Abstract

This paper examines the effects of alternative policies on the distribution of education in both partial and general equilibrium. Empirical evidence suggests a link between human capital accumulation and wages dispersion, so that policies affecting education outcomes will also have an impact on inequality, productivity and welfare. We build a life-cycle model of labor earnings with endogenous labor supply and education choice, allowing for agents' heterogeneity in several dimensions. PSID data are used to estimate relevant characteristics of education-specific dynamic earnings processes. Numerical simulations are used to compare the effects of alternative policy interventions on education participation, endogenous selection, inequality, life-cycle earnings and wealth profiles. In the current version we experiment with college tuition subsidies. It becomes apparent that while in partial equilibrium such policies can be very effective in increasing education levels and reducing inequality 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 et al. (98a,b). The inclusion of risky returns on labor earnings and the fact that labor supply is endogenous lend additional credibility to the result.
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\(^1\). We examine traditional policies, such as tuition transfers and loan subsidies\(^2\), but we also devise and evaluate alternative forms of policy intervention. The policy experiments are carried out through numerical simulations, with some of the model’s parameters directly estimated from PSID 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 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). Permanent and transitory 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 in-

\(^{1}\) Admittedly, given that labor is bought and sold on spot markets, the demand for labor is always equal to the supply. Alternatively, Acemoglu (2000) studies a model in which the demand for skills changes more than proportionally as a response to the increase in the supply of skilled workers.

\(^{2}\) 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 HC and the opportunity costs of education, altering HC investment decisions.
come, ability or locality, can generate conflicting preferences as to the kind of policies that are most desirable.

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. In fact, statements regarding the effects of policy interventions which ignore price changes induced by such interventions are misleading. Fernandez and Rogerson (1998) provide an interesting application of G.E. modelling to the evaluation of education-finance reform in the US. Later work by Cunha, Heckman, and Navarro (2004) 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 (2001). 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 two twists 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 a persistent efficiency shock. Each agent in our model represents a single individual household, conformably with the empirical analysis we produce.

Recent empirical evidence in Hyslop (2001) indicates that labor supply explains little of the rising earnings inequality for married men, but 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, the response in hours of work to changes in net wage is small for prime age male earners. However, as pointed out by Eaton and Rosen (1980) in their

\[3\] Fernandez and Rogerson (1995) consider ex-ante identical individuals who differ only in income

\[4\] 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.

\[5\] Mortality risk is also explicitly included in the model.

\[6\] Hyslop (2001) also shows that labor supply explains roughly half of the modest rise in female inequality.
seminal work on taxation and HC accumulation, even if taxes have only a limited impact upon the *quantity* of hours worked it is possible that they have an important effect on their *quality*, intended as the type of human capital. This happens because tax changes can alter the incentives for education. Moreover, even if individual labor supplies do not deviate much from some average levels, it is the case that average levels differ substantially between education groups. For given market prices, work effort represents the intensity of human capital utilization and individuals can self-select into education groups according to their preference for leisure\(^7\). Labor supply, therefore, represents an effective channel of adjustment to labor price signals and an important determinant of the relative variation in skill prices\(^8\).

The other crucial twist 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. Problems of moral hazard can be extremely severe when insuring labor risk because idiosyncratic shocks and individual ability can be partially or completely unobservable to third parties. Given these problems the market is not likely to provide insurance. Using a multiplicative form of earnings risk\(^9\) 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\(^10\); 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.

We model three levels of education obtained through formal schooling and corresponding to three types of HC which enter the production technology\(^11\). Education and employment

---

\(^7\)This selection in our model happens through permanent unobserved characteristics.

\(^8\)Consider, for example, taxes on labor earnings which reduce the return to HC investment but also the opportunity cost of being in education represented by foregone earnings. When differences in lifetime labor supply between education groups are present, the two effects are weighted by the relative intensity of HC utilization in the appropriate education group.

\(^9\)They multiply education specific earnings by a random variable.

\(^10\)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.

\(^11\)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 mutually exclusive in each period. Foregone earnings and tuition charges are the direct costs of schooling, and a utility cost comes in the form of reductions in leisure when studying.

Agents can accumulate real assets. We experiment with alternative ways to set the initial conditions for assets. We also consider different levels of correlation between ability and initial assets holdings.

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. 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 Overview

We consider an economy where a unique good is produced, and it can be either consumed or used as physical capital. We specify an overlapping generations general equilibrium model for this economy that focuses on education and labour supply. Consumers maximise an intertemporal utility function over their finite life-cycle, with respect to education, labour supply and consumption/savings. Agents can accumulate assets representing ownership shares on physical capital. They have a maximum lifetime and they plan to consume their entire assets. However they may die before that leaving accidental bequests. The maximum possible lifetime is 99 years. Individuals can work up to 65 but not beyond. They can however decide not to work before that. Retirement is financed by the accumulated assets. The population consists of 99 overlapping generations, each with an ex-ante identical distribution of heterogeneity.

Young and old households are not linked in any direct way. Bequests are pooled together and redistributed to all newly born individuals according to the steady state equilibrium wealth distribution. This reflects both inter-vivo transfers for education and actual bequests. Gale and Scholz (1994) show that inter vivos transfer for education represent only a part of total bequest. We ignore this issue in this paper and redistribute assets only among the youngest. Of special interest is the case when the initial distribution of wealth replicates the distribution prevailing among those who die in each period, as this has a realistic accidental bequest interpretation.

This can be thought of as a shortcut to incorporate the effect of parental background on ability formation, as extensively documented in the literature, see Heckman and Carneiro (2003) for a review).
quests. Since we assume that assets must always be non-zero (liquidity constraints), these transfers are the only source of funding for education, other than possible government transfers. Education can only take place at the beginning of the life-cycle and the individual can attain one of three education levels, corresponding to less than high school, high school, and college. The costs of education consist of the opportunity cost, tuition fees net of any government subsidy and the psychic/utility costs of education. In addition individuals are endowed with different abilities which lead to different efficiency units of human capital and thus earnings. Thus wage differences among people are the consequence of differences in education (between group inequality) and differences in labor efficiency (within group inequality). People are perfect substitutes within schooling groups, regardless of their individual efficiency.

There is no aggregate uncertainty in the model. Once out of school the individual has to decide on the proportion of his/her time to devote to work and on consumption. All these decisions take place in an incomplete markets environment. The individual cannot borrow against human capital and faces uninsurable idiosyncratic wage shocks.

There is an aggregate production function with four inputs: The three levels of human capital, measured in total efficiency units supplied and physical capital. We solve the model as a closed economy with the interest rate determined endogenously.

The model is partly estimated and partly calibrated. First we estimate a wage process and extract relative prices for our three human capital measures. This allows us to compute the total supply of efficiency units of human capital in each of the three groups. From the estimation of the wage process we also estimate the stochastic process of wages, net of measurement errors which we take to be the process of uncertainty facing the individual.\textsuperscript{14} The stochastic process of wages is taken to be education specific.

Next we estimate the aggregate production function which is taken initially to be Cobb-Douglas. However, it proved to be quite difficult to obtain reliable estimates, so we carry out sensitivity analysis based on a number of different production function structures. Moreover, since there seems to have been a change in the production structure mainly due to skill biased technological change we use as our basis for the Cobb-Douglas specification the average shares over the period.

\textsuperscript{14}This may overestimate the degree of uncertainty; see Cuhna, Heckman and Navarro, (2004).
To obtain the parameters characterising preferences we use risk aversion coefficients taken from the literature and then we set the preference for labour supply to match the proportion of people working in the economy. Given these parameters we then calibrate the utility costs of education to match the proportions in each education group during 1978-82. The individual discount rate is calibrated to match the ratio of physical capital over total output.

2.2 Individual preferences

We use the index $j$ to denote age. Agents have a probability to survive in each period, denoted as $s_j$, which is decreasing in age. When annuity markets are absent we use a random bequest mechanism to redistribute left over assets based on the prevailing equilibrium assets’ distribution. Negative borrowing limits open up the possibility that people dying prematurely can be in debt\textsuperscript{15}.

Agents face educational choices based on returns and costs, which depend on age, asset holdings, permanent characteristics and labor shocks. Over the life cycle they choose the labor supply and consumption path maximizing expected lifetime utility.\textsuperscript{16}

The period utility $u(c, l)$ is concave in consumption $c$ and leisure $l = (1 - n)$; it satisfies standard regularity conditions and in particular the Inada conditions. The education level is denoted by $e$, it takes three values, with $e = e_1$ the lowest and $e = e_3$ the highest. Permanent (unobserved) individual characteristics are denoted by $\theta$ and distributed over the domain $[\theta_{\text{min}}, \theta_{\text{max}}]$. We also assume that the distribution of ability $\theta$ is independent of time. We denote by $\{z\}_{j=1}^j$ the sequence of uninsurable idiosyncratic shocks. Their law of motion is summarized by a stationary transition function $\pi$ denoted as $\pi_{z_{j+1}|z_j} = \pi\{z_{j+1} | z_j\}$.

While in school individual utility is given by $u(c_j, f(\theta))$, where the function $f(\theta)$ reflects the psychic costs of schooling which may be thought of as leisure costs but may include other aspects of effort and like or dislike of the education process - hence we do not bound $f$ to lie in the unit interval. These costs will depend on ability with the idea that more able individuals will suffer lower costs.

Given some initial values $\bar{x}_1$ for the state variables, household/individual utility over

\textsuperscript{15}Yaari (1965) considers this case explicitly and proves that, with functioning credit markets not making systematic losses, the budget constraint must be such that individual can never go short on assets. We prefer to ignore this issue and let interest rates adjust appropriately.

\textsuperscript{16}A more detailed discussion of the implications of non-convexities due to binary decision variable can be found in Gallipoli (2004). Existence and uniqueness results for the solution of the model are spelled out for the finite life cycle case.
sequences of consumption and leisure, \( c = \{c_1, \ldots, c_j\} \) and \( l = \{l_1, \ldots, l_j\} \), as of age 1 is denoted \( U(\bar{x}_1, c, l) \) and can be written as the expected discounted sum of period utilities

\[
U(\bar{x}_1, c, l) = E_{x \in Z} \left\{ \sum_{j=1}^{j_{edu}} S_j \beta^{j-1} [d_j u(c_j, f(\theta)) + (1 - d_j) u(c_j, l_j)] + \sum_{j=3}^{j} S_j \beta^{j-1} u(c_j, l_j) \right\}
\]

where \( d_j \) is a binary variable which is 1 if the agent is in education and 0 otherwise, \( j_{edu} \) denotes the last period of education, \( S_j = \left( \prod_{i=1}^{j} s_i \right) \) denotes the probability of surviving to age \( j \) and \( \beta \) is the intertemporal discount factor. For the first three periods the individual may decide to be in education which is why there are two alternative forms for the utility function depending on the action they take. We restrict all education choices to take place in the first three years of life. Note that once in education \( f(\theta) \) is fixed and only depends on ability \( \theta \).

The period budget constraint is given by

\[
c_j + a_{j+1} = [1 + r (1 - \tau_k)] a_j + w e^\epsilon_j n_j (1 - \tau_{ne}) (1 - d_j) - (D_e - T_e) d_j
\]

where and \( a_j \) denotes individual asset holdings\(^{18}\) at age \( j \) and \( r \) is the riskfree interest rate. For the purposes of policy analysis we distinguish between the taxation of capital income \( \tau_k \) and the taxation of labour income \( \tau_{ne} \).\(^{19}\) the \( D_e \) is the direct cost of schooling and \( T_e \) summarizes government subsidies towards education \( e \). The term \( \epsilon^j \) is individual labor efficiency, and \( \epsilon_j \) is defined as

\[
\epsilon_j(\theta, e, z) = \theta + \xi_j(e) + z_j
\]

where \( \xi_j(e) \) is an education-specific age profile.

### 2.3 Solving the individual problem

The individual’s problem may be solved recursively by backwards induction. Denote by \( x_j \) the value of the state variables in period \( j \) and by \( W_j(x_j) \) the optimum value function at age \( j \). The state includes the current value of the shock \( z \), which is assumed to arrive at the beginning of the period, as well as permanent characteristics, past values that are relevant for predicting future outcomes and of course current assets and education levels. Following

---

\(^{18}\)Individual asset holdings satisfy: \( a_j \geq a_{min} \) for every \( j \) and \( a_{j+1} \geq 0 \). The first inequality is a borrowing constraint, whereas the second is a transversality condition for agents reaching age \( j \).

\(^{19}\)Heckman (1976) first noted the importance of this distinction when considering investments in Human Capital. Changing the cost of intertemporal substitution will affect investment decisions.
the third year of life when no more education choices can be made the individual chooses consumption and labour supply to solve the problem

\[
W_j(x_j) = \max_{c,n} \left\{ u(c_j, 1 - n_j) + s_j \beta \int_Z \pi_{z_{j+1}|z_j} W_{j+1}(x_{j+1}) \, dz_{j+1} \right\}
\]

subject to the budget constraint \(2\) with \(d_j = 0\) and subject to the constraint \(a_{t+1} \geq A\) where \(A\) is some exogenous minimum level of assets, possibly zero or possibly negative.

Previously, during the first three years of life, the individual’s problem is complicated by the fact that she/he needs to decide on whether to obtain education. In this case the problem solved is

\[
W_j(x_j) = \max_{c,n,d} \left\{ d_j u(c_j, f(\theta)) + (1 - d_j) u(c_j, l_j) + s_j \beta \int_Z \pi_{z_{j+1}|z_j} W_{j+1}(x_{j+1}) \, dz_{j+1} \right\}
\]

subject to the budget constraint, the asset constraint mentioned above and subject to the constraint that \(n = 0\) if \(d_j = 1\), i.e. we do not allow for work and education at the same time.

**2.4 Aggregate variables**

We study equilibrium allocations and assume a stationary population. The aggregate states of the economy are physical capital \(K\) and efficiency-weighted aggregate labor supplies (referred to as human capital aggregates) \(H_1\), \(H_2\), and \(H_3\). In steady state there will be no change in the relative wages between the three different types of human capital. However, the feature of the model that allows for such relative price variation is key because as the policy alters the supply of each education group, relative prices will change and this will lead to different steady state levels of supply for \(H_1\), \(H_2\), and \(H_3\). The total stock of human capital of type \(e\) is the sum of the efficiency weighted individual labor supplies of type \(e\) which are defined by

\[
h_j(\theta, e, z, a) = \epsilon^{e_j(\theta, e, z)} n_j(\theta, e, z, a)
\]

The aggregate states determine the relative prices in the economy. The demographics are stable, so that age \(j\) agents make up a constant fraction \(\zeta_j\) of the population at any point in time. The \(\zeta_j\) values are normalized to sum up to 1 and are such that \(\zeta_{j+1} = s_j \zeta_j\).
2.5 Market structure

Our setup is an incomplete markets one where idiosyncratic risk cannot be insured, other than by self-insurance through precautionary savings. However we also impose liquidity constraints, which will be biting for the more able young people, unless they have inherited wealth. There is no aggregate uncertainty. The unique physical good is used as numeraire. All unemployment in the model is voluntary.

We consider model specifications with and without missing annuity markets. In the simulations in which we allow non-zero asset holdings for new borns, we also experiment with different degrees of correlation between initial wealth and permanent characteristics $\theta$.20

2.6 Technology

Firms maximize profits using a CRS technology and set wages competitively. The aggregate technology employs physical and human capital and is denoted as $F(H, K)$ with $H = \{H_1, H_2, H_3\}$. The relationship between human capital factors $(H)$ and physical capital is expressed as a Cobb-Douglas:

$$F(H, K) = \bar{A}H^{1-\alpha}K^{\alpha}$$

(8)

$\bar{A}$ is a TFP coefficient and the general, unconstrained definition of the HC input is

$$H = \left\{A_1H_1^{\rho_1} + [A_2H_2^{\rho_2} + A_3H_3^{\rho_2}]^{\frac{\rho_2}{\rho_1}} \right\}^{\frac{1}{\rho_1}}$$

which allows for the elasticity of substitution to differ between unskilled labour $H_1$ and the other two inputs. In practice we have not been able to obtain meaningful estimates of this quite general production function and we estimate a Cobb-Douglas specification with $\rho_1 = \rho_2 = 0$. However we do use this general specification as a basis for sensitivity analysis.

The equilibrium conditions require that marginal products equal pre-tax prices so that $w_e = \frac{\partial F}{\partial H_e}$ for any education level $e$, and $r + \delta = \frac{\partial F}{\partial K}$, where $\delta$ is the depreciation rate for capital.

20 Imposing different patterns of dependence between such marginal densities turns out to be useful if ability is correlated with socio-economic background factors such as family wealth.
2.7 Government

Government has revenues from proportional taxation of labor and asset income at respectively $\tau_n$ and $\tau_k$ rate, and uses part of these revenues to subsidize education via a transfer $T_e$. We call $G$ the residual non-education general government expenditure and assume that $G$ is lost in non-productive activities. The government’s behaviour is fully described by the budget constraint, which requires that expenditures equal revenues obtained from taxation\textsuperscript{21}. The government has a balanced budget in each period.

3 Equilibrium

We use the notion of a stationary recursive competitive equilibrium in Lucas (1980)\textsuperscript{22}. Let $(X, F(X), \psi_j)$ be an age-specific measure space with state space $X$ and $F(X)$ be a $\sigma$-algebra on $X$.

Given some state vector $x \in X$, a stationary equilibrium for this economy is a set of decision rules $d_j(x)$, $a_{j+1}(x)$, $c_j(x)$, $n_j(x)$, value functions $W_j(x)$, price functions $w_e$, $r$, densities $(\psi_1, ..., \psi_j)$ and $(\zeta_1, ..., \zeta_j)$, and a law of motion $Q$, such that:

1. $d_j(x)$, $a_{j+1}(x)$, $c_j(x)$ and $n_j(x)$ are optimal decision rules and solve the household’s problem;
2. $W_j(\theta, e, z, a)$ are the associated value functions;
3. Firms employ inputs so that
   \[ w_e = F_{He} \quad \text{for } e \in \mathcal{E} \]
   \[ r + \delta = F_{K} \]
4. $\psi_j(x)$ is a stationary measure, that is $\psi_j(F) = Q(F, \psi_j)$, where $Q(\cdot, \cdot)$ is the law of motion of $\psi_j(\cdot)$ and is generated by the optimal decisions $d_j(x)$, $a_{j+1}(x)$, $c_j(x)$,\textsuperscript{23}

\textsuperscript{21}The government budget constraint is

\[ G + \sum_j \zeta_j \int_X T_e d_j(x) \, d\psi_j(x) = \]
\[ = \sum_j \zeta_j \int_X [1 - d_j(x)] \tau_n w_e h_j(x) \, d\psi_j(x) + \sum_j \zeta_j \int_A \tau_k a_{j+1} \, d\psi_j(a) \]

We assume that the government has a balanced budget in each period.

\textsuperscript{22}Our model satisfies the conditions for defining a measure $\psi_j$, such that $\mu(x, j) = \zeta_j \psi_j(x)$ is stationary as a function of the markov process $\pi\{z_{j+1} | z_j\}$ and of the decision rules $d_j(x)$ and $a_{j+1}(x)$, where $x$ is an element of the state space.

\textsuperscript{23}Given $\zeta_j$, also $\mu(x, j) = \zeta_j \psi_j(x)$ is a stationary measure.
5. The good, asset and labour markets clear.\textsuperscript{24}

The goods market clearing equation is derived by integrating the individual budget constraint.

4 Estimation

With the exception of the parameters for intertemporal substitution the remaining ones are obtained by a combination of estimation and calibration using data from the US.

We thus estimate the wage process, the distribution of unobserved heterogeneity and the aggregate production function. We then calibrate the costs of education function $f(\theta)$ to fit the distribution of education by ability in the four year period from 1978 to 1982. Our sources of data are the CPS, the PSID and NIPA.

4.1 Estimating wage equations: skill prices and age profiles

An important characteristic of the model is that the three types of human capital represent different inputs to the production function, not necessarily perfectly substitutable and may have relative prices that vary over time in response to changes in either supply or demand for skills. So as to be able to simulate our model, we need to extract from the data the distribution of unobserved heterogeneity affecting wages and education choices as well as the stochastic process of the shocks.

We start by specifying an education specific wage equation for individual $i$ wages in period $t$, $w_{eit}$

$$\ln w_{eit} = w_{et} + g_e(\text{age}_{eit}) + u_{eit}$$

where $w_{et}$ represents the log of the aggregate price of human capital for education group $e$ and where $g_e(\text{age}_{eit})$ is the education specific profile of wages. The unobservable component $u_{eit}$ is specified to be

$$u_{eit} = \theta_i + z_{eit} + m_{it}$$

where $\theta_i$ represents unobserved fixed effects, $z_{eit}$ is the (persistent) shock to wages and $m_{it}$ is measurement error, assumed iid. Self-selection implies that fixed effects are correlated with both education decisions and observed wage rates. However, a within groups transformation

\textsuperscript{24}Equilibrium definitions in the asset and good markets must include cross border asset holding $FX$ if the interest rate $r$ is constant.
eliminates the source of self-selection and identifies the changes in the returns to education over time as well as the way wages grow with age by education group. Thus we estimate by OLS

\[
(\ln w_{eit} - \ln \overline{w}_{eit}) = (\ln w_{et} - \ln \overline{w}) + g(abge_{eit}) - \overline{g(abge_{eit})} + (u_{eit} - \overline{u}_{eit}) \tag{11}
\]

where the upper-bar denotes an (individual) time average and where \( g \) is a polynomial of order two for the lowest education group and of order 4 for the two higher education groups. The term \((\ln w_{et} - \ln \overline{w})\) is modelled as time dummies. The residuals from this equation can be used to identify the persistence of wage shocks, and we discuss this below.

4.2 Wage data and results

For the estimation of wage equations we use longitudinal data from the PSID. The sample is based on annual interviews between 1968 and 1997 and on bi-annual interviews from 1999 onwards. We do not use individuals associated with the Census low income sample, the Latino sample or the New Immigrant sample and focus instead on the SRC core sample, which did not suffer any systematic additions or reductions between 1968 and 2001 and was originally representative of the US population.

The main earnings’ variable in the PSID refers to the head of the household\(^\text{25}\) and is described as total labor income of the head\(^\text{26}\). 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). A detailed description of our sample selection is reported in the appendix: in brief, we select heads of household aged 25-60 who are not self-employed and have positive labor income for at least 8 (possibly non continuous) years.

The age polynomials from the wage equation are presented in table (1).

\(^{25}\)In the PSID the head of the household is a male whenever there is a cohabiting male/female couple. Women are considered heads of household only when living on their own. We do not address the related sample issues explicitly, but any gender effects are likely to be captured in the ability estimates.

\(^{26}\)This 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: Age polynomials’ coefficients

<table>
<thead>
<tr>
<th>Dependent variable: log hourly earnings</th>
<th>coeff.</th>
<th>point estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education=LTHS</td>
<td>age</td>
<td>0.0412505</td>
<td>0.0081143</td>
</tr>
<tr>
<td></td>
<td>age²</td>
<td>-0.0004179</td>
<td>0.0000905</td>
</tr>
<tr>
<td>Education=HSG</td>
<td>age</td>
<td>0.4928285</td>
<td>0.1071015</td>
</tr>
<tr>
<td></td>
<td>age²</td>
<td>-0.0162768</td>
<td>0.0039883</td>
</tr>
<tr>
<td></td>
<td>age³</td>
<td>0.0002413</td>
<td>0.0000644</td>
</tr>
<tr>
<td></td>
<td>age⁴</td>
<td>-1.34e-06</td>
<td>3.82e-07</td>
</tr>
<tr>
<td>Education=CG</td>
<td>age</td>
<td>0.8697329</td>
<td>0.1560285</td>
</tr>
<tr>
<td></td>
<td>age²</td>
<td>-0.0282</td>
<td>0.0058548</td>
</tr>
<tr>
<td></td>
<td>age³</td>
<td>0.0004149</td>
<td>0.0000953</td>
</tr>
<tr>
<td></td>
<td>age⁴</td>
<td>-2.30e-06</td>
<td>5.69e-07</td>
</tr>
</tbody>
</table>

Figure (1) plots the age profiles implied by the polynomial estimates for different education groups.

Figure 1: Age profiles of labor efficiency by education group - age on the horizontal axis

By fitting the within group specification of the wage (log of hourly earnings) equation we also obtain ln\( \hat{w}_{et} \), estimates of the growth of log-price of labor by education and year, which are plotted in figure (2). The time effects have a natural interpretation as time varying prices of skills associated to different education groups. The fact that the relative prices vary this much is a key justification for treating the different education levels as different types of human capital.
4.3 Permanent characteristics and their distribution

For the purposes of simulation we require the unconditional distribution of ability as re-
lected by the fixed effect \( \theta_i \). We thus use the estimate

\[
\hat{\theta}_i = \frac{\sum_{t=1}^{T(i)} \ln w_{it} - \ln \bar{w}_t - g(\text{age}_{it})}{T(i)}
\]

where \( T(i) \) is the total number of observation available on agent \( i \). If we assume that the unconditional distribution of ability has not changed over the time period covered by our sample, we can use the estimated fixed-effects as an estimate of the \( \{\theta_i\} \) distribution over the working population.

Figure 3: Estimated density of log fixed effects for small (67-93) and large (67-00) samples

In figure (3) we report the empirical frequencies of \( \hat{\theta} \) obtained by aggregating both cross-
sectionally and longitudinally.
The estimation variance of \( \hat{\theta}_i \) will inflate the overall variance of unobserved heterogeneity. To mitigate this problem we have traded off some representativeness by taking individuals who are observed for at least five periods. As a further check we compare our distribution of ability with the one-off IQ reported in the 1972 sample. Figure (4) reports the measured IQ densities for the whole 1972 sample and a selected sub-sample based on our criterion. It seems that both the IQ density and the estimated fixed effect density exhibit a long left tail.

### 4.4 Analysis of labor efficiency shocks

We now use the residuals from the wage equation to estimate our assumed stochastic process for wages. First note that we can treat as observable the following:

\[
    u_{eit} = \ln w_{eit} - g_e \left( \text{age}_{eit} \right) - \ln w_{et} - \theta_i
\]

We assume that \( u_{eit} \) can be decomposed into two components

\[
    u_{eit} = z_{eit} + m_{eit}
\]

where \( z_{eit} \) is an autocorrelated error process and \( m_{eit} \) is classical measurement error \( iid \,(0, \sigma_{em}^2) \) and where \( \{z_{eit}\}_i \) is a autocorrelated process with education specific parameters

\[
    z_{eit} = \rho_e z_{eit-1} + \varepsilon_{eit}
\]

in which \( \varepsilon_{eit} \sim iid \,(0, \sigma_{e}^2) \), we can achieve identification of the autoregressive parameters in one of several ways. With an external estimate of the measurement error variance we can use the following expressions to estimate \( \sigma_{e}^2 \) and \( \rho \):

\[
    \rho = \frac{COV \left( z_{eit}, z_{eit-1} \right)}{VAR \left( z_{eit} \right)} = \frac{COV \left( u_{eit}, u_{eit-1} \right)}{VAR \left( u_{eit} \right) - VAR \left( m_{eit} \right)}
\]
\[ \text{VAR}(u_{eit}) = \frac{\sigma_e^2}{1 - \rho_e^2} + \sigma_m^2 \]

where we can substitute the covariances of \( u \) with sample analogues. However it is also possible to use the variance of \( u \) and its first two auto-covariances to identify the variance of the measurement error as well. Thus we have that

\[
\rho_e = \frac{\text{COV}(u_{eit}, u_{eit-2})}{\text{COV}(u_{eit}, u_{eit-1})}
\]

and the rest follows immediately. In practice we replace the error terms \( u \) with the residuals for the wage regression as defined in (12).

### 4.4.1 Results for the wage process

We present estimates of the autoregressive coefficients obtained using external estimates of measurement error by French (2000), who provides a lower and a upper bound estimate for measurement error (respectively 0.0172 and 0.0323). Our results are based on an average of the two. The (bootstrapped) standard errors are in parentheses.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.651</td>
<td>0.557</td>
<td>0.608</td>
<td>0.584</td>
</tr>
<tr>
<td>(0.130)</td>
<td>(0.042)</td>
<td>(0.058)</td>
<td>(0.034)</td>
</tr>
</tbody>
</table>

The estimated values for \( \hat{\rho}_e \) seem indicate that group 2 (High school graduates) experience the lowest earnings’ risk.\(^{27}\) These findings are apparently in contrast with some of the recent literature, among others Storesletten, Telmer, and Yaron (2002) and Meghir and Pistaferri (2004). However, using the upper estimate of measurement error we get parameters very close to one. Of course, with near unit-root persistence of wage shocks the identification of fixed effects suffers from severe initial conditions problems. An alternative on which we are currently working is using an exogenous distribution of ability based on 1972 test scores while assuming unit-root behaviour of labor efficiency shocks.

\(^{27}\)The point estimates of \( \sigma_e^2 \) for the pooled case is 0.01156, whereas for the LTHS, HSG and CG cases is, respectively, 0.01040, 0.01250 and 0.0098. We also perform tests based on the autocovariance structure of the AR(1) residuals, in order to check for the goodness of the specification. They validate the specification choice and are available from the authors.
4.5 Using CPS to obtain data for the aggregate production function.

Estimation of the aggregate production function requires the total wage bills for each of the education groups, and in the general CES case we also require measures of human capital in each of these groups. We use the March supplement of the Current Population Survey (CPS) to obtain these. 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 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 the growth of prices over time and we can normalise one of the initial prices to one. Now note that we have one degree of freedom. We can set the initial relative price of high school and of college graduate labour and we can then 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 rationalised either with high returns and high costs or low returns and low costs. The particular normalisation we choose will not affect the simulation of the policy changes.

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 nonresponse and undercoverage.

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.

28The 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 Labor Statistics (BLS) using data from the Current Population Survey (CPS).
29Eliminating all zero-earnings observations rules out the possibility to incorporate employment risk, which is possibly an important source of risk.
30This 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.
Figure (5) reports the number of people working in each year by education group, as reported by the CPS.

![Graphs of workers in millions by education group and year](image)

**Workers in Millions – 1=lhs 2=hsg 3=CG**

Figure 5: Employed workers in millions, by education and year

It is clear that some strong and persistent trends towards higher levels of education have characterized the sample period.

![Graphs of wage bills and average earnings by education group and year](image)

**Wage Bills (in billion 1992$) – 1=lhs 2=hsg 3=CG**  
**Average Earnings – 1992$ – 1=lhs 2=hsg 3=CG**

Figure 6: Total and average earned labor income, by education and year. Total in billions of 1992 dollars, average in units of 1992 dollars.

Figure (6) plots both the average earnings by year and total wage bills in billions of dollars. 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 (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. 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 and human
capital aggregate, as they never represent more than 5% of the working population in a
given education group (and most of the time much less than that).

Finally, dividing the wage bills by the exponentiated value of the time effects estimated
through the wage equations using PSID data we finally obtain point estimates of the value
of efficiency weighted total labor supply (human capital aggregates) by education and year.
These are plotted in figure (7).

Figure 7: Value of efficiency weighted labor supply (HC) in billions of 1992 dollars, by education and year

Notice that the evolution of human capital over time is non-monotonic, unlike the wage
bills for the two highest education groups. This is due to the large increase in the level
of estimated marginal product of these two factors in the early 1990s, which has grown
proportionally more than the total remuneration of these factors.

4.6 Aggregate Production function

In estimating technology parameters, we start from the relatively easier case of Cobb-
Douglas technology. Let aggregate output \( Y \) be produced through the following technology

\[
Y = \left( H_3^A H_2^{1-A} H_1^{1-A}(1-B) \right)^{1-\alpha} K^\alpha
\]
Using NIPA data we find the share of capital $\alpha$ to be between 0.3 and 0.35, depending on whether we correct for housing stocks. The share parameters $A$ and $B$ can be easily expressed as a function of the aggregate wage bills. If we apply this procedure separately for each year we can pinpoint the evolution of these functions over the sample period.

Figure 8: Labor shares in human capital input of technology, computed using Cobb-Douglas specification (with bounds equal to +/- 2 standard errors). Period: 1968-2000. Larger bounds after 1996 are due to changes in top-coding of income in the CPS.

Figure (8) reports the value of the share parameters (with bounds equal to 2 standard errors) for the shares associated to each human capital variety. In figure (8) the line that is increasing over the sample periods represents $A$, whereas the downward sloping one represents $(1 - A) (1 - B)$. The almost flat line on top is $(1 - A) B$.

The time average of such shares is $A = 0.33$, $(1 - A) B = 0.54$ and $(1 - A) (1 - B) = 0.14$. The evolution of the college graduates labor share over time more than doubles (from 0.2 to 0.4) whereas the share of less-than-high-school labor falls dramatically from over 0.30 to roughly 0.06. These findings confirm what we found in terms of marginal products of labor using PSID data: major shifts in technology have taken place between the late 1960s and the end of the century.

We follow up our initial findings by performing some additional inference on the technology parameters. In order to do this we first approximate the total human capital factor $H = F \{H_1, H_2, H_3\}$ by combining NIPA and CPS data on wage bills and physical capital$^{31}$ and then use a 2-step GMM method which controls for endogeneity and serial correlation of TFP to estimate the parameters. We use lagged shares as instruments.

$^{31}$Details of the model can be provided by the authors on request.
The results of the GMM estimation of our favourite specification for the log-linearized C-D technology are reported in the following tables (standard errors in parenthesis) for two alternative moment weighting matrix choices (the identity matrix and the optimal matrix.

<table>
<thead>
<tr>
<th>First Step Weighting: Identity Matrix</th>
<th>First Step Weighting: Optimal Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>0.260</td>
<td>0.783</td>
</tr>
<tr>
<td>(0.200)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>0.284</td>
<td>0.790</td>
</tr>
<tr>
<td>(0.207)</td>
<td>(0.123)</td>
</tr>
</tbody>
</table>

We also find that the linear trend included to control for TFP deterministic variation is estimated to be 0.035 and strongly significant. Given the log specification, this is equivalent to an average annual TFP growth rate of roughly 3.5% between 1967 and 1997.

The point estimates for A and B give labor shares very similar to the long-term averages we estimate using the initial Cobb-Douglas computation. The labor shares roughly sum up to one, even though we don’t impose this restriction in the estimation procedure.

5 Simulations

Estimation above has yielded the stochastic process for wages, the distribution of ability, the lifecycle growth of wages and finally the aggregate production function. In the next step we set the preference parameters and we choose the utility cost of education to match as close as possible the proportions completing each of the three levels of education.

5.1 Preferences parameters

Wealth and consumption data are probably not of sufficient quality to estimate a joint model of labour supply, consumption and education choice, particularly because we will have to observe all forms of wealth reliably, together with consumption. Thus we decided to rely on earlier Euler equation estimates for the preference parameters, as well as matching aggregate labour supply levels. Thus we specify the utility function to be of the CRRA type, i.e.

\[ u(c_j, I_j | d_j = 0) = \frac{[c_j^{1-\nu}]}{1-\lambda} \]

\[ u(c_j | d_j = 1) = \frac{[c_j^{1-\nu} f(\theta)^{1-\nu}]}{1-\lambda} \]
The parameters $\nu$ and $\lambda$ of the period utility jointly pin down the intertemporal elasticity of substitution of consumption $\frac{1}{1-\nu(1-\lambda)}$ (ISE) as well as the level of labour supply over the lifecycle. Set the ISE to 0.75 as in Blundell, Browning and Meghir (1994) and Attanasio and Weber (1993). Given this a value of $\nu = 0.33$ and hence $\lambda = 2.00$ matches the labour supply data very well.

5.2 Demographic and cost parameters

Individuals are assumed to be born at the real age of 16, and they can live a maximum of $\bar{J} = 84$ years, after which, at the real age of 99, death is certain (retirement is not modelled, so that agents die at the end of their working life). The sequence of conditional survival probabilities $\{s\}_{j=1}^{99}$ is based on mortality tables for the US and we don’t differentiate mortality rates by sex or race.

The direct cost of education $D_e$ is set to be equal to 0.3 times the average earnings in the economy, which corresponds to an estimate of average (in-state) tuition costs for public and private colleges in the US.\textsuperscript{32}

Tuition subsidies ($T_e$) as a share of average earnings have changed over the last 30 years. A long term average stands at roughly $1/2$ of the tuition costs. We run several experiments based alternative levels of tuition subsidization.

5.3 Some simple tuition experiments

The numerical experiments we report in the rest of this section are compared to a simple benchmark economy in which the discount factor $\beta$ is set to match a physical capital over output ratio of 3.0. The resulting discount factor is very close to one, with the first 3 decimals all being equal to 9. The depreciation rate is set to 0.07, which we compute from NIPA data. No negative assets are allowed in the benchmark economy. The deterministic leisure function $f^e(\theta)$ is discretised and then calibrated to match enrolment rates within different ability bins.

The initial wealth distribution is endogenously determined in this simulations: the accidental bequests are distributed to the new-borns following the steady state asset density. Thus some people are born with zero assets and others with different, positive amounts, which implies that some will be facing a tight liquidity constraint for college education. In

\textsuperscript{32}Source: Education digest, NCES, National Center for Education Statistics.
the simulations we are not correlating initial assets and permanent characteristics, although we plan to do it in the future.

The tuition subsidy experiment is implemented by giving people, ceteris paribus, a transfer (same for all) equal to a percentage of the direct cost of schooling. The extra costs are covered by extra proportional income taxation.

The top panel of the table shows the results for the benchmark economy. The bottom panel shows what happens in partial equilibrium, when prices for human capital do not change. However, taxes must change to fund the tuition subsidy and of course the underlying wealth distribution does change as well as the work behaviour. The middle panel shows the general equilibrium results where human capital process and the interest rate are allowed to change.

In partial equilibrium this universal subsidy leads to a substantial increase in college graduates from 20% to 25%. When breaking this down by ability we see that the increase is high for all ability groups, relative to their original position. In addition, this seems to have come for almost “free” since the tax on labour only needs to increase marginally. This is because the policy attracts a number of high ability and previously liquidity constrained individuals into higher education.\textsuperscript{33} They earn high levels of income which more than compensate the cost of educating them. In fact there is a substantial increase in the college level human capital aggregate from 5.41 to 6.5. This is precisely the logic underlying a number of educational subsidy programmes around the world. Thus in partial equilibrium, the policy pays for itself.

\textsuperscript{33} Admittedly, this is also due to the tight zero borrowing limit that has been imposed for all agents. In this setup people who are born with zero assets are kept out of education.
## Technology 1: Cobb-Douglas

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>Edu. Participation (aggr.shares)</th>
<th>Human Capital Aggregates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benchmark (Tuition $3105 - 30% of median income)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>Less than HS</td>
</tr>
<tr>
<td></td>
<td>0.34</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>Ability 1 (lowest)</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Ability 2</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Ability 3</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Ability 4 (highest)</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>% with wealth=0</td>
<td>0.082</td>
</tr>
</tbody>
</table>

### General Equilibrium 50% (Subsidy $1552)

| Ability 1 (lowest) | 0.92 | 0.075 | 0.005 | 0.53 | 1.0 | 0.94 |
| Ability 2 | 0.36 | 0.55 | 0.092 | 652 | 1710 | 2155 |
| Ability 3 | 0.24 | 0.52 | 0.24 | 652 | 1710 | 2155 |
| Ability 4 (highest) | 0.20 | 0.38 | 0.42 | 652 | 1710 | 2155 |
| % with wealth=0 | 0.082 | r | 0.02519 | 0.28 |

### Partial Equilibrium 50% (Subsidy $1552)

| Ability 1 (lowest) | 0.94 | 0.050 | 0.011 | 0.5575 | 1.0 | 0.95 |
| Ability 2 | 0.36 | 0.47 | 0.17 | 668 | 1767 | 2075 |
| Ability 3 | 0.23 | 0.47 | 0.30 | 668 | 1767 | 2075 |
| Ability 4 (highest) | 0.19 | 0.38 | 0.43 | 668 | 1767 | 2075 |
| % with wealth=0 | 0.086 | r | 0.0252 | 0.272 |
In General Equilibrium though the situation is quite different, at least as far as the aggregate shares are concerned. Following the policy there is a very small decline in aggregate college attendance. This is due to the decline in the marginal product of college level human capital. However, the aggregate figures hide important differences within ability groups. These show a decline in College attendance vis a vis the baseline for ability levels two and three and an increase in College attendance for the highest ability level 4. In addition there is an increase in the rates of high school graduation for the lowest level of ability in response to an increase in the relative price for high school graduates. Finally there is a decline in college for the second ability group. All this adds up to an increase in the supply of human capital for the lowest and highest education groups: The subsidy has in fact led to an increase in inequality.

Similar results have been obtained when we use a production function with a higher elasticity of substitution.
Technology 2: E.of.S.=1.54, rho=0.35

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>Edu. Participation (aggr.shares)</th>
<th>Human Capital Aggregates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benchmark (Tuition $2929 - 30% of median income)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Less than HS</td>
<td>HS</td>
</tr>
<tr>
<td>All</td>
<td>0.34</td>
<td>0.50</td>
</tr>
<tr>
<td>Ability 1 (lowest)</td>
<td>0.89</td>
<td>0.10</td>
</tr>
<tr>
<td>Ability 2</td>
<td>0.38</td>
<td>0.56</td>
</tr>
<tr>
<td>Ability 3</td>
<td>0.24</td>
<td>0.55</td>
</tr>
<tr>
<td>Ability 4 (highest)</td>
<td>0.19</td>
<td>0.45</td>
</tr>
</tbody>
</table>

| % with wealth=0 | 0.082 | r | 0.025 | Tax lab | .27 |

<table>
<thead>
<tr>
<th>Edu. Shares by ability</th>
<th>Marg. Products after Tax. and Depr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability 1 (lowest)</td>
<td>0.88</td>
</tr>
<tr>
<td>Ability 2</td>
<td>0.37</td>
</tr>
<tr>
<td>Ability 3</td>
<td>0.25</td>
</tr>
<tr>
<td>Ability 4 (highest)</td>
<td>0.21</td>
</tr>
</tbody>
</table>

| % with wealth=0 | 0.081 | r | 0.025 | Tax lab | .2734 |

<table>
<thead>
<tr>
<th>Edu. Shares by ability</th>
<th>Marg. Products after Tax. and Depr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability 1 (lowest)</td>
<td>0.89</td>
</tr>
<tr>
<td>Ability 2</td>
<td>0.40</td>
</tr>
<tr>
<td>Ability 3</td>
<td>0.27</td>
</tr>
<tr>
<td>group 4 (highest)</td>
<td>0.23</td>
</tr>
</tbody>
</table>

| % with wealth=0 | 0.084 | r | 0.025 | Tax lab | .2750 |

5.4 Some tax experiments

Here we intend to compare alternative ways of paying for the policy by comparing capital and income taxes. In addition we consider the effect of reducing one type of tax to increase the other, on human capital accumulation

6 Conclusions

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

In the current version we experiment with tuition subsidies. It becomes apparent that while in partial equilibrium such policies can be very effective in increasing education levels and reducing inequality 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 and the fact that labor supply is endogenous lend additional credibility to the result. The distributional changes in this economy under different interventions will be the focus of additional analysis. Moreover, future work includes assessing the relevance of liquidity constraints in the model economy and the equilibrium effects of artificially removing (insuring against) some of the risk components. The importance of risk in the partial equilibrium individual decision about schooling will also be the object of future extensions.
References


BECKER, G. S. (1964): Human Capital. NBER.


Hill, B. (1975): “A Simple General Approach to Approximate the Tail of a Distribution,” 


Economics*, ed. by J. Kareken, and N. Wallace. Federal Reserve Bank of Minneapolis, 
Minneapolis.

*Econometrica*, 72(1), 1–32.


Holland.

over the Life Cycle,” Forthcoming, *Journal of Monetary Economics*. 

29


A PSID Data

The Panel Study of Income Dynamics provides information on a variety of dimensions. Since the beginning it was decided that those eligible for the 1969 and following waves of interviewing would include only persons present in the prior year, including those who moved out of the original family and set up their own households. Until recently, there used to be two different releases of PSID data, Release I (also known as Early Release) and Release II (also known as Final Release). Early release data were available for all years; final release data are available (at time of writing) only between 1968 and 1993. The variables needed for our study are available in both releases. The difference is that Release II data tend to be more polished and contain additional constructed variables. We use Release II data for the period 1968-1993 and Release I data for the period 1994-2001.

Because of successive improvements in Computer Assisted Telephone Interviewing (CATI) software, the quality of the Public Release I files improved dramatically in recent waves, allowing the use of these data with confidence. The differentiation between Public Release I and Public Release II has recently been dropped altogether.

A.1 Sample selection

Unequal probabilities of selection were introduced at the beginning of the PSID (1968) when the original Office of Economic Opportunity (OEO) sample of poor families was combined with a new equal probability national sample of households selected from the Survey Research Center 1960 National Sample. Compensatory weights were developed in 1968 to account for the different sampling rates used to select the OEO and SRC components of the PSID.

The probability sample of individuals defined by the original 1968 sample of PSID families was then followed in subsequent years. 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 nonsample individuals was also made. Only original sample persons and their offspring have been

---

34 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 nonsample individuals must be made. Details about the observations on non-sample persons and their associated weights and relevance are included in the appendix.

35 We also have results obtained from a reduced sample using only Release I data for 1968-1993: estimates of the parameters of interest don’t substantially differ from the full sample estimates.
followed. These individuals are referred to as sample persons and assigned person numbers in a unique range. If other individuals resided with the sample individuals, either in original family units or in newly created family units, data were collected about them as heads, spouses/long term cohabitators or other family unit members, in order to obtain a complete picture of the economic unit represented by the family. However, these nonsample individuals were not followed if they left a PSID family.

Sample persons who are living members of a 1968 PSID family have a sample selection factor equal to the reciprocal of the selection probability for their 1968 PSID family unit. The computation of the sample selection weight factor for sample persons who are “born into” a PSID family after 1968 uses a formula that is conditional on the “sample status” of their parents. However, data for nonsample persons present a problem for longitudinal analysis since the time series for these individuals is left censored at the date at which they entered the PSID family. Furthermore, it is not likely that this left censoring is random with respect to the types of variables that might be considered in longitudinal analysis. Because of the left censoring of their data series, nonsample persons in PSID families have historically been assigned a zero value selection weight factor and a zero-value for the PSID longitudinal analysis weight. This is of course a problem when using the core SRC: non-sample people can be tracked through their Person 1968 number (that assumes values between 170 and 228) and whenever we use individual weights we control for the presence of non-sample individuals.

An additional dimension that is included in the core longitudinal weights are adjustments for panel attrition due to nonresponse and mortality. Attrition adjustments were performed in 1969 and every five years thereafter.

In general individual longitudinal weight values for PSID core sample persons are the product of three distinct sets of factors, that can be summarized as follows:

1. a single factor that represents the reciprocal of the probability by which the sample person was “selected” to the PSID panel;

2. a compound product of attrition adjustment factors developed in 1969 and every 5 years thereafter,

36Beginning with the 1993 wave, PSID is providing users with a file that includes special weights that will enable analysts to include all 1993 sample and nonsample person respondents in cross sectional analysis of the 1993 PSID data set. These weights are called cross-sectional weights (as opposed to the standard longitudinal weights that have been produced from 1969 onwards).
3. Mortality adjustment factors also developed and applied in 1969 and every 5 years thereafter.

A general formula that reflects the composite nature of the individual weights is:

\[ W_{i,1993} = W_{i,sel} \times \prod_{j=1969}^{T} [W_{i,\text{NR}(j)} \times W_{i,M(j)}] \]  

(14)

where: \( W_{i,sel} \) is the selection weight factor – the reciprocal of probability that individual \( i \) is selected to the PSID panel by membership in a 1968 PSID sample family or by birth to a PSID sample parent; \( W_{i,\text{NR}(j)} \) is the attrition adjustment factor applied to the \( i^{th} \) individual weight at time period \( j \); \( W_{i,M(j)} \) is the age, sex and race-specific mortality adjustment applied to the \( i^{th} \) individual weight at time period \( j \)\(^{37} \).

**The 1967-1992 Final Release Sample.** The 1968-1993 PSID individual file contains records on 53,013 individuals (that is, all who were ever present in the sample at least on one year) We drop members from the Latino sample added in 1990 (10,022 individuals) and keep a sample of 42,991 individuals. We then drop those who are never heads of their household and we are left with a sample of 16,028 individuals. We then drop all individuals who are younger than 25 and older than 60, which leaves us with a sample of 13,399 individuals. Dropping observations for self-employed people reduces the sample to 11,574 individuals.

We keep in our sample only people with at least 8 (possibly non continuous) observations, which leaves us with 4,529 individuals. Dropping individuals with missing, zero or top-coded earnings reduces the sample to 4,300 individuals, and dropping individuals with total hours of work that are missing, zero or larger than 5840 further reduces our sample to 4,295 individuals. We eliminate individuals with outlying earning records, defined as changes in log-earnings larger than 4 or less than -2, which leaves 4,211 individuals in the sample.

Finally, dropping people who are connected with the original SEO low-income sample leaves us with a sample of 2,371 individuals.

The composition of the sample by year and by education group is reported in the following tables.

**The 1967-2000 Mixed (Final and Early Release) Sample.** After dropping 10,607 individuals belonging to the Latino sample and 2263 individuals belonging to the new immigrant families added in 1997 and 1999, the joint 1967-2001 sample contains 50,625

\(^{37}\)Of course, non sample people have a zero weight because \( W_{i,sel} = 0 \) for them.
Table 4: Distribution of observations for the 1967-1992 sample, by year

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>783</td>
</tr>
<tr>
<td>1968</td>
<td>853</td>
</tr>
<tr>
<td>1969</td>
<td>906</td>
</tr>
<tr>
<td>1970</td>
<td>965</td>
</tr>
<tr>
<td>1971</td>
<td>1090</td>
</tr>
<tr>
<td>1972</td>
<td>1192</td>
</tr>
<tr>
<td>1973</td>
<td>1280</td>
</tr>
<tr>
<td>1974</td>
<td>1328</td>
</tr>
<tr>
<td>1975</td>
<td>1382</td>
</tr>
<tr>
<td>1976</td>
<td>1428</td>
</tr>
<tr>
<td>1977</td>
<td>1489</td>
</tr>
<tr>
<td>1978</td>
<td>1513</td>
</tr>
<tr>
<td>1979</td>
<td>1550</td>
</tr>
<tr>
<td>1980</td>
<td>1575</td>
</tr>
<tr>
<td>1981</td>
<td>1551</td>
</tr>
<tr>
<td>1982</td>
<td>1551</td>
</tr>
<tr>
<td>1983</td>
<td>1586</td>
</tr>
<tr>
<td>1984</td>
<td>1636</td>
</tr>
<tr>
<td>1985</td>
<td>1656</td>
</tr>
<tr>
<td>1986</td>
<td>1610</td>
</tr>
<tr>
<td>1987</td>
<td>1535</td>
</tr>
<tr>
<td>1988</td>
<td>1484</td>
</tr>
<tr>
<td>1989</td>
<td>1415</td>
</tr>
<tr>
<td>1990</td>
<td>1349</td>
</tr>
<tr>
<td>1991</td>
<td>1285</td>
</tr>
<tr>
<td>1992</td>
<td>1201</td>
</tr>
</tbody>
</table>

Table 5: Distribution of observations for the 1967-1992 sample, by education group

<table>
<thead>
<tr>
<th>Years of Education</th>
<th>Number of Individuals</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 12</td>
<td>330</td>
<td>4,804</td>
</tr>
<tr>
<td>12 to 15</td>
<td>1,354</td>
<td>19,902</td>
</tr>
<tr>
<td>16 or more</td>
<td>687</td>
<td>10,487</td>
</tr>
</tbody>
</table>

Individuals. After selecting only the observations on household heads we are left with 19,583 individuals. Dropping people younger than 25 or older than 60 leaves us with 16,733 people. Dropping the self employment observations leaves 13,740 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 5,559. Dropping individuals with unclear education records leaves 5,544 people in sample. Disposing of individuals with missing, top-coded or zero earnings reduces the sample to 5,112 individuals and dropping those with zero, missing or more than 5840 annual work hours brings the sample size to 5,102 individuals. We eliminate individuals with outlying earning records, defined as changes in log-earnings larger than 4 or less than -2, which leaves 4,891 individuals in the sample. Finally, dropping people connected with the SEO sample reduces the number of individuals to 2,791.

The composition of the sample by year and by education group is reported in the following tables.
Table 6: Distribution of observations for the 1967-2000 sample, by year

<table>
<thead>
<tr>
<th>year</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>776</td>
</tr>
<tr>
<td>1968</td>
<td>842</td>
</tr>
<tr>
<td>1969</td>
<td>891</td>
</tr>
<tr>
<td>1970</td>
<td>952</td>
</tr>
<tr>
<td>1971</td>
<td>1069</td>
</tr>
<tr>
<td>1972</td>
<td>1168</td>
</tr>
<tr>
<td>1973</td>
<td>1250</td>
</tr>
<tr>
<td>1974</td>
<td>1290</td>
</tr>
<tr>
<td>1975</td>
<td>1342</td>
</tr>
<tr>
<td>1976</td>
<td>1385</td>
</tr>
<tr>
<td>1977</td>
<td>1442</td>
</tr>
<tr>
<td>1978</td>
<td>1466</td>
</tr>
<tr>
<td>1979</td>
<td>1502</td>
</tr>
<tr>
<td>1980</td>
<td>1535</td>
</tr>
<tr>
<td>1981</td>
<td>1512</td>
</tr>
<tr>
<td>1982</td>
<td>1505</td>
</tr>
<tr>
<td>1983</td>
<td>1546</td>
</tr>
<tr>
<td>1984</td>
<td>1582</td>
</tr>
<tr>
<td>1985</td>
<td>1609</td>
</tr>
<tr>
<td>1986</td>
<td>1632</td>
</tr>
<tr>
<td>1987</td>
<td>1624</td>
</tr>
<tr>
<td>1988</td>
<td>1631</td>
</tr>
<tr>
<td>1989</td>
<td>1639</td>
</tr>
<tr>
<td>1990</td>
<td>1600</td>
</tr>
<tr>
<td>1991</td>
<td>1628</td>
</tr>
<tr>
<td>1992</td>
<td>1564</td>
</tr>
<tr>
<td>1993</td>
<td>1551</td>
</tr>
<tr>
<td>1994</td>
<td>1486</td>
</tr>
<tr>
<td>1995</td>
<td>1437</td>
</tr>
<tr>
<td>1996</td>
<td>1363</td>
</tr>
<tr>
<td>1998</td>
<td>1293</td>
</tr>
<tr>
<td>2000</td>
<td>1191</td>
</tr>
</tbody>
</table>

Table 7: Distribution of observations for the 1967-2000 sample, by education group

<table>
<thead>
<tr>
<th>years of education</th>
<th>Number of Individuals</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 12</td>
<td>364</td>
<td>5,358</td>
</tr>
<tr>
<td>12 to 15</td>
<td>1,621</td>
<td>25,358</td>
</tr>
<tr>
<td>16 or more</td>
<td>806</td>
<td>13,587</td>
</tr>
</tbody>
</table>