\right]d\mu_j^{COL} \\ &\quad+\sum_{j\geq j^{HSG}+1}^{j^{COL}}\int_{S_j^{COL}} g\left(a_j,\theta\right)d\mu_j^{COL}-\pi\sum_{j\geq j^{COL}+1} I_{\{n>0\}}d\mu_j^{COL} \end{aligned}$$

The government has two sources of expenditures for education and one source of revenues. First, it offers means-tested grants of size $g\left(a_j,\theta\right)$. Second, it extends credit to needy students up to $\underline{a}^{GOV}$ without requiring any interest payment. Finally, the government receives payments π from all those still with educational debt $(n_j>0)$.

8. Individual and aggregate behaviors are consistent: the vector of measures $\mu=\left\{\mu_1^{LHS},...,\mu_J^{LHS};\mu_2^{HSG},...,\mu_J^{HSG};\mu_{j^{HSG}+3}^{COL},...,\mu_J^{COL}\right\}$ is the fixed point of $\mu(S)=Q(S,\mu)$ where $Q(S,\cdot)$ is a transition function generated by the individual decision rules, the exogenous laws of motion for the shocks $\{z_j\}$, and the survival rates $\{\zeta_j\}$. And S is the generic subset of the Borel-sigma algebra $\mathcal{B}_S$ defined over the state space $\mathbf{S}$, the Cartesian product of all S_j^e.

3 Parameterization of the model

In this section we describe in detail how we parameterize the model economy. Some of the parameters are calibrated using the model, while others are estimated directly from data using model restrictions.

3.1 Demographics and preferences

Individuals are assumed to be born at the real age of 16, and they can live a maximum of $J = 99$ years, after which death is certain. Retirement occurs at the real age of 65 (model age of 50). There is no mortality risk before retirement age.

We specify the (period) utility function as a CRRA of the following type

$$\begin{aligned} u\left(c_j, l_j \mid d^e = 0\right) &= \frac{\left(c_j^{\nu} l_j^{1-\nu}\right)^{1-\gamma}}{1-\gamma} \\ u\left(c_j, \bar{l} \mid d^e = 1\right) &= \frac{\left(c_j^{\nu} \bar{l}_j^{1-\nu}\right)^{1-\gamma}}{1-\gamma} + \kappa\left(\theta\right) \end{aligned} \tag{10}$$

For the preference parameters, we rely on existing Euler equation estimates, as well as on matching aggregate labor supply levels and education enrolment rates in different ability groups. The parameters ν and γ of the period utility jointly pin down the inter-temporal elasticity of substitution of consumption $\frac{1}{1-\nu(1-\gamma)}$ (ISE) as well as the level of labor supply over the life cycle. We set the ISE to 0.75 as in Blundell, Browning, and Meghir (1994) and Attanasio and Weber (1993). The weight of leisure in preferences, ν, is chosen to match labor supply intensity of workers and is set to 0.33 (see Ríos-Rull (1993)). Hence a value of $\gamma = 2.00$ is chosen to match the inter-temporal elasticity of substitution.

3.2 Education cost parameters

Our main source of information on education costs and funding opportunities in the US are figures published by the National Center for Education Statistics The direct cost of education ϕ is meant to represent tuition and fees in 4-year colleges (we consider only public universities and private, non-for profits universities which cater to populations mostly below age 24. Private for-profit institutions are mostly involved in adult education).

We set annual monetary cost of education (tuition costs plus provisions of academic materials) to be 30% of the median income in the economy, which corresponds to an estimate of the long-term average costs for public and private colleges in the US. We are aware that tuition costs have been soaring in the past decades and we use a long term average to pin down the direct cost of schooling which is consistent with the long-term real costs in the 30 years post 1970s. [11]

The two major sources of financial aid awarded to students are *grants* and *loans*. They are cumulable. Grants are awarded by the federal government, states and institutions. Loans are almost entirely administered by the federal government. The amount of aid received is increasing in the price of attendance and decreasing in family income. These two patterns reflect the need-based formula used to award financial aid in the vast majority of aid programs.

The average tuition fees paid are not very related to family income of the student. For example, an annual family income in the bracket \$20,000-\$40,000 implies average tuition costs of \$4,000 (\$15,000) at public (private) colleges; on the other hand, an average family income between \$80,000 and \$100,000 implies average tuition costs of \$4,400 (\$17,000) at public (private) institutions.

Tuition subsidies $g(a_j, \theta)$ often depend on ability and family income. Roughly half (44.4%) of students receive a grant in each given year. Taking into account all grant types, the average grant size stands at roughly 90% of average tuition costs.

In the numerical section we design alternative experiments which describe different types of subsidization of post-secondary education. We also plan to build counterfactual scenarios based on different assumptions on direct costs which are useful to understand the differences between higher subsidies vis-a-vis lower direct costs.

[11] The cost of attending university includes both tuition and fees and non-tuition expenses like room, board and other supplies. However non tuition expenses can be mostly considered as regular consumption so we do not include them in the direct cost of schooling. Tuition costs vary in private and public institution: for example, in the year 2000 the tuition and fees costs in private institutions were \$15,000 versus \$4,300 in private institutions. In the same year, roughly 2/3 of the students were attending public institution. The average cost of tuition in that year is \$7,900. Source: "Education Digest" and "Student Financing of Undergraduate Education:1999-2000", published by the National Centre for Education Statistics (NCES), which provides a wealth of information regarding both costs of and financial aid towards post-secondary education in the United States.Information about Federal aid programs can also be found in the 'Guide to US Department of Education Programs'.

3.3 Ability gradient, age profiles and labor efficiency shocks

An important characteristic of the model is that the three types of human capital represent different inputs to the production function, not necessarily substitutable.[12] They may have relative prices that vary over time in response to changes in either supply or demand for skills. In particular, supply of skills does not only depend on the number of people with a certain education level but also on their relative efficiency. So as to be able to simulate our model, we need to quantify the effects of individual heterogeneity on observed productivity, as reflected into wages, and, more broadly, on the aggregate supply of human capital. Heterogeneity includes ability as well as the stochastic process of labor-efficiency shocks.

We start by specifying an education specific wage equation. For individual i with education e, the wage rate in period t is denoted as w_{it}^e,

$$\ln w_{it}^e = \log\left(w_t^e\right) + \lambda^e \ln\left(\theta_i\right) + \xi^e\left(age_{it}\right) + u_{it}^e \tag{11}$$

where $\log\left(w_t^e\right)$ represents the log of the marginal product of one efficiency unit of human capital of education-type e; θ_i and λ^{edu} denote, respectively, permanent individual heterogeneity and its gradient, and $\xi^e\left(age_{it}\right)$ is an education specific age-profile for wages.[13]

The unobservable shock u_{it}^e can be specified as the sum of two independent components

$$u_{it}^e = z_{it}^e + m_{it} \tag{12}$$

where z_{it}^e is a (persistent) shock, assumed to have an AR(1) structure

$$\begin{aligned} z_{it}^e &= \rho z_{it-1}^e + \varpi_{it}^e \\ \varpi_{it}^e &\sim N\left(0, \sigma_{\varpi}^e\right) \end{aligned}$$

and m_{it} is *i.i.d.* measurement error (a transitory shock). The persistent z_{it}^e shock is observed before making any consumption or education choices. The decomposition of the

[12] The degree of substitutability is important in determining the size of the G.E. effects. We estimate the elasticity of substitution between labor types using CPS data and we also experiment with alternative specifications in the simulations. A more detailed discussion is presented in the section about identification and estimation of aggregate technology parameters.

[13] We assume that $\xi^e\left(age_{it}\right)$ is a polynomial in age of order 4, that is

$$\xi^e\left(age_{it}\right) = \sum_{x=1}^{4} \alpha_l^e age_{it}^x$$

unobserved heterogeneity term u_{it}^e does not include a permanent shock because we assume that all permanent heterogeneity is captured by θ_i. Self-selection based on permanent heterogeneity (and, to a smaller extent, persistent heterogeneity) impacts on both education decisions and observed wage rates. However, under our shock-structure assumption (12), a within-groups estimator will be sufficient to control for any self-selection associated to fixed effects. Moreover, if one estimates wage equations from individual panel data sets, selection bias attributable to persistent shocks becomes less severe.[14] Another important issue in estimating the wage equations relates to finding a satisfactory way to approximate permanent heterogeneity θ_i: an appropriate data set has to provide panel observations on individual wages and a measure of permanent heterogeneity (ability) which has a measurable impact on wages. The NLSY79 has both these characteristics, as it provides different measures of individual wages and earnings, as well as information about the AFQT (Armed Forces Qualification Test) of most sample members. The AFQT is a test score derived from the combination of different psycho-metric scores (see Appendix for details). We use NLSY79 data to estimate education-specific wage equations like (11): however, this data set provides observations only for workers between age 14 and 45, which makes it hard to identify the whole span of the age-earning profiles. Therefore we use wage data from the PSID to estimate age polynomials for different education groups: the age profiles are then used to filter out age effects from the wage observations in the NLSY79.

3.3.1 Estimating age effects from PSID wage data

The PSID provides information on earnings and hours worked for workers aged 18 to 65. We use this data to estimate age-earning profiles for different education groups. All the waves of the survey between 1968 and 2001 are included. We estimate fourth-degree age polynomials for different education groups and residually generate wage series that are free of age effects. We also provide estimates for a pooled age-earning profile, based on all education groups

Details about our sample selection are reported in section A of the appendix , with

[14] The issue of selection bias ensuing from persistent shocks is related to the so-called "incidental parameters problem" discussed in Heckman (1981). The severity of the incidental parameters problem becomes smaller as the number of panel observation for each given individual in a sample increases.

the estimated age polynomials. We only use data from workers who appear in at least 8 waves.

3.3.2 Ability gradients estimates from the NLSY79

We use data from the NLSY79 to estimate gradients of ability on individual wages. We filter out age effects on wages by using polynomials estimated from PSID data.[15] Denoting the age-free wages as $\tilde{w}_{it}^{edu}$, we are left with the following wage components

$$\ln \tilde{w}_{it}^{e} = \log\left(w_t^e\right) + \lambda^e \ln\left(\theta_i\right) + u_{it}^e$$

Conditioning on education and assuming that the unobserved error term is uncorrelated with θ_i, we can identify the parameters λ^e, $e \in \{1,2,3\}$, by running simple OLS regressions. We use AFQT89 as a measure for θ_i and provide results for different wage measures available in the NLSY79.

We estimate the above equation for the cross-sectional representative sample as well as the full sample of people in the NLSY79, which includes oversamples for minorities and disadvantaged groups. The sample selection and results are reported in section C of the appendix and are presented both by education groups and for the pooled group. Results for the raw, unfiltered wages are also presented. The estimated ability gradient does not change dramatically when we do not purge out age effects.

3.3.3 Labor efficiency shocks

We use the residuals from the wage equations to analyze the stochastic component of wages.[16] First note that, after estimating wage equations, we can observe the following residual:

$$u_{it}^e = \ln \tilde{w}_{it}^e - \log\left(\widehat{w}_t^e\right) + \hat{\lambda}^{edu} \ln\left(\theta_i\right) \tag{13}$$

We assume that u_{it}^e can be decomposed into two components

$$u_{it}^e = z_{it}^e + m_{it}^e$$

[15]We set the intercept of the age polynomial to zero. This is a normalization on marginal products of human capital.

[16]For a review of the relevant literature on wages uncertainty and labor supply see, among others, MaCurdy (1981), Abowd and Card (1989) and **?**

where z_{it}^e is an autocorrelated error process and $m_{it}^e \sim iid\,(0, \sigma_m^e)$ is a transitory shock (interpreted as classical measurement error). We assume that $\{z_i^e\}$ is an AR(1) process with education specific parameters of the following type

$$z_{it}^e = \rho^e z_{it-1}^e + \varpi_{it}^e$$
$$\varpi_{it}^e \sim iid\,(0, \sigma_\varpi^e)$$

We use a Minimum Distance Estimator (MDE), (see Chamberlain (1984) and Heathcote, Storesletten, and Violante (2004), to estimate the basic parameters of both persistent and transitory shocks for each education group. Table (18) in section C of the appendix reports estimates of the year-specific variance of both transitory and persistent shocks to wages, as well as estimates of the autoregressive coefficients ρ^e and of the initial condition for the variance of the persistent shocks. The estimates are based on on the CPS-type wages reported in the NLSY79 panel.

3.4 Permanent heterogeneity

In what follows we use the term 'ability'to describe a set of permanent characteristics which affect lifetime earnings as well as education attainment. For the purpose of measuring the distribution of ability over the population we use NLSY data. The NLSY79 provides IQ test scores for both mothers and children: by linking children's measures of ability to their mothers', one can estimate ability transition matrices.

Moreover, the NLSY test scores can be linked to wage data, so to quantify the effect of measured ability on lifetime earnings. Finally, the NLSY also allows us to measure education enrolment rates in different ability groups, which we use for our calibration.

3.4.1 Measuring mother-to-child ability transition

Using the "Children of the NLSY79" survey, we build pairs of mother and child test-score measurements. For mothers we use AFQT89 measurements whereas for children we choose the PIAT Math test-scores.[17] Mothers and children are ranked using their test

[17] No AFQT measure is available for children, and the"piat_math" is considered to be the most accurate measure of future ability among the available test-scores. In some cases the test was administered at different ages to the same child, so that different measurements are available: in these cases we use the latest available measurement as we wish to approximate the distribution of ability at age 16.

scores and then split into "bins" corresponding to different quintile groups.[18] We compute a 'quintile-transition' matrix, which assigns a probability to the event that a child ends in a given ability group, given the observed ability rank of the parent. More details about the procedure used to compute the ability-transition matrix can be found in section C of the appendix. The estimated ability transition matrix for a 5-bins decomposition is reported in table (13).[19]

3.4.2 Approximating the stationary distribution of ability

The ability transition matrix describes a mapping from maternal quintiles to children quintiles. However, we also need to approximate an equilibrium distribution of ability which takes values over some given test-scores range. We use the distribution of normalized AFQT89 (in logs) from the whole cross-sectional sample of the NLSY to approximate the quantiles of the unconditional distribution of ability in the population. This also helps to relate ability to earnings, as the normalized logs of the AFQT89 are also used to estimate ability gradients in wage equations.

Tables $(16 - 15)$ in section C of the appendix document some facts about the distribution of AFQT89 over the subsample of mothers we use in the analysis of ability transition as well as over the whole cross-sectional sample of the NLSY79. There is very little difference in the distribution of AFQT test-scores over these two samples.

We also compute education enrolment rates for different quintiles of the ability distribution, which we use to calibrate the relative supply of different types of human capital in the economy.

3.5 Using CPS data to measure aggregate human capital inputs

Estimation of the aggregate production function requires the total wage bills for each of the education groups. In the general CES case we also need measures of human capital in each of these groups. To this purpose we use the March supplement of the Current

[18]The percentiles used to rank mothers and children are based on the sample populations. We estimate transition matrices based on 5 bins decompositions, as well as decompositions with 10 ability bins.

[19]The estimated transition matrix does not assign exactly 1/5 of the children to each quintile of the children distribution because of clustering of observations around quantile values. In the numerical work we rescale the transition probabilities to keep quintile sizes constant over successive iterations, which is necessary for a stationary distribution.

Population Survey (CPS). The wage bills are straightforward to obtain. We just add up the earnings of each of the three education groups and then scale up the figures to match the entire US economy. When we need to estimate a CES production function the issue is more involved because we also need to estimate the *quantity of human capital* in each year. To achieve this we need an aggregate price series for each of the education groups; our estimates from the PSID provide measures of price growth over time, but a normalization assumption on each price is necessary. Any normalization will correspond to a set of relative prices at a given point in time. However, we still have one degree of freedom: in fact, after setting the initial relative price of high school and of college graduate labor we can choose the utility costs of education to match the proportions going into each of the educational categories. In other words with unobserved costs the data can be rationalized either with high returns and high costs or low returns and low costs. The particular normalization we choose will not affect the simulation of the policy changes. Given a series of log prices for different labor types, it is possible to divide the wage bills by the exponentiated value of such prices to obtain point estimates of the value of efficiency weighted total labor supply (human capital aggregates) by education and year.

3.6 Aggregate Production Function

The relationship between human and physical capital is expressed as a Cobb-Douglas

$$Y = F(K, \mathcal{H}) = MK^{\phi}H^{1-\phi} \tag{14}$$

where the factor M is a TFP coefficient. Aggregate human capital stock $\mathcal{H}$ is the product of a CES aggregator

$$\mathcal{H} = (s_{1t}H_1^{\rho} + s_{2t}H_2^{\rho} + s_{3t}H_3^{\rho})^{\frac{1}{\rho}} \tag{15}$$

where H_e is the stock of human capital associated with education level edu and $s_{3t} = (1 - s_{1t} - s_{2t})$. The equilibrium conditions require that marginal products of human capital (MPHC) equal pre-tax prices, so that $w^e = MPHC^e = \frac{\partial F}{\partial H_{edu}}$ for any education level e, and $r + \delta = \frac{\partial F}{\partial K}$.

¿From the iso-elastic CES specification for the human capital aggregate in equation (15) we can derive log-linearized expressions for the wage bills . For education groups j

and i, for example, we can write

$$\log\left(WB_t^j/WB_t^i\right) = \log\left(\frac{s_{jt}}{s_{it}}\right) + \rho\log\left(\frac{H_{jt}}{H_{it}}\right)$$

where WB_t^j and H_{jt} denote total wage bill and aggregate human capital for education group j in year t. Given the strong shifts in education enrolment and wage bills over the period considered, we express share parameters s_{jt} $(j = 1, 2, 3)$ as the product of a constant and a time-varying component: $s_{jt} = s_j \exp\{g_j t\}$, where t denotes calendar year and g_j captures the change in each human capital share j over time. The log-linearized equation, for arbitrary education groups j and i, can be written as

$$\log\left(WB_t^j/WB_t^i\right) = \log\left(\frac{s_{jt}}{s_{it}}\right) + \log\left(\frac{g_j}{g_i}\right)t + \rho\log\left(\frac{H_{jt}}{H_{it}}\right) \tag{16}$$

We use equation (16) to identify the ratio of share parameters in 1968 $\left(\frac{s_{jt}}{s_{it}}\right)$, their rates of growth in every subsequent year $\left(\frac{g_j}{g_i}\right)$ and the elasticity of substitution between human capital inputs, (ρ).[20] The estimate value for ρ ranges between 0.36 and 0.68, which corresponds to an elasticity of substitution between 1.6 and 3.1. Using a simple skilled/unskilled classification Katz and Murphy estimate the elasticity of substitution in production to be 1.41, while Heckman, Lochner, and Taber (1998a) report a favorite estimate of the elasticity of substitution between skilled and unskilled equal to 1.44; Johnson (1970) has an old estimate equal to 1.50. Card and Lemieux (2001) find that the elasticity of substitution between different age groups is large but finite (around 5) while the elasticity of substitution between college and high school workers is about 2.5. Notice that

[20] In equation (16) the ratio $\log\left(\frac{s_{jt}}{s_{it}}\right)$ contains information about the intercept of the age polynomial $\frac{\alpha_0^2}{\alpha_0^1}$ as defined in equation (11): the amount of log hourly wage that is attributed to marginal product of labor cannot be distinguished from the amount attributed to a constant component of the age polynomial. This can be seen by way of example. We know that where $\frac{H_2}{H_1} = \frac{\exp\left(\alpha_0^2\right)}{\exp\left(\alpha_0^1\right)}\frac{\tilde{H}_2}{\tilde{H}_1}$ where $\frac{\tilde{H}_2}{\tilde{H}_1}$ are the ratio of human capital aggregates obtained under the assumption that the education specific α_0^1 and α_0^2 are equal to zero. We can rewrite equation (16) as

$$\log\left(WB^2/WB^1\right) = \left[\log\left(\frac{s_2}{s_1}\right) + \rho\log\frac{\exp\left(\alpha_0^2\right)}{\exp\left(\alpha_0^1\right)}\right] + \log\left(\frac{g_2}{g_1}\right)t + \rho\log\left(\frac{\tilde{H}_2}{\tilde{H}_1}\right) \tag{17}$$

The ρ parameter in equation is identified under any rescaling of the ratio $\frac{H_2}{H_1}$, because the log transformation isolates rescaling factors in the constant term. This means that the estimation of ρ does not change with alternative price normalizations. A normalization is necessary on the *ratio* of aggregate human capital types: we assume that all the constants in the age polynomials are zero, which is consistent with the procedure we adopted to purge out age effect from PSID data.

our elasticity estimates provide a measure of substitutability between 3 different types of workers, rather than two simple skill groups.

3.7 Inter-vivos transfers of resources

The NLSY97 provides information regarding family transfers received by young individuals. In particular, it asks respondents about any gifts in the form of cash or a check (not including any loans) from parents. Given the length of the sample we can observe such transfers for people between the age of 16 and 22. We use this information to evaluate the relative size of early transfers, which are relevant for education financing, as a share of available measures of family income and wealth. Section (D) of the appendix describes the sample we use to measure early inter-vivos transfers and summarizes the basic facts about parental gifts to young individuals, as recorded in the NLSY97.

Since we model early inter-vivos transfers as a one-off, lump sum gift from parents to their child, we are interested in the total monetary transfer between age 16 and 22. We find that the average transfer is $1,531 per year, which sums up to an average total transfer of $10,717 per youth between age 16 and 22. The median transfer per year is $486 and corresponds to a median total transfer between age 16 and 22 of $3402. When we condition on the highest education between residential parents, the total mean transfer for households in the highest education group is $14,105 (median $3,500), versus a mean of $7,714 (median $3,353) for households of High School graduates and a mean of $6,510 (median $2,800) for households with guardians who are all drop-outs.

4 Simulations - Preliminary[21]

This section describes the benchmark economy and presents the results of our policy experiments. We start by describing the main features of our benchmark economy in some detail. All results are reported in normalized model units, rather than dollars (however an easy translation into dollars can be obtained by multiplying values by a factor of 10,000).

[21] The following numerical results are based on a version of the model in which annuity markets are absent and the initial wealth distribution is given by the equilibrium distribution of accidental bequests. Results for the alternative specification with intervivos transfers and annuity markets will be added when available.

The numerical results are preliminary insofar they refer to the case in which no inter-vivos transfers take place: instead, the distribution of initial wealth among youths is obtained from the distribution of accidental bequests. In order to evaluate the effects of possible correlation between initial wealth and ability, we use copulas to parameterize a degree of correlation based on recent research using the NLSY79 by Zagorsky (2007).

We are currently working on simulations which incorporate parental altruism as the source of initial wealth in life, as described in the model. All along we compare general and partial equilibrium effects of education policies: there are several alternative ways to interpret the differences between the two cases. One might, for example, think of partial equilibrium outcomes as the results of a small pilot-run of a possibly larger intervention (a local experiment, limited to a province or city) or as the outcomes which would be observed in the short-term, while the necessary adjustments for price-effects take place. One might also think of partial equilibrium outcomes as being the outcomes for economies with wage rigidities or in the case of small-open economies which take wage rates as given. However, one striking result of our experiments is that relatively small percentage changes in the returns to HC (in the order of less than 1%) are sufficient to almost cancel out the effects of a subsidy. This result is largely due to selection based on ability: even very small price adjustment can induce a crowding out of low ability types by high ability types (not unlike some of the results documented in the literature on entrepreneurship, where better ability matches are developed when credit constraints are released, inducing higher efficiency of allocations).

4.1 Calibration of the benchmark model

Not all the parameters in our model are estimated: some parameters are in fact chosen with the objective to build a numerical counterpart of our model which is able to reproduce some relevant features of the US economy.

Wealth plays a pivotal role in determining equilibrium outcomes. The availability of assets and access to credit to smooth consumption is a crucial factor for education decisions. We set time-preference and borrowing limit parameters in order to obtain a benchmark with a realistic distribution of wealth: in particular, we want our model to generate a share of wealth-poor people in line with the observed share for the target year. The distribution of workers over education outcomes is equally important, because it determines the relative returns to the education investments. However, the aggregate education shares are not sufficient by themselves to pin down relative returns because the relative ability of workers is key in determining aggregate human capital inputs in the production function. Therefore we target not only the aggregate education shares in the target year, but also education shares by ability. The additional benefit of this calibration approach is that we are able to assess the composition effects of potential policies by looking at selection over ability as well as wealth.

The remainder of this section describes our calibration approach in more detail.

Demographics. Each period represents one full year. An individual is born at age 16. After retirement there is an age-related probability to die in each period that we take from the US life tables for 1989-1991.

Discount factor. The discount factor β is chosen to produce a wealth income ratio equal to that for US households up to the 99% percentile. (Wolff 2000) estimates the value of this ratio to be roughly 3.5 in 1983. The implied value of the discount factor is 0.962.

Wealth distribution of the youngest. We assume that the wealth distribution among the youngest corresponds to the distribution of the accidental bequests in the economy. However, no agent is endowed with negative assets, so we censor the bequests' distribution at zero and appropriately modify the average bequest so that the total bequeathed wealth is held constant. We calibrate the benchmark under two alternative assumptions on the degree of correlation between initial assets and ability: either we assume no correlation or we assume a 10% correlation, as documented by Zagorsky (2007). We replicate identical policy experiments on the benchmark economy under each of the two assumptions.

Borrowing Limit. The exogenous borrowing limit $\underline{a}^{PRV}$ for private loans is calibrated to match the share of workers (all agents excluding students) with zero or negative wealth. (Wolff 2000) provides an estimate of 15.5% for this share, which implies a borrowing limit of about 40% of median post-tax income (sum of earned and capital income). We assume no difference between public and private loans at this stage and we assume students have access to the same credit markets as working adults. In the future we will differentiate between private and public provision of credit.

Government. We use flat tax rates for both labor and capital income and, following (Domeij and Heathcote 2003), we set $t_l = 0.27$ and $t_k = 0.4$. For simplicity, the pension is assumed to be a constant lump sum for all agents, regardless of their education and previous earnings. The replacement rate for the lump-sum is set to 16.4% of average post-tax labor earnings like in (Heathcote, Storesletten, and Violante 2004).

Distribution of permanent characteristic (ability). We use the distribution of AFQT89 scores over the whole cross-sectional sample of the NLSY79 (for which we have computed wage gradients). For expositional simplicity we split the range of ability in 5 quintiles. Such ability bins are used to characterize policy effects on different agents in the ability distribution.

Direct Cost of Education. The direct cost of college education is chosen to match the value of tuition costs as a proportion of average pre-tax earning. The National Center for Education Statistics provides several measures of tuition costs and we use our PSID sample for an estimate of average pre-tax earnings. Over the sample period the real college tuition costs have been been steadily growing, increasing from less than 5% to over 15% of our selected measure of earnings. We choose to set the college tuition costs to be 10% of average *post-tax* earnings. Given the labor tax rate in our model, this is equivalent to a college tuition cost roughly equal to 8% of average *pre-tax* earnings. For the value of High School direct costs we have set them to be just 1% of average post-tax earnings, in order to account for expenses incurred for studying equipment and other necessities. There does not seem to be not much information on such costs.

Education Shares among Workers. Education rates are matched both in the aggregate and by ability groups. The distinction is important because the same aggregate shares are consistent with many different distributions of ability over education and, therefore, many different relative marginal returns between different types of labor. Moreover, the policy experiments are likely to alter the distribution of ability in each education group and it is desirable that the benchmark reproduce the distribution of ability types over education outcomes. In order to approximate such distribution we use information from the NLSY79 which provides data on educational attainment of agents as well as their score in the AFQT test. We assign people to 5 different ability bins, with bin 1 comprising those with the lowest IQ scores and bin 5 those with the highest. The education shares for each ability bin are reported in table (17). However, the aggregate education shares based on the NLSY do not represent the true shares of aggregate enrolment in the US economy in our sample period.[22] In order to reproduce the aggregate education distribution in the economy we gross-up (by the same proportion) the ability-specific rates so that their aggregation gives back the average overall education rates for workers in the US economy between 1977 and 1983[23]. In aggregate, the average fraction of workers with no High School degree is 0.25. The fraction of High School graduates is 0.576 and the College graduate share is 0.174.We use ability-specific quasi-linear utility terms $\kappa(\theta)$ to shift the value of education for different ability bins and match the education shares.

The value of the parameters calibrated in the benchmark are reported in table (25) in section E of the appendix . The characteristics of benchmark economy are reported in table (26). We report the results for the economy with zero correlation between ability and initial wealth because they are almost identical to the ones for the economy with positive (10%) ability-wealth correlation. However, as we will see, the ability-wealth correlation does play a role when running policy experiments. In summary the benchmark economy has the following characteristics:

- the marginal return to an efficiency unit of College labor is 25% higher than to an efficiency unit of HS labor. Similarly the marginal return to HS labor is 26% higher than the marginal return to the labor of HS drop-outs;
- after accounting for selection and relative efficiency of different workers, hourly wages of HS graduates are 57% larger than HS drop-outs wages. College wages are 60% larger than High School wages;
- income inequality is only slightly larger than wage inequality for College graduates vis-a-vis HS graduates;
- overall, the patterns of inequality in labor and income are roughly consistent with the averages over the sample period considered.

[22] One reason for this problem is attrition which can unequally affect people with different education in the NLSY, altering the aggregate education shares. Moreover, our sampling procedure is likely to exclude some workers.

[23] We use this time period because we calibrate the other parameters circa 1980. Moreover the education shares lie very close to the sample averages for the period 1967-2001.

4.2 Experimenting with conditional subsidies

In what follows we report the results of different counterfactual policy experiment based on the introduction of subsidies: we consider subsidies conditional on current wealth of potential students (a proxy for family income) as well as subsidies conditional on ability. We compare the outcomes of education policies in partial equilibrium (when the returns to HC are held constant to their benchmark level) vis-a-vis their general equilibrium outcomes (when HC returns are allowed to adjust). All along we maintain that the elasticity of substitution among HC types is 3.1, which corresponds to the highest (and least favorable to price-effects) of our estimates. Moreover, we also compare the effects of policies in the case in which we allow for positive correlation between initial wealth endowments and ability.

The annual grant size, conditional on the award of a subsidy, is set to 94.5% of the annual monetary cost of College, based on information from the NCES. Students receive the grant in each year of their curriculum. To provide a measure of the cost of the policy we assume that all subsidies are financed through changes in the labor earnings or capital income tax (if efficiency gains and tax receipts are large enough a subsidy can pay for itself and even induce a reduction in tax rates).

4.2.1 Wealth-dependent grants

The first set of experiments consider the introduction of a grant for which only wealth-poor youths are eligible: eligibility is restricted to youths with assets 50% (or less) the median value of assets in the benchmark economy.[24] Subsidy expenditure is financed through changes in the labor income tax rate.

Table (27) documents a set of outcomes for the case of zero ability-endowment correlation as well as for positive ability-endowment correlation.

Case 1: zero correlation between initial asset endowment and ability. The partial and general equilibrium outcomes of this conditional subsidy are strikingly different; in partial equilibrium the subsidy increases output by 2.5% partly thanks to a large increase in skilled labor: college graduates account for over 27% of total workers

[24] Since median wealth endowments for youths are 65% of median wealth in the benchmark economy, this subsidy rule implies that only youths with wealth equal or below of the 77th percentile in their cohort are eligible for the subsidy.

and aggregate HC of college type almost doubles. This is accompanied by a reduction in inequality, with the college premium (vis-a-vis High School) shrinking to 55% from the original 60%. The subsidy more than pays for itself, generating a drop in the labor tax rate to a marginal of 26.34%. Roughly 60% of University students receive the subsidy in the P.E. case. These results are not robust to changes in the marginal returns to HC: in general equilibrium the subsidy generates an increase in the marginal return to HS human capital of 2% and a drop in the marginal return to University-level human capital of 1.6%. These apparently small changes are sufficient to undo the partial equilibrium results: the share of college-level workers goes back down to 18% and aggregate HC of the same workers almost reverts to its benchmark level. The change in wage inequality is greatly undone as well, as the college-high school premium goes back to 58%. The tax rate associated to the subsidy in G.E. is almost unchanged with respect to the benchmark, at 26.92%, as is the aggregate output which goes up only by 0.7%. The marginal improvement in production efficiency is attributable to the changes in the education distribution by ability: although the aggregate education shares are almost identical to the benchmark, the composition by ability is very different. the subsidy originally induces more people to acquire education and, when marginal returns adjust, the first people to find education unprofitable are the lower ability ones. This results in a much larger number of high ability workers in high education jobs: the matching of ability and education is positive in terms of efficiency.

Case 2: positive (10%) correlation between initial asset endowment and ability. As it might be expected, the effect of a (means-tested) subsidy is weaker in this case because the financing constraints to high ability individuals are less binding. In this sense both the partial and general equilibrium outcomes are closer to the benchmark. In the P.E. case about 55% of University students receive the subsidy,. the output gain is just above 2% and the share of college-educated workers goes up by less than in 'Case 1' above. However we also notice that the GE effects in this case completely undo the efficiency gains observed in P.E.: output is just .2% larger than in the benchmark, and taxes are larger than the benchmark 27% (at 27.11%). There is less scope for matching good workers to high-skill jobs, as part of the efficiency gain follows naturally from the fact that higher ability workers are born with more resources and are less constrained in their education choices. This implies that the subsidy has virtually no effect in G.E..

4.2.2 Ability-dependent grants

We now turn to subsidies which are conditional on the student having at least a minimum level of ability. We consider grants for which only students above the 80th percentile level of ability are eligible. This restriction generates the same share of grant recipients as in the wealth-dependent subsidy case, which makes comparison easier. Also in this case the financing of subsidies comes from changes in labor tax rate. Table (29) describes the basic facts about the effects of this subsidy structure.

Case 1: zero correlation between initial asset endowment and ability. The share of students benefitting from the grant in partial equilibrium is around 60% (similar to the wealth-based subsidy example). This subsidy increases the share of college-educated worked to 21% in P.E. which is well below the wealth-dependent subsidy. This is partly because more able agents were already self-selecting themselves into higher education before the transfer and the number of marginal individuals who react to the subsidy is relatively smaller. For the same reasons this type of subsidy has a smaller effect on aggregate output: in P.E. the subsidy generates only a 1.4% increase in output compared to the benchmark. This policy substantially increases wage inequality: the college premium jumps to 62% in partial equilibrium, as only more able people are eligible for the subsidy and less able agents are left behind. This selection exacerbates inequality. The labor tax rate goes down to 26.76%. In G.E. the share of college-educated workers goes back to its benchmark level of 17%, and the increase in aggregate output drops to 0.6%. The tax rate is virtually unchanged at 26.97%. However the selection of more able people into education is reflected in a shift of the composition of college-workers towards the highest ability group. This is reflected in the college premium which jumps to 64%, even higher than the benchmark. .

Case 2: positive (10%) correlation between initial asset endowment and ability. The share of college students receiving a grant is around 60% and roughly comparable to the share receiving the asset-dependent grant. The effects of the subsidy are very similar to 'Case 1' above: the effects on aggregate output are comparatively smaller, but the magnifying effect on inequality is even larger with the college wage-premium jumping to 65% (vis-a-vis High School wages). This policy, together with the positive correlation between endowments and ability, pushes up inequality with very little gain in efficiency.

4.3 Experimenting with changes in marginal tax rates

A possible alternative to subsidizing education achievement could be to reduce the taxation of the returns to human capital, namely labor income taxes. Intuitively this channel sounds promising, as it positively affects returns to HC by reducing a distortionary tax. In order to test this hypothesis we design an experiment in which the marginal tax rate on wages is reduced to 25%. This reduction is financed by increasing the marginal tax rate on capital until the government budget is fully balanced: this is likely to reduce the incentive to accumulate physical capital and might have a detrimental effect on output. The experiment is run only under the G.E. specification, as it would be hard to imagine a case in which such policy could hold as a P.E. case. In order to explore alternative avenues we also run the opposite experiment: that is, we reduce the capital tax rate to 38% and finance the ensuing drop in tax revenues by pushing up labor income taxes.

4.3.1 Lowering the labor income tax rate

We perturb the benchmark equilibrium by reducing the marginal tax rate on labor income to 25%. This reduction is financed by adjusting the capital tax rate.

Case 1: zero correlation between initial asset endowment and ability. The impact of this reshuffling of the tax code is relatively small (see table 31): there is very little change in aggregate education shares (a marginal increase of college equivalents) but overall output goes down by 1.4% relative to the benchmark. This effect is due to the jump in capital tax rate to 47.64% (from 40% in the benchmark) which reduces the amount of productive capital in equilibrium. This experiment also generates a decrease in wage inequality, with the college premium going down to 57.5%.

Case 2: positive (10%) correlation between initial asset endowment and ability. In this case the effects of the tax changes are more extreme: human capital accumulation actually increased in aggregate (with college-equivalent share going to 18%). Education by ability also changes towards a more even distribution of education outcomes by ability group. Wage inequality drops substantially, with a college wage premium of only 51%. The draw back is an even larger decrease in output, of 1.9% relative to the benchmark. It seems that in this case the change in tax structure is able to reshuffle HC accumulation, with a substantial increase for the lower ability groups. However this reshuffling, while reducing inequality, reduces the match-up between able people and high-skill job. This, together with the drop in productive capital, generates a rather substantial drop in efficiency.

4.3.2 Lowering the capital income tax rate

We also experiment with a decrease of the capital tax rate from 40% to 33%.

Case 1: zero correlation between initial asset endowment and ability. The offsetting increase in labor income tax is small (to 28.68%) and the distribution of education achievement, wage inequality and aggregate human capital are virtually unchanged relative to the benchmark. However output goes up by 1% because of the higher accumulation of productive capital. Overall this tax change is not effective in fostering HC accumulation but is able to increase output efficiency, as we might have expected.

Case 2: positive (10%) correlation between initial asset endowment and ability. The outcome of this experiment are very similar to the case of no correlation between ability and endowments. The effects on HC accumulation are negligible, just like the changes in wage inequality. The only noticeable effect, as before, is the increase in physical capital and aggregate output.

5 Conclusions

We combine estimation and calibration techniques to obtain an overlapping generations general equilibrium model with heterogeneous agents and idiosyncratic uncertainty. Individuals choose education levels, labor supply and consumption within an incomplete markets set-up. We use this model to evaluate alternative educational interventions.

We experiment with two types of tuition subsidies, respectively conditional on wealth and ability. It becomes apparent that while in partial equilibrium such policies can be very effective in increasing education levels, in general equilibrium the results are much less encouraging: the main effect of a subsidy there is to increase the supply of human capital as one would expect. However, it is the more able but liquidity constrained individuals who take up extra education, while the education levels of the less able can actually decrease (they are crowded out). Thus the subsidy acts on the composition of those in education. In many respects this is very much in line with results found by Heckman, Lochner, and Taber (1998a). The inclusion of risky returns on labor earnings, the fact that labor supply is endogenous and the explicit modelling of the wealth distribution of youths lend additional credibility to the result. The wealth-based subsidy seems to be able to generate larger efficiency gains and reduce inequality. The ability-tested subsidy actually increases inequality both in partial and general equilibrium: the effect is particularly strong when we also allow for positive correlation between ability and initial wealth endowments. Overall, wealth-based grants seem to be preferable, even though their general equilibrium impact is still relatively small.

We also experiment with changes in tax rates. First, we reduce the labor income tax rate (which is financed by off-setting changes in capital income tax): in this case we find very little effect on aggregate human capital accumulation but a drop in output-efficiency due to a smaller stock of productive capital. Viceversa, when we reduce the capital income tax to 33% (offsetting through an increase in labor income tax) we find almost no effect on the accumulation of HC but a relatively large (1%) positive effect on output, due to higher accumulation of productive physical capital. It would be interesting to run the same tax experiments using an offsetting consumption tax.

We conclude that equilibrium effects exist, can be triggered by relatively small changes in marginal returns and generally undo the partial equilibrium outcomes. We find that wealth-based subsidies are more effective at increasing HC accumulation and output than ability-based subsidies, both in P.E. and G.E., especially in the presence of positive correlation between initial wealth and ability. Finally we argue that reshuffling the tax burden from labor to capital, or viceversa, has very little effect on the equilibrium level of HC but, as might be expected, strongly affects the stock of productive physical capital and aggregate output.

References

Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

Abraham, A. (2001): "Wage Inequality and Education Policy with Skill-biased Technological Change in OG Setting," Ph.D. thesis, Universitat Pompeu Fabra.

Altonji, J. G., F. Hayashi, and L. J. Kotlikoff (1997): "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, 105(6), 1121–66.

Attanasio, O., and G. Weber (1993): "Consumption Growth, the Interest Rate and Aggregation," *Review of Economic Studies*, 60(3), 631–649.

Becker, G. S. (1964): *Human Capital*. NBER.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), 352–365.

Blinder, A. S., and Y. Weiss (1976): "Human Capital and Labor Supply: A Synthesis," *Journal of Political Economy*, 84(3), 449–472.

Blundell, R., M. Browning, and C. Meghir (1994): "Consumer Demand and the Life-Cycle Allocation of Household Expenditure," *Review of Economic Studies*, 61(1), 57–80.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116(2), 705–746.

Chamberlain, G. (1984): "Panel data," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. D. Intriligator, vol. 2 of *Handbook of Econometrics*, chap. 22, pp. 1247–1318. Elsevier.

Cunha, F., J. Heckman, and S. Navarro (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57(2), 191–261.

De Nardi, M. (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71(3), 743–768.

Domeij, D., and J. Heathcote (2003): "On The Distributional Effects Of Reducing Capital Taxes," Mimeo.

Eaton, J., and H. S. Rosen (1980): "Taxation, Human Capital and Uncertainty," *American Economic Review*, 70(4), 705–715.

Fernandez, R., and R. Rogerson (1995): "On the Political Economy of Education Subsidy," *Review of Economic Studies*, 62, 249–262.

——— (1998): "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review*, 88(4), 813–833.

Gale, W., and J. Scholz (1994): "Intergenerational Transfers and the Accumulation of Wealth," *Journal of Economic Perspectives*, 8(4), 145–160.

Gallipoli, G., and L. Nesheim (2007): "The problem of non convexities in life-cycle models," Mimeo.

Heathcote, J., K. Storesletten, and G. Violante (2004): "The Macroeconomic Implications of Rising Wage Inequality in the US," Mimeo.

Heckman, J. (1981): "The Incidental Parameters Problem and the Problem of Initial Condition in Estimating a Discrete Time-Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data and Econometric Applications*, ed. by C. F. Manski, and D. L. McFadden. MIT Press.

Heckman, J., and P. Carneiro (2003): "Human Capital Policy," NBER Working Paper, 9055.

Heckman, J., L. Lochner, and C. Taber (1998a): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

——— (1998b): "General Equilibrium Treatment Effects: A Study of Tuition Policy," *American Economic Review, Papers and Proceedings*, 88(2), 381–386.

——— (1998c): "Tax Policy and Human Capital Formation," *American Economic Review, Papers and Proceedings*, 88(2), 293–297.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2005): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," NBER Working Papers 11544, National Bureau of Economic Research, Inc.

Hill, B. (1975): "A Simple General Approach to Approximate the Tail of a Distribution," *Annals of statistics*, 3, 1163–1174.

Hyslop, D. (2001): "Rising U.S. Earnings Inequality and Family Labor Supply: the Covariance Structure of Intrafamily Earnings," *American Economic Review*, 91(4), 755–777.

Johnson, G. E. (1970): "The Demand for Labor by Educational Category," *Southern Economic Journal*, 37, 190–204.

Keane, M., and K. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Lee, D. (2005): "An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice," *International Economic Review*, 46, 1–34.

Lee, D., and K. Wolpin (2006): "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, 47, 1–46.

Levhari, D., and Y. Weiss (1974): "The effect of risk on the investment in human capital," *American Economic Review*, 64, 950–63.

MaCurdy, T. E. (1981): "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, 89(6), 1059–85.

Mincer, J. (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66(4), 281–302.

Mincer, J. (1994): "Investment in U.S. Education and Training," Discussion Paper NBER WP4844, NBER, Working Paper Series.

Polivka, A. E. (2000): "Using Earnings Data from the Monthly Current Population Survey," mimeo, Bureau of Labor Statistics.

Ríos-Rull, J.-V. (1993): "Working in the Market, Working at Home and the Acquisition of Skills: A General Equilibrium Approach," *American Economic Review*, 83(4), 893–907.

Rosen, S. (1977): "Human Capital: Relations between Education and Earnings," in *Frontiers of Quantitative Economics*, ed. by M. D. Intriligator, vol. 3b. Amsterdam: North Holland.

West, S. A. (1985): "Estimation of the Mean from Censored Income Data," mimeo.

Willis, R., and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5), s7–s36, pt. 2.

Wolff, E. N. (2000): "Recent trends in wealth ownership, 1983-1998," working paper.

Zagorsky, J. L. (2007): "Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress," *Intelligence*, 35, 489–501.

A PSID Data

The Panel Study of Income Dynamics is a survey of the US population started in 1968 and repeated with annual frequency. Following waves of interviews include only persons present in the prior year, including those who moved out of the original family and set up their own households.[25] Each wave provides information on the previous year. We use data for the waves from 1968 to 2002 (referring to 1967 to 2001). Since 1997 the PSID has become biannual. The PSID contains different samples with unequal probabilities of selection: at the beginning of the PSID (1968) the original Survey of Economic Opportunity (SEO) sample of poor families was combined with a new equal probability national sample of households selected from the Survey Research Center 1960 National The SRC was originally representative of the US population. In 1990 an over-sample of Latino families was added. Similarly, in 1997 and 1999 another over-sample of new immigrant families became part of the study population.

A.1 Sample selection and estimated age profiles

The main earnings' variable in the PSID refers to the head of the household, and is described as total labor income of the head.[26] We use this measure, deflated into 1992 dollars by the CPI-U for all urban consumers. By selecting only heads of household we ignore other potential earners in a family unit and restrict our attention to people with relatively strong attachment to the labor force. We include both men and women as well as whites and non-whites.

Information on the highest grade completed is used to allocate individuals to three education groups: high school drop-outs (LTHS), high school graduates (HSG) and college graduates (CG).

We choose not to use the over-sample of Latino families and new immigrant families. After dropping 10,607 individuals belonging to the Latino sample and 2263 individuals

[25] A distinction between original sample individuals, including their offspring if born into a responding panel family during the course of the study (i.e., both those born to or adopted by a sample individual), and non-sample individuals must be made. Details about the observations on non-sample persons and their associated weights and relevance are included in the appendix.

[26] In the PSID the head of the household is a male whenever there is a cohabiting male/female couple. The earnings variable includes the labor part of both farm and business income, wages, bonuses, overtime, commissions, professional practice and others. Labor earnings data are retrospective, as the questions refer to previous year's earnings, which means that 1968 data refer to 1967 earnings.

Table 1: Distribution of observations for the 1967-2000 PSID sample, by education group

years of education	Number of Individuals	Number of Observations
less than 12	430	6,546
12 to 15	1,792	29,229
16 or more	863	14,945

belonging to the new immigrant families added in 1997 and 1999, the joint 1967-2001 sample contains 50,583 individuals. After selecting only the observations on household heads we are left with 19,583 individuals. Dropping people younger than 25 or older than 65 leaves us with 18,186 people. Dropping the self employment observations leaves 14,866 persons in the sample. We then select only the individuals with at least 8 (possibly non continuous) observations, which further reduces the people in the sample to 6228. Dropping individuals with unclear education records leaves 6,213 people in sample. Disposing of individuals with missing, top-coded or zero earnings reduces the sample to 5,671 individuals and dropping those with zero, missing or more than 5840 annual work hours brings the sample size to 5,660 individuals. We eliminate individuals with outlying earning records, defined as changes in log-earnings larger than 4 or less than -2, which leaves 5,477 individuals in the sample. Finally, dropping people connected with the original SEO sample reduces the number of individuals to 3,085.

The age polynomials are presented in Table (3) for different education groups and the pooled sample.

A.2 Time changing relative labor prices and their normalized level

Equation (11) allows explicitly for time changing labor prices for different education groups, denoted as w_t^e. These can be interpreted as marginal products of different types of efficiency units of labor. Using the wage data directly to estimate the time series of different human capital prices would not take into account changes in ability composition over time. However, we can exploit the fact that ability enters linearly in equation (11) and use first differences of wages to estimate the time series of price growth in each education group. Figure 1 reports the point estimated of price growth by education group (tables of estimates and standard errors are available upon request).

Table 2: Distribution of observations for the 1967-2000 PSID sample, by year

year	Number of Observations	year	Number of Observations
1967	933	1983	1775
1968	1015	1984	1802
1969	1109	1985	1808
1970	1181	1986	1829
1971	1294	1987	1837
1972	1395	1988	1840
1973	1508	1989	1838
1974	1543	1990	1809
1975	1601	1991	1780
1976	1635	1992	1697
1977	1685	1993	1698
1978	1705	1994	1638
1979	1737	1995	1588
1980	1755	1996	1510
1981	1734	1998	1425
1982	1718	2000	1298

Given a normalization one can recover spot prices: these, in conjunction with aggregate wage bills (total labor earnings for different education groups) can be used to back out aggregate supplies of human capital, since the aggregate wage bills are defined as $WB_t^e = MPHC_t^e \times H_t^e$ (where MPHC stands for marginal product of human capital). We use a normalization based on the relative hourly wages observed in our PSID sample in 1989: first we compute average wages by education group for 1989; second we correct for ability composition using information from the NLSY79 (AFQT test together with their gradient on wages). We choose 1989 because people from the NLSY79 are between age 23 and 31, which means most of them are already working.[27] Figure (2) reports the logs of the normalized prices (marginal products) for the three education groups.

In section B of the appendix we use a normalization to obtain price *levels* for different types of human capital and use these prices to approximate total supply of human capital of different types in each given year.

[27]Details regarding the normalization and the ability adjustment are available upon demand.

Table 3: Age polynomials' coefficients

Dependent variable: real log hourly earnings ($1992)				
	Less Than HS	**High School**	**College**	**Pooled**
	Coefficient	Coefficient	Coefficient	Coefficient
	(Std. Err.)	(Std. Err.)	(Std. Err.)	(Std. Err.)
age	.2	.41	.67	.46
	(.015)	(.06)	(.10)	(.05)
age^2	-.01	-.013	-.02	-.014
	(.001)	(.002)	(.004)	(.002)
age^3	1.e-4	2.e-4	3.e-4	2.e-4
	(1.e-5)	(4.e-5)	(6.e-5)	(3.e-5)
age^4	-8.e-7	-1.e-6	-1.6e-6	-1.2e-6
	(2.e-7)	(2.e-7)	(3.7e-7)	(1.8e-7)

B CPS Data

The CPS is a monthly survey of about 50,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics. The survey has been conducted for more than 50 years. Statistics on the employment status of the population and related data are compiled by the Bureau of Labor Statistics (BLS) using data from the Current Population Survey (CPS). The adult universe (i.e., population of marriageable age) is comprised of persons 15 years old and over for March supplement data and for CPS labor force data. Each household and person has a weight that we use in producing population-level statistics. The weight reflects the probability sampling process and estimation procedures designed to account for non-response and under-coverage. We use the CPI for all urban consumer (with base year 1992) to deflate the CPS earnings data and drop all observations that have missing or zero earnings. Since the earning data are top-coded for confidentiality issues, we have extrapolated the average of the top-coded values by using a tail approximations based on a Pareto distribution.[28]

Figure (5) reports the number of people working in each year by education group, as reported in the CPS. It is clear that some strong and persistent trends towards higher

[28]This procedure is based on a general approach to inference about the tail of a distribution originally developed by Hill (1975). This approach has been proposed as an effective way to approximate the mean of top-coded CPS earning data by West (1985); Polivka (2000) provides evidence that this method closely approximates the average of the top-coded tails by validating the fitted data through undisclosed and confidential non top-coded data available only at the BLS.

levels of education have characterized the sample period.

Figure (6) plots both the average earnings by year and total wage bills in billions of dollars for the 3 education groups. Since CPS earning data until 1996 are top coded we report both the censored mean and a mean adjusted by using a method suggested by the BLS (see West (1985)) which is based on the original Hill's estimator to approximate exponential tails. The difference between the two averages is larger for the most educated people who tend to be more affected by top-coding. [29].

B.1 Aggregate technology estimation

In estimating technology parameters, we start from the relatively easier case of a Cobb-Douglas function for aggregate human capital. We define aggregate human capital H as

$$H = H_1^{s_{1t}} H_2^{(1-s_{1t})s_{2t}} H_3^{(1-s_{1t})(1-s_{2t})}$$

Under the assumption of competitive markets one can use aggregate yearly wage bills for different education groups in order to obtain point-estimates of the share parameters s_{1t}, $(1 - s_{1t})\, s_{2t}$ and $(1 - s_{1t})\,(1 - s_{2t})$ under a Cobb-Douglas specification of aggregate technology like equation (**??**). Figure (7) reports these estimated share parameters for each human capital type between 1968 and 2001.[30] The college graduates' labor share more than doubles over this time interval (from 0.2 to 0.4) whereas high-school drop-outs' share falls dramatically from over 0.3 to roughly 0.06. The Cobb-Douglas specification restricts the elasticity of substitution to be equal to one. Retaining the assumption of iso-elasticity between human capital factors, we choose to work with a more general CES specification for the aggregate human capital factor H, like the one in equation (15).

[29] We include also self-employed people in the computation of these aggregates; however, their exclusion has almost no effect on the value of the wage bills, as they never represent more than 5% of the working population in a given education group

[30] Using NIPA data we find the share of physical capital (ϕ) is between 0.3 and 0.35, depending on whether we correct for housing stocks. The long-term averages for human capital shares are 0.33 for college graduates, 0.54 for high school graduates and 0.14 for high school dropouts.

Given the isoelasticity assumption we can express the ratios of wage bills (WB^{edu}) as:

$$WB^3/WB^1 = \frac{(1 - s_{1t} - s_{2t})}{s_{1t}} \left(\frac{H_3}{H_1}\right)^{\rho} \tag{18}$$

$$WB^3/WB^2 = \frac{(1 - s_{1t} - s_{2t})}{s_{2t}} \left(\frac{H_3}{H_2}\right)^{\rho} \tag{19}$$

$$WB^2/WB^1 = \frac{s_{2t}}{s_{1t}} \left(\frac{H_2}{H_1}\right)^{\rho} \tag{20}$$

B.1.1 Human capital aggregates

Dividing the wage bills by the (normalized) marginal products of human capital estimated through from PSID data (see section of the A of the appendix) we obtain point estimates of total efficiency weighted labor supply (human capital aggregates) by education and year. These are plotted in figure (8).

Notice that the estimated stock of college-equivalent human capital does not trend as strongly as the wage bill for college graduates. This is partially due to changes in the marginal product of this factor (see figure 2). However, the time series of human capital stocks give an insight also on the quantitative importance of selection: despite a doubling of both the total number and wage bill of high school graduates, their human capital aggregate has been almost flat over the sample period, suggesting that for this group there has been a reduction in average per worker efficiency. A similar conclusion can be drawn for the college graduates, as their total number went up by almost four times over the sample period, their marginal product also went up whereas their human capital aggregate increased roughly by a factor of two. Big shifts in the distribution of people of different ability over educational outcomes have probably taken place over the sample period.

We incidentally notice that the monetary value of human capital stocks shows a pattern that is very similar to the shares of human capital estimated using the Cobb-Douglas technology specification (figure 7).

B.1.2 Estimation results for aggregate technology parameters

We estimate equation (16) for each of the 3 wage bill ratios. We use two different specifications to estimate the parameters of interest:

1. the first specification does not require any normalization on the *level* of human capital aggregates, but only delivers estimates for the elasticity of substitution between human capital types and the growth rate of the shares' ratio $\frac{g_j}{g_i}$. The initial values of the shares' ratios are not identified in this specification.[31] This specification is based on time-differencing of wage bill ratios in equation (16):

$$\log\left(WB_t^j/WB_t^i\right)-\log\left(WB_{t-1}^j/WB_{t-1}^i\right)=\log\left(\frac{g_j}{gi}\right)+\rho\left[\log\left(\frac{H_{jt}}{H_{it}}\right)-\log\left(\frac{H_{jt-1}}{H_{it-1}}\right)\right]$$

The advantage of this method is that the right-hand side variable $\log\frac{\left(\frac{H_{jt}}{H_{it}}\right)}{\left(\frac{H_{jt-1}}{H_{it-1}}\right)}$ can be approximately measured as the difference between the growth rate in wage bills' ratio and the growth rate in the ratio of marginal products estimated using PSID data (see section A of the appendix).

2. The second specification estimates (16) directly, after backing out the values of $\log\left(\frac{H_2}{H_1}\right)$ through a normalization of the marginal products $MPHC_t^{edu}$ for $edu \in \{1,2,3\}$ and given year t. We choose to normalize marginal products using the average wages of different education groups for year 1968, as observed in our PSID sample.

In both estimation procedures we control for possible endogeneity of the human capital inputs in the production function through an IV approach with lagged regressors (lags up to 5 periods back are included in the first step, depending on the specification). The results, for both methods and for each wage bill ratio, are reported in table (4) with standard errors in parenthesis. The estimation procedure is based on a stacking method which allows to test for differences in the elasticity parameters in different wage ratios (like in a Chow test).

Table (B.1.2) reports the results of an F-test for specification (1) in differences with instruments going back to the 4th lag and for specification (3), in levels with instruments going back to the 5th lag.[32] We are unable to reject the null hypothesis that the aggregate technology is iso-elastic at 5% level of significance. The null hypothesis cannot be rejected by a much larger margin in the growth rates specification.

[31]One would have to make a normalization assumption on the share parameters (and, by implication, on the human capital aggregates) to back out the share parameters' values for each year in the sample.

[32]Results for the isoleasticity test for the other choice of instruments are available upon request.

	Specification : growth rates		Specification : levels	
	(1)	**(2)**	**(3)**	**(4)**
First stage IV	up to 4 lags	up to 3 lags	up to 5 lags	up to 4 lags
Number of obs.	75	78	75	78
	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)
$\rho_{2,1}$	0.540	.145	0.589	.476
	(0.183)	(.324)	(0.207)	(.224)
$\rho_{3,2}$	0.582	.542	0.506	.441
	(0.352)	(.351)	(0.114)	(.113)
$\rho_{3,1}$	0.454	.394	0.893	.900
	(0.193)	(.263)	(0.118)	(.117)
$g_{2,1}$	0.021	.043	0.018	.026
	(0.012)	(.019)	(0.013)	(.014)
$g_{3,2}$	0.012	.013	0.015	.016
	(0.009)	(.010)	(0.002)	(.002)
$g_{3,1}$	0.041	.045	0.008	.008
	(0.016)	(.022)	(0.009)	(.009)
$s_{2,1}$			0.449	.452
			(0.046)	(.049)
$s_{3,2}$			-0.424	-.483
			(0.117)	(.119)
$s_{3,1}$			0.355	.360
			(0.099)	(.100)

Table 4: Estimation results : aggregate technology (isoelastic CES spec.), Unrestricted ρ

Testing the isoelastic restriction

	(1): growth rates (4 lags)		**(3): levels (5 lags)**	
Null Hypothesis	F-stat.	Prob.>F-stat.	F-stat.	Prob.>F-stat.
$\rho_{(2/1)} = \rho_{(3/2)}$	$F_{(1,69)} = 0.01$	0.916	$F_{(1,66)} = 0.12$	0.726
$\rho_{(3/2)} = \rho_{(3/1)}$	$F_{(1,69)} = 0.10$	0.751	$F_{(1,66)} = 5.54$	0.022
$\rho_{(2/1)} = \rho_{(3/1)}$	$F_{(1,69)} = 0.10$	0.748	$F_{(1,66)} = 1.63$	0.207
$\rho_{(2/1)} = \rho_{(3/1)} = \rho_{(3/1)}$	$F_{(2,69)} = 0.08$	0.927	$F_{(2,66)} = 2.87$	0.064

Table 5: Tests for equality of elasticities of substitution among human capital inputs

Next, we estimate a restricted version of equations (16) with a unique ρ for all wage-bill ratios. This improves the efficiency of the estimator, which is particularly valuable since we are using a relatively short time series (approximately 30 observations). The results for this specification are reported in table (6).

	Specification : growth rates		**Specification : levels**	
	(1)	**(2)**	**(3)**	**(4)**
First stage IV	up to 4 lags	up to 3 lags	up to 5 lags	up to 4 lags
Number of obs.	75	78	75	78
	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)
ρ	0.510	.357	0.677	.641
	(0.121)	(.170)	(0.079)	(.079)
$g_{2,1}$	0.023	.031	0.013	.016
	(0.009)	(.012)	(0.005)	(.005)
$g_{3,2}$	0.014	.017	0.012	.012
	(0.006)	(.007)	(0.002)	(.002)
$g_{3,1}$	0.036	.048	0.025	.028
	(0.011)	(.015)	(0.006)	(.006)
$s_{2,1}$			0.431	.419
			(0.027)	(.027)
$s_{3,2}$			-0.252	-.275
			(0.082)	(.085)
$s_{3,1}$			0.180	.143
			(0.068)	(.070)

Table 6: Estimation results : aggregate technology (isoelastic CES spec.), Restricted ρ

The initial conditions for the share parameters of the CES production function can be identified by using the estimated constant ratios $\frac{s_j}{s_i}$. Solving for the share values in 1968 we obtain: $\widehat{s}_1 = \frac{1}{1+\widehat{\left(\frac{s_2}{s_1}\right)}+\widehat{\left(\frac{s_3}{s_1}\right)}}$, $\widehat{s}_2 = \frac{\widehat{\left(\frac{s_2}{s_1}\right)}}{1+\widehat{\left(\frac{s_2}{s_1}\right)}+\widehat{\left(\frac{s_3}{s_1}\right)}}$ and $\widehat{s}_3 = \frac{\widehat{\left(\frac{s_3}{s_1}\right)}}{1+\widehat{\left(\frac{s_2}{s_1}\right)}+\widehat{\left(\frac{s_3}{s_1}\right)}}$. By construction we have $\widehat{s}_1 + \widehat{s}_2 + \widehat{s}_3 = 1$, and the human capital shares sum up to 1. Using the estimated time trend components we can compute a set of share ratios for each year in the sample; denoting $S_{j,i} = \left[\widehat{\left(\frac{s_j}{s_i}\right)} + \left(\frac{\widehat{g_j}}{gi}\right) * (year - 1968)\right]$, we have that in general

Year	LTHS	HS	College	Year	LTHS	HS	College
1968	0.26	0.41	0.32	1985	0.21	0.40	0.39
1969	0.26	0.41	0.33	1986	0.20	0.40	0.39
1970	0.26	0.41	0.33	1987	0.20	0.40	0.40
1971	0.25	0.41	0.34	1988	0.20	0.40	0.40
1972	0.25	0.41	0.34	1989	0.20	0.40	0.40
1973	0.25	0.41	0.34	1990	0.19	0.40	0.41
1974	0.24	0.41	0.35	1991	0.19	0.40	0.41
1975	0.24	0.41	0.35	1992	0.19	0.40	0.42
1976	0.24	0.41	0.35	1993	0.18	0.40	0.42
1977	0.23	0.41	0.36	1994	0.18	0.39	0.42
1978	0.23	0.41	0.36	1995	0.18	0.39	0.43
1979	0.23	0.41	0.37	1996	0.17	0.39	0.43
1980	0.22	0.41	0.37	1997	0.17	0.39	0.44
1981	0.22	0.41	0.37	1998	0.17	0.39	0.44
1982	0.22	0.41	0.38	1999	0.17	0.39	0.44
1983	0.21	0.40	0.38	2000	0.16	0.39	0.45
1984	0.21	0.40	0.39	2001	0.16	0.39	0.45

Table 7: Shares of different types of human capital by year. CES human capital aggregation based on estimates from specification (3). LTHS=Less than high school; HS=High School.

$$
\begin{aligned}
\widehat{s}_{1t} &= \frac{1}{1+S_{2,1}+S_{3,1}} \\
\widehat{s}_{2t} &= \frac{S_{2,1}}{1+S_{2,1}+S_{3,1}} \\
\widehat{s}_{3t} &= \frac{S_{3,1}}{1+S_{2,1}+S_{3,1}}
\end{aligned}
$$

Figure (9) plots the evolution of the shares parameters estimated from the restricted specification (3). Table (7) reports the corresponding point estimates.

C NLSY79 Data

The NLSY79 is a representative sample of 12,686 American young men and women who were 14-22 years old when they were first surveyed in 1979. Data was collected yearly from 1979 to 1994, and biennially from 1996 to the present.

The following three subsamples comprise the NLSY79: (1) a cross-sectional sample of 6,111 respondents designed to be representative of the non-institutionalized civilian segment of young people living in the United States in 1979 and born between January 1, 1957, and December 31, 1964 (ages 14–21 as of December 31, 1978) (2) a supplemental sample of 5,295 respondents designed to oversample civilian Hispanic, black, and economically disadvantaged non-black/non-Hispanic youth living in the United States during 1979 and born between January 1, 1957, and December 31, 1964; (3) a sample of 1,280 respondents designed to represent the population born between January 1, 1957, and December 31, 1961 (ages 17–21 as of December 31, 1978), and who were enlisted in one of the four branches of the military as of September 30, 1978

C.1 The ASVAB tests and AFQT measures

During the summer and fall of 1980 NLSY79 respondents participated in an effort of the U.S. Departments of Defense and Military Services to update the norms of the Armed Services Vocational Aptitude Battery (ASVAB). A total of 11,914 civilian and military NLSY79 respondents completed this test. The ASVAB consists of a battery of 10 tests that measure knowledge and skill in the following areas: (1) general science; (2) arithmetic reasoning; (3) word knowledge; (4) paragraph comprehension; (5) numerical operations; (6) coding speed; (7) auto and shop information; (8) mathematics knowledge; (9) mechanical comprehension; and (10) electronics information. A composite score derived from select sections of the battery can be used to construct an approximate and unofficial Armed Forces Qualifications Test score (AFQT) for each youth. The AFQT is a general measure of trainability and a primary criterion of enlistment eligibility for the Armed Forces. Two methods of calculating AFQT scores, developed by the U.S. Department of Defense, have been used by CHRR to create two percentile scores, an AFQT80 and an AFQT89, for each respondent.

For each sample member we compute both AFQT80 and AFQT89, as well as their percentile distribution.

C.2 Sample Selection for wage equations and estimates of ability gradients

In the analysis of NLSY wage data we use 3 different measures for hourly wages. Specifically we use:

- a wage variable corresponding to the hourly rate of pay of the current or most recent job. This measure is based on the same question which is used to record hourly wages in the CPS. This wage measure is available only from 1979 to 1994.
- a wage variable corresponding to hourly rate of pay in the first reported job. This measure is available in every wave between 1979 and 2002.
- a hourly wage rate obtained dividing total earning by total hours worked in the previous calendar year. This variable can constructed for each wave between 1979 and 2002. The earnings' measure includes wages, salary, commissions or tips from all jobs, before deductions for taxes.

Some of the wage measure are censored in some waves and depending on the measure of wages we use, we select a different sample of workers.

C.2.1 Sample selection for different measures of wages

Sample 1: Annual earnings divided by annual hours of work

The initial sample includes 11878 individuals who are out of school. We start by dropping observations which relate to study periods, which leaves the number of sample members unchanged because all individuals work during at least one sample year. We then get rid of individuals who report total yearly earnings which are either missing or topcoded in at least one sample year: this keeps consistency with the PSID sampling procedure. This leaves us with 11522 sample members. We keep only those observations which report positive earnings, which further reduces the sample size to 11173 individuals. We drop observations which refer to years in which the individual worked less than 400 or more than 5840 hours, which reduces the sample to 10904 individuals. Then we get rid of agents who are officially classified as unemployed, out-of-labor-force and in the military, which leaves us with 10358 individuals. We drop individuals who report extremely high or low real hourly wages (more than \$400 or less than \$1 in 1992 dollars) which leaves 9452 individuals in the sample. We also drop individuals who report a log growth in wages larger than 4 or smaller than -2, which brings the sample to 9346 individuals. Finally, to keep consistency of the education groups, we drop individuals who change the highest completed grade of education during their working lives, which reduces the sample substantially to 7175 people. This final sample is then split into 3 education groups: less than high-school (1119 individuals), high school graduates (5001 individuals) and college graduates (1052 individuals). If we restrict the sample to people who are in the cross-sectional (representative) sample, the total number of individuals more than halves to 3504.

Sample 2: "Current/most recent job" measure of wages (CPS-type)

Again we start with 11878 individuals and we get rid of observation for current for students and people who have a missing wage measure, leaving the sample size unchanged. We then drop those observations which have a zero wage, leaving only 11224 individuals in our sample. We drop observations with reported annual work hours which are missing, below 400 or large than 5840: the sample reduces to 10937 individuals. We also keep only people who are formally employed, and drop individuals who are reported as unemployed, out-of-labor-force and in the military. this reduces the sample to 10592 individuals. Dropping individuals who report (at least once) hourly wages above \$400 or below \$1 further reduces the sample to 10202. We also get rid of agents who report log wage increases larger than for or smaller than -2, which leaves 10056 workers in the sample. Finally, we drop people who change their education level during their working life, which gives us a final sample of 7954 individuals. When we split this sample in 3 education groups, we get a HS drop-outs sample of 1341 individuals, a HS graduates sample of 5403 individuals and a college graduates sample of 1206 individuals. If we consider only workers from the cross-sectional sample we end up with a total size of 3983 individuals.

Sample 3: Wage of "Job 1" (first job reported)

The initial sample is always 11878 individual workers. Students and missing wage observations are dropped. When we drop observations with zero wages we go down to 11423 individuals. Dropping observations with zero, missing or larger than 5840 hours worked per year we go down to 11211 individuals. When we restrict the sample to people who are formally employed we get 10758 individuals. We drop people who report hourly wages which are below $1 or above $400 dollars in real 1992 terms: the sample goes to 10343. We drop people whose log wages record changes above 4 or below -2, reducing the sample to 10197. Finally we get rid of people who report changes in highest degree of education during working life, bringing down the sample to 7799 members, who are split into 1282 HS drop-outs, 5348 HS graduates and 1165 college graduates. Workers from the cross-sectional sample are only 3855, roughly half of the final sample of 7799.

C.3 Estimated gradients of AFQT89 on hourly wages

Here we report the details of the estimation of the gradient of ability as measured by the AFQT89: we use specifications with time dummies to control for variation in market wages, but the estimated effects are almost identical to the estimates obtained without time dummies.[33]

We use all workers including NLSY79 over-samples in our estimation in order to maximize the number of observations: a dummy is introduced to control for possible hourly wage differences of workers from the over-samples.[34] We also run specifications based on measures of wages which are not purged of age effects: the estimates based on these measures are generally close to the ones obtained for age-free wages reported below. Complete estimation results are available on request from the authors. All standard errors are corrected for individual clustering. Results are reported for pooled samples as well as by education group (LTHS=Less Than High School,HSG=High School Graduates,CG=College Graduates). In summary, we have 3 tables:

- Table (8) reports estimates based on sample 1 (wage rates computed from annual earnings purged of age components, including over-samples, all year from 1979 to 1998).

[33] We have also estimated gradients for two sub-samples referring to the periods before and after 1988: it is apparent that the return to ability as measured by the AFQT89 have changed over time. For all wage measures we find that the difference across education groups in returns to ability (as measured by AFQT89 scores) has shrank over time. On the other hand estimates of the pooled effect are larger for the more recent sample of workers (1988-1998). The return to ability seems to have gone up in aggregate, and become more homogenous across education groups!

[34] The over-sample dummies are not significant in most cases and, even when significant, they are small in size.

Table 8: **Estimated ability gradient. Sample 1: Wage = earnings divided by hours worked**

Education group	Gradient (S.E.)	# of obs.	# of workers
LTHS	.46 (.07)	7,897	1,119
HSG	.61 (.03)	5,003	42,916
CG	.78 (.09)	1,052	8,655
pooled	.76 (.02)	7,175	59,499

- Table (9) reports estimates based on sample 2 (CPS-type wage rates based on most recent job purged of age components, including over-samples, all year from 1979 to 1994).

Table 9: **Estimated ability gradient. Sample 2: Wage = CPS-type**

Education group	Gradient (S.E.)	# of obs.	# of workers
LTHS	.36 (.06)	1,341	8,982
HSG	.54 (.03)	5,403	42,270
CG	.89 (.09)	1,206	8,719
pooled	.71 (.02)	7,954	60,009

- Table (10) reports estimates based on sample 3 (first reported job) purged of age components, including over-samples, all year from 1979 to 1998).

Table 10: **Estimated ability gradient. Sample 3: Wage = first job reported**

Education group	Gradient (S.E.)	# of obs.	# of workers
LTHS	.39 (.06)	1,282	9,281
HSG	.57 (.03)	5,350	46,755
CG	.93 (.10)	1,165	9,713
pooled	.77 (.02)	7,799	65,787

C.4 The distribution of permanent characteristics (ability)

C.4.1 Children of NLSY79

The Children of the NLSY79 survey began in 1986. The expanded mother-child data collection has occurred biennially since then. This survey consists of detailed information on the development of children born to NLSY79 women During these biennial surveys, a battery of child cognitive, socio-emotional, and physiological assessments are administered to NLSY79 mothers and their children. In addition to these assessments, the Children of the NLSY79 are also asked a number of questions in an interview setting. In 1994, children age 15 and older, the "Young Adults," first responded to a separate survey with questions similar to those asked of their mothers and a wide array of attitudinal and behavioral questions tailored to their age group. The number of children born to interviewed mothers has increased from 5,255 in 1986 to 8,323 in 2002. Interviews were completed during 2002 with 7,467 children, or 90 percent of children born to interviewed NLSY79 mothers.

C.4.2 Sample selection of the mother-child-pairs

The original NLSY79 sample includes 12,686 individuals, of whom only 11,878 took the tests which allow to compute AFQT scores. For such individuals we are able to construct two types of AFQT scores: the AFQT80 and the AFQT89. We use the latter score in our analysis, which is also the ability measure used in the estimation of the wage equations.

In the total NLSY sample there are 11,340 children born to the total 6,283 female respondents of the NLSY79 (not all of them had children: only 4,890 of them are mothers meaning they have at least one child). We link the children's file to the main data file using the individual identifier for mothers. Each child has observations taken in different years; however many child/year combinations do not have any test score observations. The child test scores are the PIAT Math, the PIAT reading comprehension, the PIAT Reading Recognition, and the PPVT score. We use only the most recent PIAT Math test scores to rank children's ability: in particular, we use standardized scores of the PIAT Math test, which are derived on an age-specific basis from the child's raw score and are comparable across ages. We get rid of the mother-child pairs which refer to earlier PIAT scores: this leaves us with 3,389 mothers and 7,589 mother-child pairs.

Given the presence of sampling problems for the children of NLSY over-sample members, we restrict our attention only to mothers who are part of the cross-sectional (nationally representative) sample of the NLSY79, which further reduces our mother-child pairs to 4,455 and the total number of mothers to 2,087. Table (11) reports the distribution of children's age at the time of test in our final sample.

Finally we use the test-scores to assign individual specific percentiles to both mothers and children, according to the relative ranking of their scores (AFQT89 for mothers,

Table 11: **Child's age at time of test (relative frequency)**

Age	Number	Per cent	Age	Number	Per cent
5	98	2.2	12	331	7.4
6	202	4.5	13	1,208	27.1
7	194	4.4	14	1,081	24.3
8	231	5.2	15	87	2.0
9	251	5.6	16	49	1.1
10	301	6.8	17	45	1.0
11	368	8.3	18	9	0.2

Total number of mother-child pairs: 4,455

PIAT Math for children) in the sample. These percentiles are used to split the sample population of mother and children in ability groups.

The fact that children took the PIAT test at different ages should have no relevance because we use standardized scores which control for the age of the test-subject. However, in order to verify the robustness of the estimated transition matrices, we also use a smaller sample including only mother-child pairs in which the child was at least 13 years of age at the time of the test. This sample consists of 2,479 mother-child pairs and of 1,412 mothers. The age distribution of children at the time of the test for this sample is reported in table (12).

Table 12: **Child's age at time of test (relative frequency) - only children tested at age 13 or later**

Age	Number	Per cent
13	1,208.0	48.7
14	1,081.0	43.6
15	87.0	3.5
16	49.0	2.0
17	45.0	1.8
18	9.0	0.4

C.4.3 Ability transition matrices

After splitting mothers and children into quintiles according to their relative score in the sample, we compute the conditional probabilities of transiting from a given mother's quintile to a given child's quintile. Results for the larger sample (including test scores for all test-ages) are reported in table (13). Quintile number 1 is the lowest, while quintile number 5 is the highest.

For each maternal quintile, the first row reports the number of sample children in each quintile, the second row reports the conditional probability of ending up in that quintile.

Table 13: **Ability transition, by quintile**

	Children					
Mothers	1	2	3	4	5	Total
1	416	218	180	59	42	915
	45.5%	23.8%	19.7%	6.5%	4.7%	100.0%
2	228	219	219	143	100	909
	25.8%	24.2%	24.2%	15.7%	11.0%	100.0%
3	146	203	247	173	143	912
	16.0%	22.3%	27.1%	19.0%	15.7%	100.0%
4	100	150	225	183	218	876
	11.4%	17.1%	25.7%	20.9%	24.9%	100.0%
5	61	64	164	204	350	843
	7.2%	7.6%	19.5%	24.2%	41.5%	100.0%
Total	951	854	1,035	762	853	4,455
	21.3%	19.2%	23.2%	17.1%	19.2%	100.0%

Each cell reports absolute number and conditional probability

We also compute a transition matrix for the smaller sample which excludes mother-child pairs where the child was younger than 13 when taking the test. The transition matrix based on this sample is summarized in table (14).

One can easily check that restricting the test-age of children implies very small differences in the ability transition probabilities.

C.4.4 The stationary distribution of ability

Table (15) reports relevant statistics for the distribution of the logs of AFQT89 for the set of mothers used to compute the transition matrix of ability. The statistics are presented by quintiles of the distribution.

Similarly, table (16) reports descriptive statistics for the distribution of AFQT89 test-scores (in logs) over the whole cross-sectional sample of the NLSY79. It appears that the

Table 14: **Ability transition, by quintile - only children tested at age 13 or later**

	Children					
Mothers	1	2	3	4	5	Total
1	228	127	88	53	25	521
	43.8%	24.3%	16.9%	10.2%	4.8%	100.0%
2	123	129	111	78	55	496
	24.8%	26.0%	22.4 %	15.7%	11.1%	100.0%
3	86	108	136	81	80	491
	17.5%	22.0%	27.7%	16.5%	16.3%	100.0%
4	53	83	113	133	128	510
	10.4%	16.3%	22.2%	26.1%	25.1%	100.0%
5	35	39	77	125	185	461
	7.6%	8.5%	16.7%	27.1%	40.1%	100.0%
Total	525	486	525	470	473	2,479
	21.2%	19.6%	21.2%	19.0%	19.1%	100.0%

Each cell reports absolute number and conditional probability

Table 15: **Descriptive statistics by quintile: mothers' AFQT89 (logs)**

quintile	**min**	**max**	**mean**	**median**
1	-0.68	-0.19	-0.31	-0.30
2	-0.19	-0.02	-0.09	-0.08
3	-0.02	0.09	0.04	0.04
4	0.10	0.19	0.14	0.14
5	0.20	0.32	0.25	0.25
Total	-0.68	0.32	0.00	0.03

distribution of AFQT scores among mothers is extremely similar to the distribution of AFQT scores in the whole cross-sectional sample.

Table 16: **Descriptive statistics by quintile: all cross-sectional sample's AFQT89 (logs)**

quintile	min	max	mean	median
1	-0.68	-0.20	-0.32	-0.31
2	-0.19	-0.02	-0.09	-0.08
3	-0.01	0.09	0.04	0.04
4	0.09	0.20	0.14	0.15
5	0.20	0.32	0.25	0.25
Total	-0.68	0.32	0.00	0.03

The AFQT89 scores (over the cross-sectional sample of the NLSY79) can be matched with information about the education levels of the subjects in order to measure education shares by ability level. The implied education shares are reported in table (17).

Table 17: **Education shares (%) by AFQT89 quintile. Full cross-sectional sample NLSY79**

	quintile (AFQT89)					
Education	1	2	3	4	5	Total
less than H.S.	32.00	9.21	3.94	0.96	0.28	9.48
H.S.	66.21	83.08	78.51	61.86	30.96	64.81
College	1.79	7.71	17.55	37.19	68.76	25.72
Total	100.00	100.00	100.00	100.00	100.00	100.00

C.5 Estimates of Labor Shock Processes

In order to identify the parameters of the persistent and transitory shocks to wages we use the Minimum Distance Estimator originally proposed by Chamberlain (1984). A detailed description of the estimation method is presented in Heathcote, Storesletten, and Violante (2004). For each year in which wage data are available, the method allows to identify and estimate the following parameters of the persistent shock process $z_{it}^{edu} = \rho^{edu} z_{it-1}^{edu} + \varepsilon_{it}^{edu}$:

- autoregressive coefficient ρ^{edu};

- year-specific variance of the innovation ε_{it}^{edu}, denoted as $\sigma_{\varepsilon}^{2}(t)^{edu}$;

- initial condition for the variance of the innovation ε_{it}^{edu}, denoted as $\sigma_{\varepsilon}^{2}(0)^{edu}$.

The MDE also allows us to estimate year-specific values for the variance of the transitory shocks m_{it}^{edu}, which we denote as $\sigma_{m}^{2}(t)^{edu}$

Table (18) reports the estimates of these parameters obtained from the sample of CPS wages (sample 2). Estimates are for the period 1979 to 1993 (details available from the authors).

Table 18: **Estimated Variances and autoregressive coefficients for the transitory and persistent shocks to wages - NLSY data using CPS-type wage measures.**

	H.S. dropouts		H.S. graduates		Coll. graduates	
	ρ	0.936	ρ	0.951	ρ	0.945
	$\sigma_{\varepsilon}^{2}(0)$	0.105	$\sigma_{\varepsilon}^{2}(0)$	0.101	$\sigma_{\varepsilon}^{2}(0)$	0.128
	$\sigma_{\varepsilon}^{2}(t)$	$\sigma_{m}^{2}(t)$	$\sigma_{\varepsilon}^{2}(t)$	$\sigma_{m}^{2}(t)$	$\sigma_{\varepsilon}^{2}(t)$	$\sigma_{m}^{2}(t)$
YEAR						
1979	0.012	0.026	0.010	0.036	0.012	0.002
1980	0.016	0.060	0.010	0.045	0.014	0.004
1981	0.005	0.055	0.015	0.052	0.016	0.045
1982	0.013	0.053	0.012	0.059	0.003	0.072
1983	0.014	0.086	0.008	0.055	0.013	0.067
1984	0.023	0.095	0.018	0.069	0.011	0.058
1985	0.018	0.081	0.021	0.060	0.021	0.059
1986	0.012	0.056	0.021	0.064	0.018	0.075
1987	0.043	0.078	0.023	0.062	0.016	0.082
1988	0.006	0.086	0.022	0.064	0.032	0.059
1989	0.022	0.068	0.019	0.068	0.019	0.075
1990	0.043	0.060	0.016	0.052	0.024	0.040
1991	0.004	0.101	0.017	0.051	0.025	0.060
1992	0.026	0.064	0.022	0.036	0.027	0.043
1993	0.039	0.074	0.025	0.047	0.055	0.051
MEAN	*0.020*	*0.070*	0.017	*0.055*	*0.020*	*0.052*

D NLSY97 and the measurements of inter-vivos transfers

Inter-vivos transfers (i.e.gifts) from parents to their children are captured by a set of variables in the NLSY97 which is found in the 'Income' subsection of the survey.[35] These refer to *all* income, transferred from parents or guardians to youths, that are neither loans nor regular allowance. This is elicited through a series of questions beginning with the following: "(YINC-5600) Do you live with your mother $figure$ and father $figure$". Respondents have the option of responding to this question with either a yes or a no. If the respondent answers no, they are asked a further battery of questions whether either of their biological parents are alive. If the respondent answers yes to any of these questions, then he is asked to specify the exact and estimated value of the inter-vivos transfers. This is phrased as the following when the respondent lives with both parents: "(YINC-5700) Other than allowance, did your parents give you any money in [insert year]? Please include any gifts in the form of cash or a check but do not include any loans from your parents". For youths that are living at home, inter-vivos transfers also contain an imputed value for rent which is based on the mean value paid by independent youths of the same age.

In order to relate the size of the transfer to the characteristics of the giving guardians we use information about parental wages, total household income, youth reported net worth and highest level of education attained by either parent collected from the NLSY 1997. Net worth is composed by subtracting liabilities from assets, where assets include real estate and other property ownership, pensions, savings and stocks. Liabilities include mortgage, student loans and other debts. Although parent reported net worth would likely better capture actual household value, it is elicited only once in the initial phase of ths survey and therefore is less useful when measuring variation in yearly transfers.

D.1 Sampling procedure

We use waves from 1997 to 2003. Data for 2004 are dropped as there are no inter-vivos amounts available after that year. This gives us an initial sample of 12,686 youths who

[35] Also the 'College Experience' subsection has some information about income transferred from parents to children that is earmarked as financial aid while attending a post-secondary academic institution. However these transfers are not fully consistent with the information in the 'Income' section, and contain many skips. Most importantly, they do not cover all transfers.

were between age 12 and 16 in 1997. Only respondents that are part of the cross-sectional (representative) sample are kept, which leaves 6,748 individuals.

Furthermore, we drop observations for youths below age 16, which gives us a sample with 6,346 youths and a total number of observations equal to 21,149. We drop 13 cases reporting positive inter-vivos transfers which are more than twice the size of their households' negative net worth: these observations are very likely to be misreported. This creates a final sample of 6,346 youths and 21,136 observations.

In the total sample, 35% of youths report living in households with both biological parents as guardians, 7% live in two-parents households with the biological mother, 2% live in two-parents households with the biological father and 0.5% live in adoptive parents households. 18% of youths live in single parent households, 16% single mothers and 2% single fathers. 0.1% constitute children living with foster parents, 1.2% no parents but living with another relative and 35% report living in a household where the relationship to the guardian cannot be described by any of the above.

The age distribution in our final youths' sample including the proportions of those enrolled in college for each age is reported in table (19).

Table 19: **Age distribution of final NLSY97 sample**

age	Number	Per cent	# Enrolled	Per cent	# Enrolled & Live @ Home	Per cent
16	1,004	5	26	3	23	88
17	3,381	16	1,247	37	1,091	87
18	5,743	27	2,446	43	2,030	83
19	4,538	21	1,847	41	1,402	76
20	3,328	16	1,306	39	903	69
21	2,150	10	643	30	374	58
22	527	5	217	22	109	50
Overall	21,136	100	7,732	37	5,932	77

Overall, 37% of the sample are enrolled in college, and from this group of college enrollees, 77% live at home. College enrollment in the population begins at age 17 and begins to drop off after 18 which may as well be a function of survey attrition. Those who live at home form the majority among college attendees for all ages only reaching a minimum of 50% at 22 years old.

In principle, observations should be weighted when tabulating sample characteristics in order to describe the represented population. However the use of weights without other adjustments is inappropriate when using samples generated after dropping observations reporting item non-responses. We do use the BLS custom weighting engine to construct specific weights for our sample but our results change only marginally when we use weights. Therefore we use only results from the unweighted sample.

D.2 Early transfers and family characteristics

In the final sample, 32.4% of observations report positive intervivos transfers elicited from the relevant survey questions, meaning 67.6% did not receive any transfers. 75.1% of observations reported positive intervivos transfers when imputed rent is included with the amount. The value of imputed rent varies from age to age with a minimum of $4,733 for 16 year olds and a maximum of $6,615 for 22 year olds. We express all intervivos transfers in year 2000 dollars. Table (20) reports the average and median yearly transfer amount by age group and standard deviation of the distribution of transfers, with and without rent imputation, and with and without observations reporting zero transfers.

It is evident there is a large divide in mean and median values with and without rent. There are 13,880 cases that report living at home and as such a majority of cases integrate imputed rent with the amount of intervivos transfers, even if they received no monetary intervivos. The median value for intervivos transfers including rent is higher because youths living at or away from home are integrated in the final sample, and the amounts transferred to each independent youth pulls down the mean by being less than the value of imputed rent. This phenomenon is observed throughout the summary statistics.

For the sample of positive transfers only where rent is included, the average transfer is $5,054 per year, and over the period from age 16 to age 22 this sums up to an average total transfer of $35,378 per youth. The median transfer is higher and equal to $5,282 over all age groups: this corresponds to a median total transfer between age 16 and 22 of $36,974.

In order to have an idea of the relative magnitude of the transfers we use information regarding parental wages, household income and net worth, and education of the most educated residential parent/guardian. In these tables, transfers are measured on a yearly basis. Each table contains summary statistics with and without rent, and with and

Table 20: **Descriptive statistics: distribution of inter-vivos transfers by age of youth.**

Positive Transfers only								
	Rent				No Rent			
age	**mean**	**median**	**stand.dev.**	**obs**	**mean**	**median**	**stand.dev.**	**obs**
16	4,801	4,966	1,601	812	706	310	1,302	372
17	4,707	4,765	1,565	2,824	860	423	1,543	1,184
18	5,013	5,014	1,863	4,711	1,073	479	2,068	2,027
19	5,209	5,368	2,258	3,408	1,305	500	2,386	1,450
20	5,261	5,484	2,626	2,299	1,601	500	2,783	1,009
21	5,053	5,318	2,833	1,288	1,725	486	3,199	573
22	5,773	6,615	3,262	527	1,921	670	3,489	234
Overall	5,054	5,282	2,179	15,869	1,227	486	2,342	6,849
Whole sample								
	Rent				No Rent			
16	3,883	4,966	2,375	1,004	263	0	863	1,004
17	3,931	4,765	2,257	3,381	301	0	1,001	3,381
18	4,112	5,014	2,559	5,743	379	0	1,331	5,743
19	3,912	5,368	2,984	4,538	417	0	1,479	4,538
20	3,634	5,484	3,268	3,328	485	0	1,700	3,328
21	3,027	1,945	3,308	2,150	460	0	1,818	2,150
22	3,067	287	3,736	992	453	0	1,878	992
Overall	3,795	5,014	2,889	21,136	398	0	1,452	21,136

without zero transfers.

The wage information is available for 3,978 observations as it is only asked every year up to 2001 and refers to the responding parent/guardian. Household income and net worth data are available for all years up to 2003 from the youth survey. Household income and net worth are reported for 17,243 observations. Top coding for parental wages, household income and net worth are conducted at the top 2% for each year, which leads to inconsistent truncation levels and skewing of the sample distribution. Therefore to reduce this effect, 555 observations where household income is above $240,000 are excluded. Additionally, 43 observations where net worth exceeds $700,000 and 101 cases where parental wage exceeds $150,000 are excluded from summary statistic analysis to avoid similar distributional skewing. Exclusion in this context refers to changing their responses to missing rather than dropping them entirely. Education of residential parents is available for all sample observations.

We report the mean, median, standard deviation and number of observations of the transfers' sample: (1) by quartiles of parent/guardian wage in table (21); (2) by quartiles of household income in table (22); (3) by quartiles of youth reported household net worth in table (23); and (4) by education group in table (24).

Across the results, the general trend is that intervivos transfers increase as income, parental wages, household net worth and maximum parental education increase regardless of sample restrictions.

When the analysis is modified such that only people who are currently enrolled in college are examined, the broad patterns across all these variables and the various sample restrictions are replicated. Intervivos transfers within each category naturally increase by anywhere from $500 to well over $1,000 since college enrolled youths are more likely to receive intervivos transfers. The main difference is with respect to parental wage quartiles where mean transfers in the 2nd quartile are larger than those from the lowest quartile. Further experimentation is pursued where the rent imputation is removed from youths aged 16 or 17 years old based on the wisdom that high school aged youths remain at home as a matter of course. Whether the sample is restricted to college attendees only or not, the effect on intervivos transfers is marginal since these youths make up a minority of the total sample.

Table 21: **Descriptive statistics: inter-vivos transfers by parental wage quartile.**

Positive Transfers only								
	Rent				No Rent			
age	**mean**	**median**	**stand.dev.**	**obs**	**mean**	**median**	**stand.dev.**	**obs**
q1	5,113	5,014	1,473	923	949	317	1,812	382
q2	5,263	5,014	1,578	913	1,085	500	1,984	375
q3	5,341	5,027	1,629	896	1,070	500	1,978	373
q4	5,405	5,100	1,815	908	1,170	500	2,233	375
Overall	5,279	5,014	1,631	3,640	1,068	475	2,006	1,505
Whole sample								
	Rent				No Rent			
q1	4,578	5,014	2,103	974	316	0	1,108	974
q2	4,928	5,014	1,999	975	388	0	1,319	975
q3	5,093	5,014	1,938	959	454	0	1,384	959
q4	5,232	5,065	1,995	969	502	0	1,561	969
Overall	4,957	5,014	2,024	3,877	415	0	1,354	3,877

Table 22: **Descriptive statistics: inter-vivos transfers by household income quartile.**

Positive Transfers only								
	Rent				No Rent			
age	**mean**	**median**	**stand.dev.**	**obs**	**mean**	**median**	**stand.dev.**	**obs**
q1	4,091	5,014	2,688	3,116	1,186	479	2,333	1,408
q2	4,967	5,214	1,980	3,117	1,131	479	2,191	1,407
q3	5,473	5,368	1,613	3,105	1,119	486	1,982	1,416
q4	5,699	5,368	1,928	3,112	1,414	517	2,584	1,396
Overall	5,057	5,306	2,179	12,450	1,212	486	2,284	5,627
Whole sample								
	Rent				No Rent			
q1	2,072	146	2,785	4,205	372	0	1,403	4,205
q2	3,060	4,765	2,877	4,144	343	0	1,334	4,144
q3	4,531	5,114	2,514	4,167	397	0	1,334	4,167
q4	5,438	5,368	2,106	4,172	522	0	1,675	4,172
Overall	3,773	5,014	2,896	16,688	409	0	1,445	16,688

Table 23: **Descriptive statistics: inter-vivos transfers by household net worth.**

Positive Transfers only								
	Rent				No Rent			
age	**mean**	**median**	**stand.dev.**	**obs**	**mean**	**median**	**stand.dev.**	**obs**
q1	4,875	5,017	1,701	2,290	838	400	1,512	930
q2	4,893	5,014	2,000	1,977	974	414	2,029	930
q3	4,990	5,018	1,982	2,134	1,116	486	2,049	925
q4	5,175	5,086	2,083	2,133	1,300	500	2,437	928
Overall	4,983	5,014	1,945	8,534	1,057	479	2,039	3,713
Whole sample								
	Rent				No Rent			
q1	3,785	5,014	2,524	2,949	264	0	934	2,949
q2	3,913	4,976	2,619	2,357	318	0	1,230	2,357
q3	4,057	5,014	2,665	2,650	398	0	1,338	2,650
q4	4,295	5,014	2,716	2,651	505	0	1,645	2,651
Overall	4,009	5,014	2,636	10,607	370	0	1,308	10,607

Table 24: **Descriptive statistics: inter-vivos transfers by maximum residential parent education.**

Positive Transfers only								
	Rent				No Rent			
age	**mean**	**median**	**stand.dev.**	**obs**	**mean**	**median**	**stand.dev.**	**obs**
LHS	5,050	5,115	1,721	1,055	944	383	1,887	349
HSG	4,978	5,293	1,978	6,070	1,032	479	1,913	2,611
CG	5,108	5,293	2,353	8,744	1,383	500	2,613	3,889
Overall	5,054	5,282	2,179	15,869	1,227	486	2,342	6,849
Whole sample								
	Rent				No Rent			
LHS	3,675	5,014	2,686	1,450	227	0	1,009	1,450
HSG	3,761	5,014	2,745	8,035	335	0	1,193	8,035
CG	3,833	5,014	3,007	11,651	462	0	1,645	11,651
Overall	3,795	5,014	2,889	21,136	398	0	1,452	21,136

Calibrated Parameter Values for Benchmark

Parameter	*Value*	*Moment to Match*
J	79	Max model age (between age 16 and age 95)
j^{RET}	50	Maximum years of working life
$\{\zeta_j\}$	-	Survival rates (from US Life Tables)
β	0.962	Match Wealth-Income ratio (3.5) excluding top 1%
ϕ_{HS}		Direct cost of High School: 1% of post-tax median income
ϕ_{COL}		Direct cost of College: 10% of post-tax median income
$\underline{a}^{PRV}$		Match fraction of households with net worth ≤ 0
α	0.35	Capital share in total output
δ	6.5%	Depreciation rate
p^e	16.4%	Pension replacement rate (same for all edu. groups)
t_l	27%	Labor income tax (flat)
t_K	40%	Capital income tax (flat)

Table 25: Value of Parameters Calibrated in Benchmark

E Numerical results

Here we report some details about the numerical analysis and calibration of the benchmark economy.

Table (25) reports the values for a set of parameters which are not directly estimated.

The values of the linear utility terms $\kappa(\theta)$ are available upon request (there are 10 in total: 5 for High School and 5 for College. We omit for sake of brevity).

Table (27) reports the results of a subsidy conditional on current assets of youths. Selected outcomes of the policy in both partial and general equilibrium are reported.

Table (28) documents the outcomes in terms of education achievement that are associated to the means-tested subsidy under consideration.

Table (29) reports the main effects of ability dependent grants.

The equilibrium education shares associated to the ability-tested conditional subsidy are reported in table (30).

We explore the effects of changing the tax rate on HC accumulationa, efficiency and inequality. Table (31) reports the main results.

Benchmark Economy			
	LHS	*HS*	*COL*
Aggregate HC	3.28	11.08	3.99
Returns to HC	.79	1.0	1.25
Average wage	.76	1.19	1.91
Average ability	-.34	.06	.18
Average income	.28	.44	.73
Median income	.37		
Aggregate output	39.4		
Educ. Distrib. (source: NLSY 79 and CPS)			
Ability	*LHS*	*HS*	*COL*
Quintile 1 (low)	.86	.13	.01
Quintile 2	.25	.70	.05
Quintile 3	.11	.78	.11
Quintile 4	.03	.73	.24
Quintile 5 (high)	.01	.54	.45
Aggregate	.25	.58	.17

Table 26: Characteristics of the benchmark economy

Means-tested (assets) subsidy: inequality and output						
Case 1: zero correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
	LHS	*HS*	*COL*	*LHS*	*HS*	*COL*
Aggregate HC	3.48	8.90	6.17	3.30	10.76	4.31
Returns to HC	.79	1.09	1.11	.79	1.02	1.23
Average wage	.77	1.20	1.86	.77	1.21	1.91
Average ability	-.33	.06	.15	-.34	.05	.21
Average income	.28	.44	.72	.28	.45	.74
Median income	.39			.38		
Aggregate output	40.4			39.7		
Wage tax rate	26.34% (down .66)			26.92% (down .08)		
Case 2: 10% correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
	LHS	*HS*	*COL*	*LHS*	*HS*	*COL*
Aggregate HC	3.40	9.30	5.83	3.36	10.80	4.23
Returns to HC	.80	1.07	1.13	.79	1.01	1.23
Average wage	.76	1.20	1.85	.76	1.20	1.88
Average ability	-.33	.05	.15	-.34	.05	.20
Average income	.28	.44	.71	.27	.45	.73
Median income	.39			.38		
Aggregate output	40.25			39.5		
Wage tax rate	26.43% (down .57)			27.11% (up .11)		

Table 27: Selected outcomes of a means-tested education subsidy (based on current assets of youths).

Means-tested (assets) subsidy: education outcomes

Case 1: zero correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
Ability	*LHS*	*HS*	*COL*	*LHS*	*HS*	*COL*
Quintile 1 (low)	.90	.06	.04	.92	.06	.02
Quintile 2	.27	.61	.12	.20	.78	.02
Quintile 3	.11	.68	.21	.12	.84	.04
Quintile 4	.03	.62	.35	.03	.75	.22
Quintile 5 (high)	.01	.35	.64	.01	.38	.61
Aggregate	.27	.46	.27	.26	.56	.18
Case 2: 10% correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
Ability	*LHS*	*HS*	*COL*	*LHS*	*HS*	*COL*
Quintile 1 (low)	.87	.09	.04	.88	.10	.02
Quintile 2	.26	.62	.12	.23	.73	.04
Quintile 3	.11	.69	.20	.01	.84	.05
Quintile 4	.03	.63	.34	.03	.73	.24
Quintile 5 (high)	.01	.39	.60	.01	.43	.56
Aggregate	.26	.48	.26	.26	.56	.18

Table 28: Education achievement under means-tested education subsidy (based on current assets of youths).

Ability-dependent subsidy: inequality and output

Case 1: zero correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
	LHS	*HS*	*COL*	*LHS*	*HS*	*COL*
Aggregate HC	3.44	10.09	4.92	3.30	10.95	4.15
Returns to HC	.79	1.04	1.18	.80	1.01	1.24
Average wage	.77	1.19	1.93	.77	1.20	1.97
Average ability	-.33	.05	.19	-.34	.05	.22
Average income	.28	.44	.74	.28	.45	.76
Median income	.38			.38		
Aggregate output	39.95			39.64		
Wage tax rate	26.76% (down .24)			26.97% (down .03)		
Case 2: 10% correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
	LHS	*HS*	*COL*	*LHS*	*HS*	*COL*
Aggregate HC	3.37	10.17	4.94	3.35	10.90	4.16
Returns to HC	.79	1.04	1.18	.79	1.01	1.24
Average wage	.76	1.18	1.93	.76	1.19	1.96
Average ability	-.33	.05	.19	-.34	.05	.22
Average income	.28	.44	.74	.27	.44	.75
Median income	.38			.38		
Aggregate output	39.87			39.55		
Wage tax rate	26.72% (down .28)			27.09 (up .09)		

Table 29: Selected outcomes of a education policy (subsidy) based on youths' ability

Ability-dependent subsidy: education outcomes

Case 1: zero correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
Ability	*LHS*	*HS*	*COL*	*LHS*	*HS*	*COL*
Quintile 1 (low)	.91	.08	.01	.91	.09	.00
Quintile 2	.26	.69	.05	.21	.78	.01
Quintile 3	.11	.77	.12	.10	.86	.04
Quintile 4	.03	.73	.24	.03	.78	.19
Quintile 5 (high)	.01	.35	.64	.01	.36	.63
Aggregate	.26	.53	.21	.26	.57	.17

Case 2: 10% correlation of ability and initial endowments						
	Partial Equilibrium			General Equilibrium		
Ability	*LHS*	*HS*	*COL*	*LHS*	*HS*	*COL*
Quintile 1 (low)	.87	.12	.01	.88	.12	.00
Quintile 2	.25	.70	.05	.24	.75	.01
Quintile 3	.11	.78	.11	.11	.86	.03
Quintile 4	.03	.73	.24	.03	.78	.19
Quintile 5 (high)	.01	.34	.65	.01	.35	.64
Aggregate	.26	.53	.21	.26	.57	.17

Table 30: Education achievement under ability-dependent education subsidy.

Ability-dependent subsidy: inequality and output

Case 1: zero correlation of ability and initial endowments							
	Labor tax rate: 25%				Capital tax rate: 33%		
	LHS	*HS*	*COL*		*LHS*	*HS*	*COL*
Aggregate HC	3.30	11.22	4.03		3.28	10.97	3.94
Returns to HC	.80	1.00	1.26		.79	1.00	1.25
Average wage	.77	1.20	1.89		.76	1.19	1.91
Average ability	-.35	.06	.16		-.33	.06	.18
Average income	.27	.44	.72		.28	.44	.74
Median income	.38				.37		
Aggregate output	38.9				39.78		
Capital tax	47.64% (up 7.64)			Labor tax	28.68% (up 1.68)		
Case 2: 10% correlation of ability and initial endowments							
	Labor tax rate: 25%				Capital tax rate: 33%		
	LHS	*HS*	*COL*		*LHS*	*HS*	*COL*
Aggregate HC	3.37	11.15	4.00		3.35	10.90	3.91
Returns to HC	.79	1.00	1.26		.78	.99	1.25
Average wage	.76	1.21	1.83		.75	1.19	1.85
Average ability	-.35	.07	.13		-.33	.07	.15
Average income	.27	.45	.70		.28	.45	.72
Median income	.38				.37		
Aggregate output	38.66				39.59		
Capital tax	48.11% (up 8.11)			Labor tax	28.83% (up 1.83)		

Table 31: Selected outcomes of changes in tax rates.

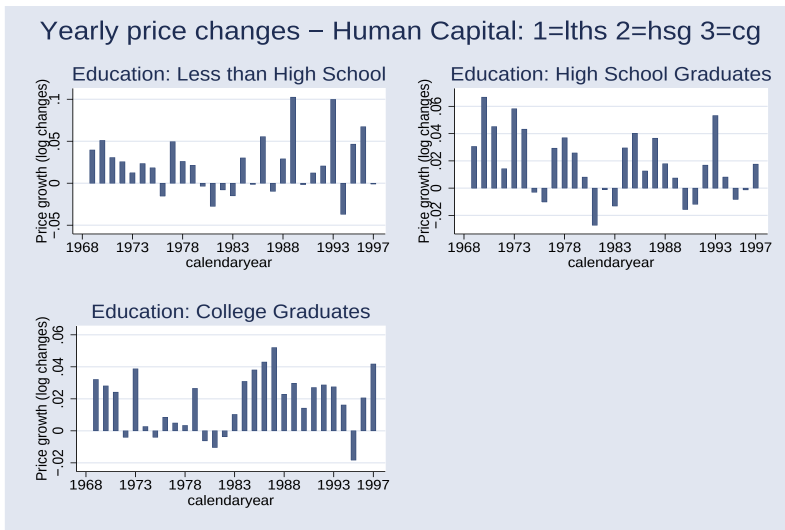

Figure 1: Estimated price growth by education group. PSID 1968-1997

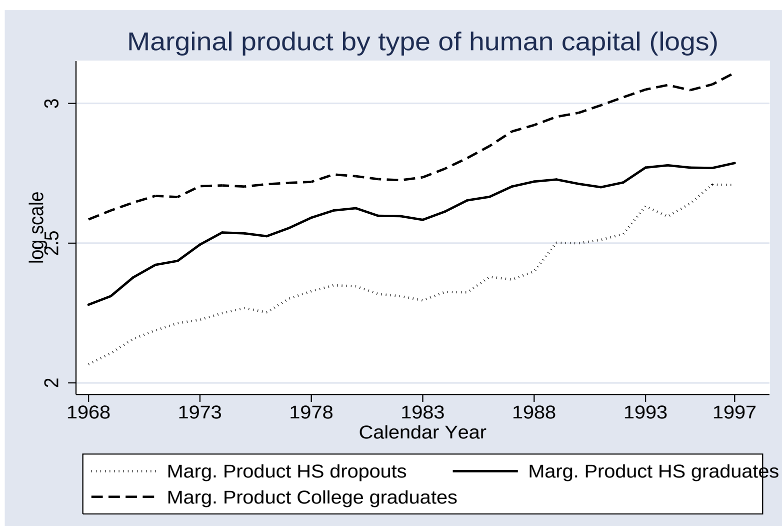

Figure 2: Estimated log price of labor by education group. PSID 1968-1997

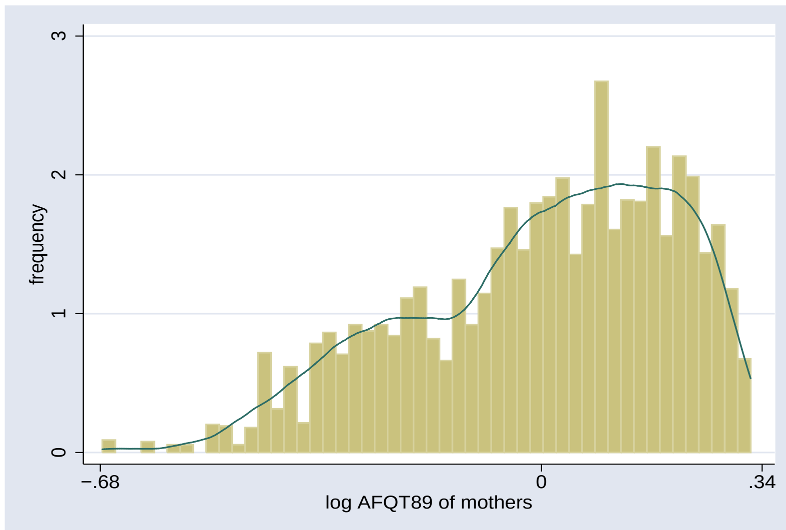

Figure 3: Distribution of AFQT89 (logs), mothers. Histogram and kernel density.

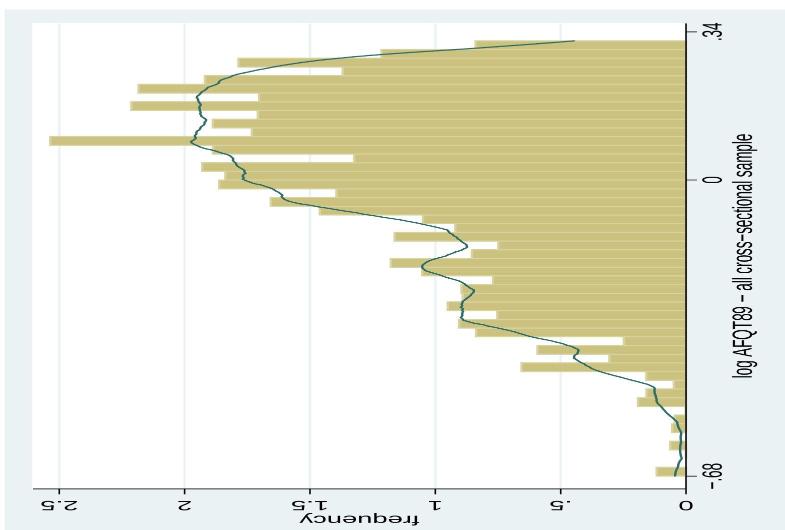

Figure 4: Distribution of AFQT89 (logs), all cross-sectional sample. Histogram and kernel density.

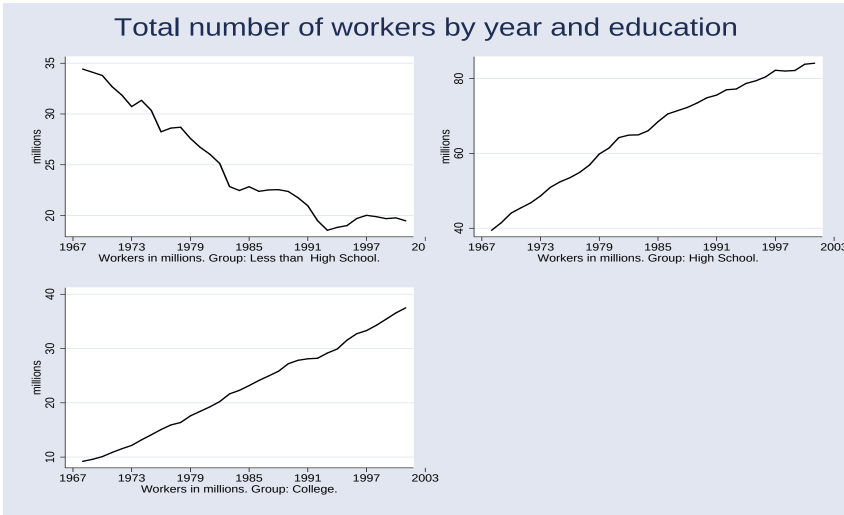

Figure 5: Workers in millions, by education and year

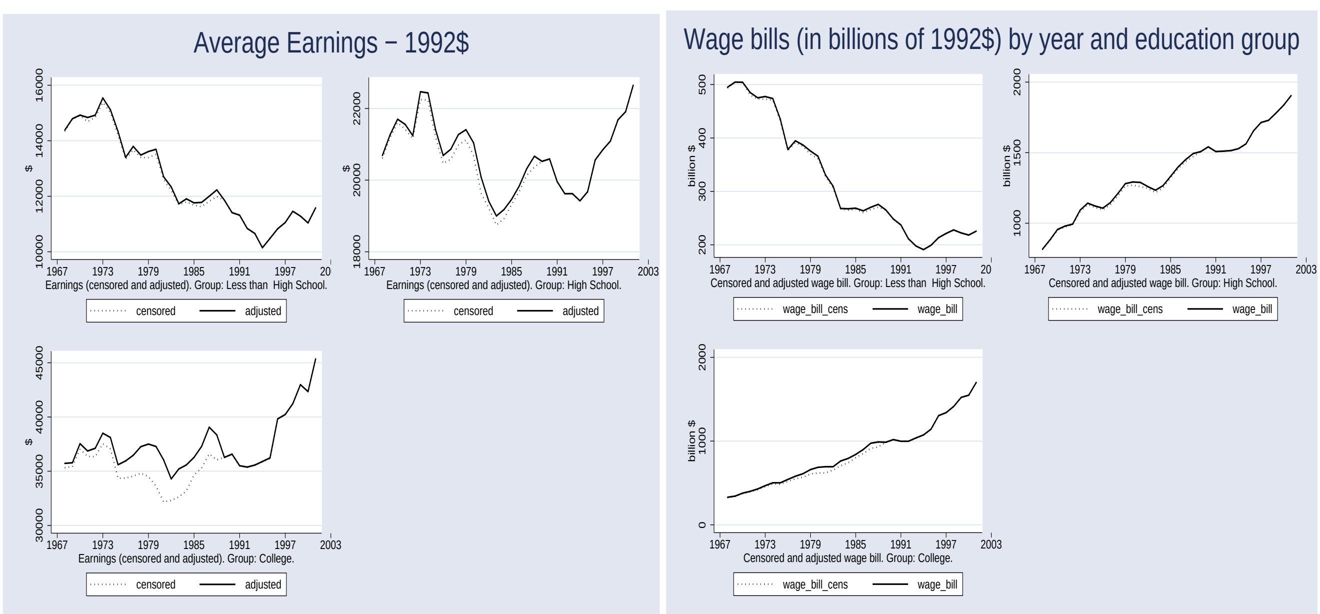

Figure 6: Aggregate wage bills and individual average labor income, by education and year.

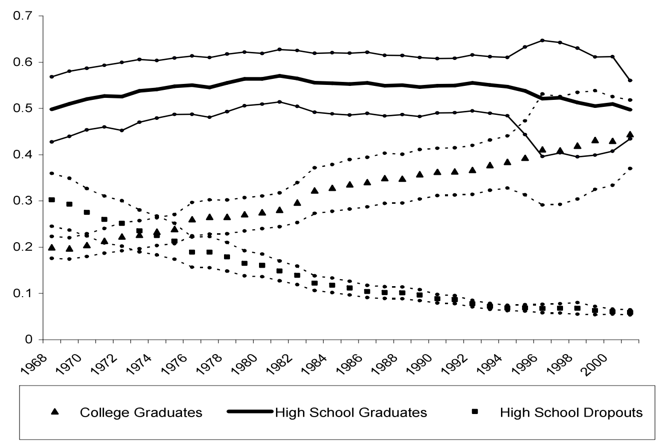

Figure 7: Cobb-Douglas specification for aggregate human capital: estimated shares of human capital inputs by type and year

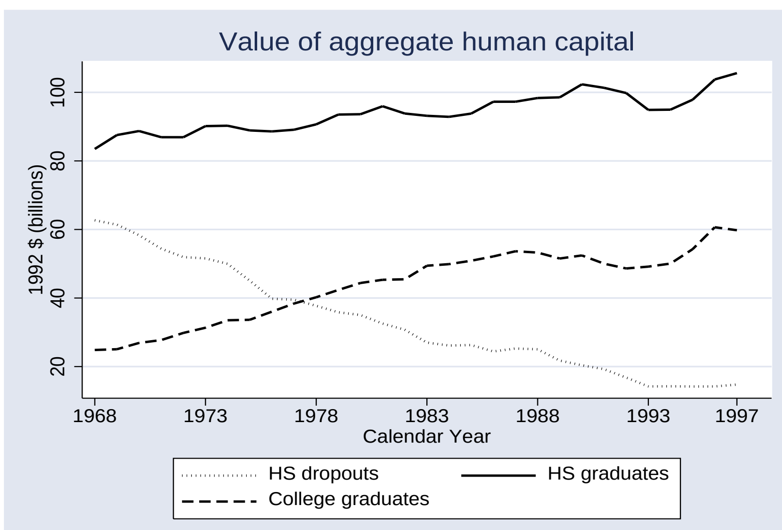

Figure 8: Value of efficiency weighted labor supply (human capital) in billions of 1992 dollars, by education and year

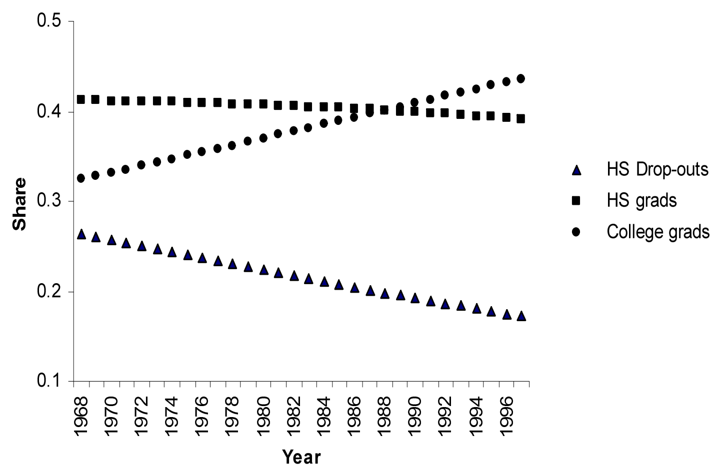

Figure 9: CES specification for aggregate human capital: estimated shares of human capital inputs by type and year