

Education Decisions, Equilibrium Policies and Wages Dispersion

Giovanni Gallipoli
Costas Meghir
Gianluca Violante*[†]

December 2004

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 (HC) 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 and endogenous education choice, allowing for agents' heterogeneity in several dimensions. PSID and CPS data are used to estimate the parameters of a production function with different kinds of human capital and idiosyncratic labor risk processes. A by-product of the estimation procedure is an empirical density of individual permanent efficiency inferred from the distribution of observed wages. These estimates are used to pin down some of the model's parameters. Numerical simulations are used to compare the effects of alternative policy interventions on education participation, endogenous selection, life-cycle earnings and wealth profiles, inequality and productivity.

*University College London and IFS/CEMMAP; University College London and IFS; New York University.

[†]Special thanks go to Lars Nesheim for many discussions on both theoretical and computational aspects of the problem. We are also grateful to Orazio Attanasio, Marco Bassetto, Michele Boldrin, Martin Browning, Monica Costa Dias, Marco Cozzi, Maria Cristina DeNardi, Hide Ichimura, Ken Judd, Hamish Low, Robert Moffitt, Nicola Pavoni, Josep Pijoan-Mas, Jean-Marc Robin, Victor Rios-Rull, Robert Townsend and Fabrizio Zilibotti for comments and suggestions. All errors are ours.

Correspondence: Centre for Microdata Methods and Practice, Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE. *E-mail:* g.gallipoli@ucl.ac.uk .

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 people¹. We examine traditional policies, such as tuition transfers and loan subsidies², 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 tuned 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-

¹Admittedly, 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.

²Standard 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 desirable³.

Studies on the evaluation of policy interventions in equilibrium are more recent. Heckman, Lochner, and Taber (1998b, 1998c) have led the way in advocating a novel approach to policy evaluation which does not overlook equilibrium effects induced by the policy⁴. 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 shock⁵. Each agent in our model represents a single-unit household, conformably with the empirical analysis we produce.

Recent empirical evidence 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's⁶. 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 seminal work on

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

⁴Heckman, Lochner, and Taber estimate and simulate a dynamic general equilibrium model of education accumulation, assets accumulation and labor earnings with skill-biased technological change.

⁵Mortality risk is also explicitly included in the model.

⁶Hyslop (2001) also shows that labor supply explains roughly half of the modest rise in female inequality.

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 prices⁸.

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 risk⁹ 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 investment¹⁰; 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 technology¹¹. Education and employment are mutually exclusive in each period. Foregone earnings and tuition charges are the direct

⁷This selection in our model happens through permanent unobserved characteristics.

⁸Consider, 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.

⁹They multiply education specific earnings by a random variable.

¹⁰As 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.

¹¹We 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.

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 assets¹². We also consider different levels of correlation between ability and initial assets holdings¹³.

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

We derive the optimal consumption and schooling choices for an individual of given ability who supplies labor in a competitive market. A unique good is produced in the economy, and it can be either consumed or used as physical capital. Different kinds of human capital are an input to the aggregate production function and command different returns. Wage differences among people are the consequence of differences in education (between group inequality) and differences in labor efficiency (within group inequality). People with different labor efficiencies are perfect substitutes within schooling groups. Agents can accumulate assets representing ownership shares on physical capital.

2.1 Demographics and preferences

Each agent represents a (non-altruistic) single-unit household, whose life starts at age 1 and lasts \bar{j} periods, after which death is certain¹⁴. Therefore the population consists of \bar{j} overlapping generations, each with an ex-ante identical distribution of heterogeneity. 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. Negative borrowing limits open up the

¹²Gale 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.

¹³This 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).

¹⁴Retirement is not modelled. We experiment also with different labor life lengths.

possibility that people dying prematurely can be in debt¹⁵.

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 path maximizing expected lifetime utility.¹⁶

The education level is denoted as e , with $e = e_1$ the lowest and $e = e_3$ the highest. Permanent individual characteristics are denoted as θ and distributed over the domain $[\theta_{\min}, \theta_{\max}]$. We let $\{z\}_{j=1}^{\bar{j}}$ be a sequence of uninsurable idiosyncratic shocks. We also assume that the distributions of ability θ and of idiosyncratic shocks z are independent of time.

The spot market wage for education e_l is w_{e_l} , individual labor supply is $n \in [0, 1]$. Leisure l is the complement to one of n , and is chosen optimally, however the amount of leisure consumed by students is a deterministic function of ability, defined as $f^S = f^S(\theta)$.

Agents pay proportional taxes τ_n and τ_k on, respectively, labor and asset income: the distinction is kept to capture the effect of physical capital taxation on HC accumulation, due to a substitution between the two forms of investment¹⁷. Aggregate physical capital is denoted as K . The risk free, pre-tax return on assets is r and the depreciation rate of K is $\delta > 0$.

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 law of motion for the idiosyncratic efficiency shocks is summarized by a transition function π denoted as $\pi_{z_{j+1}|z_j} = \pi\{z_{j+1} | z_j\}$.

Given some initial conditions \bar{x}_1 for the state variables, the age 1 household's utility over sequences of consumption and leisure, $c = \{c_1, \dots, c_{\bar{j}}\}$ and $l = \{l_1, \dots, l_{\bar{j}}\}$, is denoted as $U(\bar{x}_1, c, l)$ and can be written as the expected discounted sum of period utilities

$$U(\bar{x}_1, c, l) = E_{z \in Z} \sum_{j=1}^{\bar{j}} S_j \beta^{j-1} u(c_j, l_j) \quad (1)$$

where $S_j = \left(\prod_{i=1}^j s_i\right)$ and β is the intertemporal discount factor. The period budget

¹⁵Yaari (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.

¹⁶A 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.

¹⁷This effect was first noted by Heckman (1976) and is very intuitive if we think of investment as a way to transfer resources intertemporally. Changing the price of intertemporal substitution affects the quantity and quality of investments.

constraint is

$$c_j + a_{j+1} = [1 + r(1 - \tau_k)]a_j + w_e \exp^{\epsilon_j} n_j (1 - \tau_{ne}) (1 - d_j) - (D_e - T_e) d_j^{18} \quad (2)$$

where d_j is a binary variable which is 1 if the agent is in education and 0 otherwise, and a_j denotes individual asset holdings¹⁹ at age j . D_e is the direct cost of schooling, T_e summarizes government subsidies towards education e .

The term \exp^{ϵ_j} is individual labor efficiency, and ϵ_j is defined as

$$\epsilon_j(\theta, e, z) = \theta + \xi_j(e) + z_j \quad (3)$$

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

2.2 Household's problem

The education choice depends on relative returns to different education levels, current pecuniary cost of schooling, permanent characteristics θ and current labor shock and asset holdings, z and a ²⁰. Agents can go back to school, and their incentive to do so varies with time preference and individual states.

To pass from e_l to e_{l+1} an agent must continuously be a student for a (exogenously given) number of periods in school²¹. No schooling is possible after e_3 has been achieved.

Given prices and direct costs of schooling, the binary function $d_j = d_j(\theta, e, z, a)$ describes schooling choice as a mapping from the space of individual states to set $\{0, 1\}$, where 1 stands for full-time education status and 0 stands for full-time worker²².

Conditional on entering the labor market, the labor supply policy of an agent is $n_j = n_j(\theta, e, z, a \mid d_j = 0)$ ²³.

2.3 Optimality and value functions

Without conditioning on the current education decision, the optimal policy of an agent can be represented as a vector $p_j = (d_j, a_{j+1})$, where a_{j+1} is the optimal saving policy and d_j the binary education decision.

¹⁹Individual asset holdings satisfy: $a_j \geq a_{\min}$ for every j and $a_{\bar{j}+1} \geq 0$. The first inequality is a borrowing constraint, whereas the second is a transversality condition for agents reaching age \bar{j} .

²⁰To keep the model as simple as possible we do not explicitly model the sector producing education.

²¹A state variable in this problem is therefore the number of previous years of education already under the belt: we do not use additional notation for it, although this variable implicitly determines future budget constraints and utility.

²²This sorting mechanism hinges on the assumption that progress from one educational level to the next may require more than just one study spell. In the simulations we set the lengths of the required study spells to match the features of the educational system under investigation.

²³Using the intratemporal margin condition it is possible to express the individual labor supply as a function of optimal consumption and real wage. The analytical details of the labour/leisure intertemporal choice are provided in the Appendix.

A functional equation is an equivalent and unique approach to the household's sequence problem. We use value functions to characterize the optimal path²⁴.

The functional equation can be written as

$$J(x_j, p_{j-1}) = \sup_{p_j} v(p_j) + s_j \beta^{j-1} \int_Z \pi_{z_{j+1}|z_j} J(x_{j+1}, p_j) dz_{j+1} \quad (4)$$

for given initial condition \bar{x}_1 .²⁵

We call $J^*(x_j, p_{j-1})$ the unconditional value function because it is defined over all possible education choices.

In order to characterize the unconditional value function it helps to consider two conditional value functions which are obtained by assigning a value to (current) binary choice d_j . The conditional versions of $J^*(x_j, p_{j-1})$ are then the value of employment ($d_j = 0$) and the value of education ($d_j = 1$).

We denote the *conditional* value function as $J^*(x_j, p_{j-1} | \text{condition})$, with the condition being the value of d_j .²⁶ The unconditional functional equation $J^*(x_j, p_{j-1})$ is the upper envelope of the conditional values of employment and education. Without loss of generality, we can reduce the complexity of the value functions associated to different kinds of employment by making the choice of employment sector irreversible²⁷.

Of course if agents had the possibility to return to education after a working spell they could choose a different employment sector, so we also assume that agents don't go back to education.²⁸

The conditional value of employment is denoted as $J^*(x_j, p_{j-1} | d_j = 0) = W_j(\theta, e, z, a)$ and is unique.

If we rule out returning to education after a working spell, this value is defined as

$$W_j(\theta, e, z, a) = \max_{a', n} u(c, 1 - n) + s_j \beta \int_Z \pi_{z'|z} W_{j+1}(\theta, e, z', a') dz' \quad (5)$$

In the class of employment value functions, special attention must be devoted to the

²⁴In this section we use an hyphen "' " to identify next period unknown values and omit the age/time subscripts when possible.

²⁵Our problem satisfies regularity conditions, described in Gallipoli (2004), under which a value function $J^*(x_j, p_{j-1})$ satisfying the functional equation (4) exists.

²⁶Such notation allows to summarize education status for both current and past periods (d_{j-1} and d_j), which suffices for our analysis.

²⁷We assume that the costs of reverting to different, feasible 'careers' are sufficiently high.

²⁸This assumption also reduces the degree of non concavity of the value functions describing the solutions, making the numerical work easier.

When agents are allowed to go back to education, the discounted value of the future in (5) must be extended to include the value of additional education.

value function of newly employed agents. This conditional value is

$$J^*(x_j, p_{j-1} = (d_j = 1, a_j) \mid d_j = 0) = \max_e \{W_j(\theta, e, z, a)\}_{e=e_1}^* \quad (6)$$

where e^* is the agent's education level: the conditional value of first-time employed equals the highest employment value among those available. It is possible to prove that the conditional value of employment is monotonous, concave and smooth, and the optimal policy is single valued and continuous.

The conditional value of being in education, $V_j = V_j(\theta, e, z, a)$, also exists, is unique and is defined as

$$J^*(x_j, p_{j-1} \mid d_j = 1) = V_j(\theta, e, z, a) = \max_{a'} u(c, f_e(\theta)) + \quad (7)$$

$$+ s_j \beta \int_Z \pi_{z'|z} \max \left\{ V_{j+1}(\theta, e', z', a'), \{W_{j+1}(\theta, e, z', a')\}_{e=e_1}^* \right\} dz'$$

where e^* is the education level of the agent. Both V_j and W_j are subject to (2).²⁹ The education value is not generally concave on the state space.³⁰

The *unconditional* choice problem of an agent can be summarized as

$$\max_{\{a_{j+1}, d_j\}} \{V_j, W_j\}$$

We call this the unconditional problem because we are not restricting the value of the binary choice d_j . It is possible for the conditional values V_j and W_j to cross at some state space locations, which may give to the unconditional value a peculiar “butterfly” shape. To guarantee that education decisions are uniquely determined, we assume that whenever the present value of education is at least as large as the present value of employment, education is chosen over employment. Using this assumption and the set of results obtained for the conditional optimal policies, we argue that the unconditional optimal policy is uniquely determined and piecewise continuous.³¹

2.4 Aggregate variables

We study equilibrium allocations and assume a stationary population. The aggregate states of the economy are aggregate physical capital K and efficiency-weighted aggregate labor

²⁹The conditional value of education $e = e_2$ for people in the last year of schooling is such that $V_{j+1}(\theta, e^3, z', a') < W_{j+1}(\theta, e, z', a')$ for any e , which satisfies the assumption that no further schooling is possible after reaching e^3 .

³⁰This is potentially troubling because it raises issues of non-uniqueness of the optimal policies. This and related issues for this class of models are discussed in Gallipoli (2004). A discussion for the infinitely lived case is in Townsend and Ueda (2003).

³¹The discontinuities in the asset policies occur at the switch points because of the jumps in marginal utility at such locations. Nonetheless, the optimal policy duplet $p_j = (a_{j+1}, d_j)$ is continuous between successive switch points.

supplies (referred to as human capital aggregates) 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

$$H_e = \sum_j \zeta_j \int_X h_j(x) d\psi_j(x) \quad (8)$$

where $\psi_j(x) = \psi_j(\theta, e, z, a)$.

Individual efficiency-weighted labor supplies are denoted as h_j and defined as

$$h_j(\theta, e, z, a) = \exp^{\varepsilon_j(\theta, e, z)} n_j(\theta, e, z, a) \quad (9)$$

We define a measure space $(X, \mathcal{F}(X), \psi_j)$, where X is the individual state space and $\mathcal{F}(X)$ is a σ -algebra on X . For each set $F \subseteq \mathcal{F}(X)$, let $\psi_j(F)$ represent the normalised measure of age j agents whose individual states lie in F as a proportion of all age j agents. Calling ζ_j the fraction of age j agents in the economy we define

$$\mu = \mu(F, j) = \zeta_j \psi_j(F)$$

as a measure of agents belonging to age group j with individual state vector $(\theta, e, z, a) \in F$. The aggregate states determine the relative prices in the economy. The demographics are stable, so that age j agents make up a constant fraction ζ_j of the population at any point in time. The ζ_j values are normalized to sum up to 1 and are such that $\zeta_{j+1} = s_j \zeta_j$.

2.5 Markets structure

The unique physical good is used as numeraire. Such good can be either consumed or saved. In this economy savings a represent ownership rights over physical capital K . We do not model entrepreneurial choices directly, but maintain that entrepreneurs behave optimally in managing firms.

We consider model specifications with and without missing annuity markets. In the simulations in which we allow non-degenerate asset holdings for new borns, we also experiment with different degrees of correlation between initial wealth and permanent characteristics θ .³²

We do not model involuntary unemployment, but people can choose to consume all their leisure endowment if, for example, the market value of their time is too low.

³²Imposing 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.

2.6 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 :

$$F(H, K) = \bar{A}H^{1-\alpha}K^\alpha \quad (10)$$

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

$$H = \{A_1H_1^\rho + A_2H_2^\rho + A_3H_3^\rho\}^{\frac{h}{\rho}} \quad (11)$$

with $h = 1$ given the CRS assumption.³³

In this specification (A_1, A_2, A_3) are share parameters, while ρ pins down the Allen elasticity of substitution among labor inputs.³⁴ When ρ is equal to zero the technology is Cobb-Douglas, whereas values of ρ greater than zero indicate more substitutability than in the Cobb-Douglas case. An alternative and interesting specification we consider is³⁵

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

which has a symmetry property imposing that the elasticity of substitution between H_2 and H_3 is the same as the that between H_3 and H_1 . Therefore, if $\rho_2 > \rho_1$ we have that H_3 is more complementary with H_1 than with H_2 . Also, the grouping allows separate parts of the above technology to be Cobb-Douglas, when either ρ_2 or ρ_1 tend to zero.

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 $r + \delta = \frac{\partial F}{\partial K}$.

2.7 Government

Government has revenues from proportional taxation of labor and asset income at respectively τ_{ne} and τ_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

³³For strict quasi-concavity of the production function ρ has to lie within $(-\infty, 1)$.

³⁴In the CES case, the Allen elasticity of substitution between any two inputs is $\frac{1}{1-\rho}$. This is also known as the Allen/Uzawa E. of S. and is the most widely used. However, Blackorby and Russell (1981) show that there is no intuition about what it measures. Blackorby and Russell advocate the use of the so-called Morishima E. of S., and another alternative for multisector models would be the so-called direct E. of S. proposed by McFadden. In what follows we just use the Allen E. of S. as a simple approximation.

³⁵Hamermesh (1993) attributes this ‘grouping’ production function to Sato (1967).

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 taxation³⁶. The government has a balanced budget in each period.

3 Equilibrium

We use a notion of equilibrium in which the measure $\mu(x, j)$ remains unchanged over time. This notion of equilibrium is known as stationary recursive competitive equilibrium Lucas (1980).³⁷

3.1 Equilibrium definition

Let $(X, F(X), \psi_j)$ be an age-specific measure space with state space X and $F(X)$ be a σ -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 $V_j(x)$, $W_j(x)$, price functions $w_e(\mu)$, $r(\mu)$, densities $(\psi_1, \dots, \psi_{\bar{j}})$ and $(\zeta_1, \dots, \zeta_{\bar{j}})$, 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. $V_j(\theta, e, z, a)$, $W_j(\theta, e, z, a)$ are the associated value functions
3. Firms employ inputs so that

$$w_e = F_{H_e} \quad \text{for } e \in \mathfrak{S}$$

$$r + \delta = F_K$$

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 $d_j(x)$, $a_{j+1}(x)$, $c_j(x)$.³⁸
5. The good, asset and labour markets clear³⁹

³⁶The government budget constraint is

$$G + \sum_j \zeta_j \int_X T_e d_j(x) d\psi_j(x) =$$

$$= \sum_j \zeta_j \int_X [1 - d_j(x)] \tau_n e w_e h_j(x) d\psi_j(x) + \sum_j \zeta_j \int_{\bar{A}} r \tau_k a_j d\psi_j(a)$$

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

³⁷Our model stifies the conditions for defining a measure ψ_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.

³⁸Given ζ_j , also $\mu(x, j) = \zeta_j \psi_j(x)$ is a stationary measure.

³⁹Equilibrium definitions in the asset and good markets must include cross border asset holding FX if the interest rate r is constant.

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

4 Identification and Estimation

We estimate values for a set of production and efficiency parameters using US data. This section describes the procedures used to identify and estimate:

- education specific age-earning profiles and idiosyncratic labor shock processes;
- the empirical density of idiosyncratic permanent characteristics θ over the working population;⁴⁰
- the aggregate technology parameters determining shares and, whenever possible, substitution elasticities for aggregate inputs.

Different data sets are used in the process, namely CPS, PSID and NIPA.

4.1 Estimating wage equations: skill prices and age profiles

Skill prices and age-earning profiles can be estimated by imposing some structure on the data, which we do by using our model. For each education group we study a wage equation that is consistent with the individual earning mechanism of the model. The (log-linear) specification of individual hourly wages is

$$\ln w_{eit} = w_{et} + g(\text{age}_{eit}) + u_{eit} \quad (12)$$

where

$$u_{eit} = \theta_i + z_{eit} + m_{it} \quad (13)$$

In this notation w_{eit} denotes observed hourly wage rate for individual i at time t in education group e , w_{et} is a (time dependent) hourly return to the specific human capital e , θ_i is an individual fixed effect, $g(\text{age}_{eit})$ is a education-specific function of age⁴¹ and z_{eit} is a education-specific idiosyncratic shock, possibly autocorrelated. Finally, the term m_{it} denotes a measurement error component in wage.⁴²

⁴⁰It must be stressed that we refer to ability as a set of observable and unobservable characteristics that have a direct impact on households' earnings but are not explicitly modelled.

⁴¹We do not include returns to experience. Experience is the difference between age and years of schooling, and agents belonging to a given education group have roughly the same number of years of schooling. Therefore the age effects end up capturing returns from experience as well as seniority.

⁴²Different identification results can be obtained depending on whether the error term is correlated over time. However, a necessary identification condition is that measurement error is orthogonal to all observed and unobserved characteristics.

Our steady state model does not include any time variation. However, recovering time varying prices w_t from data is of pivotal importance for the identification of age effects, residual terms and, in a different context, to pin down the evolution of HC aggregates over time. Since we are primarily interested in non-demographic determinants of education decisions, we do not model cohort effects explicitly. Self selection implies that fixed effects are correlated with both education decisions and observed wage rates. By estimating a distinct wage equation for each education group we control for the education self-selection problem, but heterogeneity still represents an obstacle to identification.

Assuming linearity of the permanent error components, we identify our model parameters by adopting a within group specification for wage equations. We estimate the following specification

$$(\ln w_{eit} - \ln \bar{w}_{ei}) = (\ln w_{et} - \ln \bar{w}) + g(\text{age}_{eit} - \overline{\text{age}}_{ei}) + (u_{eit} - \bar{u}_{ei}) \quad (14)$$

where the upper-bar denotes an (individual) time average. This delivers consistent estimates of time effects and age profiles⁴³.

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. All interviews are retrospective, providing data on the previous year. The sample for this study combines a cross-section sample of nearly 3,000 families, representative of the US population, selected from the Survey Research Center's master sampling frame, and a subsample of about 1,900 families interviewed previously by the Bureau of the Census for the Office of Economic Opportunity. The subsample drawn from the OEO-Census study was limited to low-income families, and compensatory weights were developed in 1968 to account for the different sampling rates used to select the OEO sample component as opposed to the SRC component⁴⁴. A subsample of Latino (Latin American origins) families was added in 1990 and dropped in 1995. Additional immigrant families were added in 1997 and 1999. Moreover, in 1997 some families belonging to the OEO-Census sample

⁴³A normalization of the estimation results is necessary to obtain age-earning profiles, skill prices and estimates of permanent heterogeneity. These are then used in numerical simulations. The normalization is bound to be arbitrary because we don't have any 'metric' to measure and compare the relative contributions of age, skills and permanent heterogeneity in determining the final wage rate. A description is included in the appendix.

⁴⁴In fact the original 1968 wave data must be weighted unless one uses only the SRC representative cross section sample.

component were dropped.

We do not use individuals associated with the Census low income sample, the Latino sample or the New Immigrant sample⁴⁵ 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 household⁴⁶ and is described as total labor income of the head⁴⁷. We use this measure, deflated by the CPI-U for all urban consumers, as the reference earning variable. 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.⁴⁸

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.

Figure (1) plots the evolution of mean and variance of log annual earnings over time in our PSID sample, both by education group and for the whole sample.

The plot shows that average annual earnings have experienced a drop in real values during the early 1980s, but a steep rise in the 1990s drove them to be almost 20% higher in 2000 than in 1967. Big differences in the time evolution of earnings are evident for different education group: college graduates earnings rose by almost 30% over the sample period, whereas high school drop-outs earnings barely managed to stay the same thanks to a recovery in the late 1990s. The 1980s' drop in real earnings was accompanied by higher dispersion in all groups. However, visual inspection suggests that while for college graduates the earnings variance settled at the higher level, for other groups it did revert to lower values (although still higher than in the 1970s). As college graduates represent an

⁴⁵Lillard and Willis (1978) make the case that the SEO low income sample should be dropped because of endogenous selection problems.

⁴⁶In 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.

⁴⁷This 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.

⁴⁸Using heads to approximate households' behaviour finds some support in recent work by Hyslop (2001), who provides evidence of very strong and positive assortive matching by couples and shows that such matching is based on permanent individual characteristics.

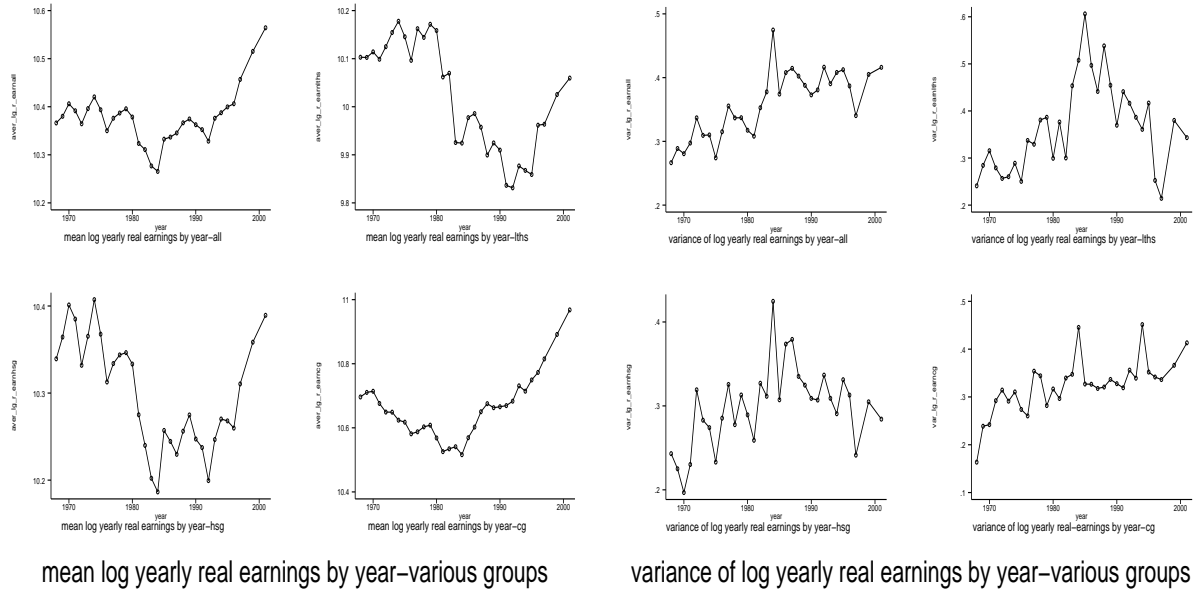


Figure 1: Mean and variance of log real annual earnings (in 1992 dollars), both by education group and combined

ever increasing share of workers over time, this might suggest that less skilled people have self-selected into the highest education group inducing a rise in within group inequality. In fact, the relative size of the different education groups has changed substantially

In order to give a rough idea of the importance of labor supply decision in shaping earnings differences, we report in figure (2) the mean and variance of hourly wage rates, both for the combined sample and by education groups. Individual hourly wages are computed as the ratio of yearly earnings and yearly hours worked, which introduces additional noise in the measure (especially when earnings and hours are measured with error). Nonetheless, the variances are lower than in the case of annual earnings, which we interpret as evidence that labor supply decisions are a significant mechanism linking wage changes and earnings dispersion. Interestingly, the evolution of average hourly wage rates is remarkably similar to annual earnings. This preliminary evidence points to the possibility that differences in permanent characteristics (e.g., taste for leisure) might induce self-selection into education groups, which in turn drives differences in dispersion.

We use hourly wage rates as dependent variable in the wage equations.⁴⁹ We estimate a wage specification as in(14) and use a 4th degree polynomial in age to approximate the

⁴⁹The case for using hourly rates rather than annual earnings is presented by Heathcote, Storesletten, and Violante (2004). They also point out the importance of labor supply in shaping dispersion.

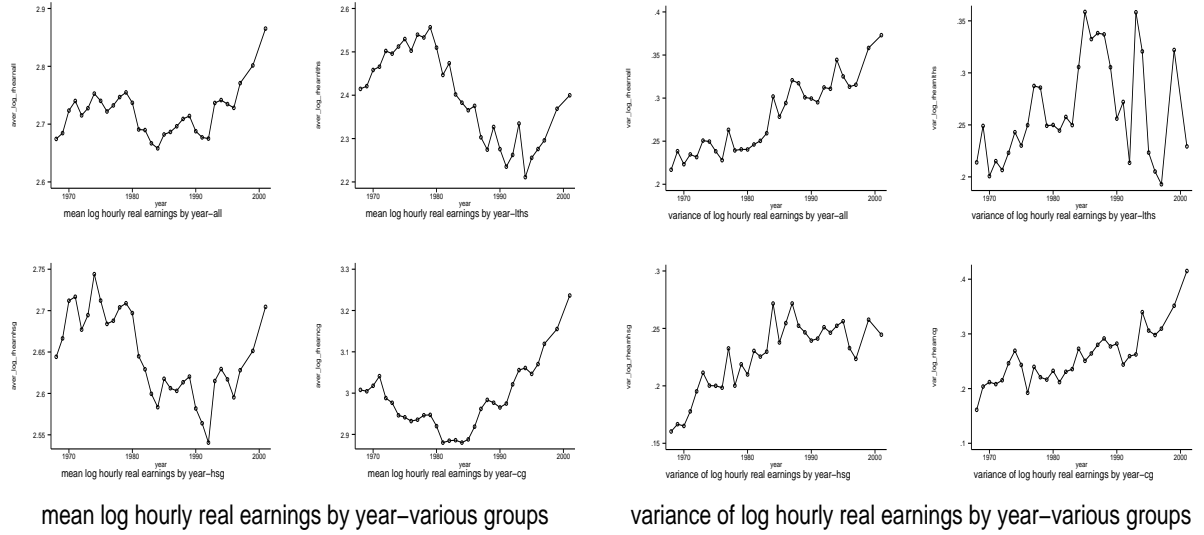


Figure 2: Mean and variance of log real hourly wage rates (in 1992 dollars), both by education group and combined

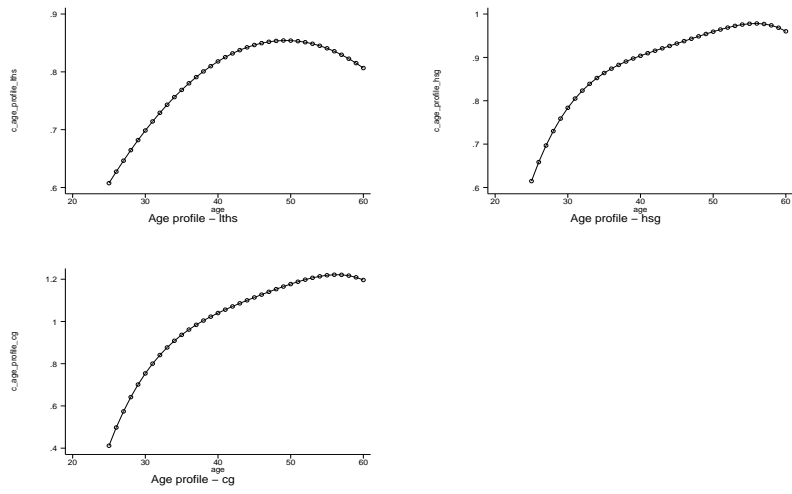
possibly non linear (age_{it}) functions (for the LTHS group just a 2nd degree polynomial is sufficient to characterize the age profile). The age polynomials' coefficients are presented in table (1).

Table 1: Age polynomials' coefficients

Dependent variable: log hourly earnings		
coeff.	point estimate	S.E.
<i>Education=LTHS</i>		
age	0.0412505	0.0081143
age ²	-0.0004179	0.0000905
<i>Education=HSG</i>		
age	0.4928285	0.1071015
age ²	-0.0162768	0.0039883
age ³	0.0002413	0.0000644
age ⁴	-1.34e-06	3.82e-07
<i>Education=CG</i>		
age	0.8697329	0.1560285
age ²	-0.0282	0.0058548
age ³	0.0004149	0.0000953
age ⁴	-2.30e-06	5.69e-07

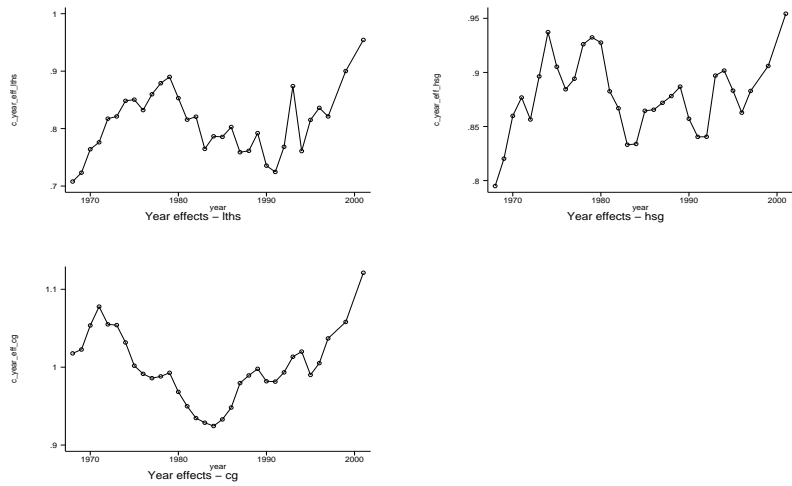
Figure (??) plots the age profiles implied by the polynomial estimates for different education groups.

By fitting the within group specification of the wage equation we also obtain $\ln \hat{w}_{et}$, estimates of the log-price of labor by education and year, which are plotted in figure (4). The time effects have a natural interpretation as time varying prices of skills associated to



AGE profiles – various groups

Figure 3: Age profiles of labor efficiency by education group - age on the horizontal axis



Year effects – various groups

Figure 4: Estimated log of marginal labor productivity, by education and year

different education groups and will be used to identify the supply of human capital in the economy.⁵⁰ The log of marginal returns indicate that, after controlling for fixed effects and age components, the marginal return to each single labor type has gone up, from 10% in the HS drop-outs case case to 30% for college graduates. Moreover, we can now rationalize the 1980s plunge observed in the raw earnings and wages series as a time effect: this roughly corresponds to the time in which the effects of the baby boom generation came to their full fruition.

4.3 Education Specific Labor Inputs

Consider the problem of studying the production function

$$Y_t = F(H_{1t}, H_{2t}, H_{3t}, K_t)$$

From national accounts data (NIPA) we can obtain long time series of aggregate output. We also have observations on aggregate physical capital, K_t , and on the wage bills that are paid in each year to different education groups (denoted as WB_t^e)⁵¹.

In order to retrieve technology parameters it is necessary to identify and estimate HC aggregates, H_{et} ⁵², which are defined as an efficiency-weighted sum of individual labor supplies. The crucial question is whether we can recover (H_1, H_2, H_3) for a reasonably long number of periods.

Our approach is to identify (H_1, H_2, H_3) by combining information on observable wage bills (WB^1, WB^2, WB^3) and estimated year specific skill prices, w_{et} . In fact, by definition we have that

$$WB_t^e = w_{et}H_{et} \tag{15}$$

We can identify HC aggregates by using the estimated time series of skill prices \widehat{w}_{et} obtained from the wage equations.

The main problem with this identification procedure is a data measurement problem: in the US fringe benefits and other employer's contributions are often not recorded as straightforward earnings⁵³ and can account for a sizeable proportion of yearly earnings.

⁵⁰An arbitrary normalization has been made to set the scale of the various wage components. Details available from the authors.

⁵¹The (yearly) wage bill for a given education group is the total labor payments received by people within that education group in a given year.

⁵²The subscripts e and t stand for human capital type (as implied by the educational level) and date of observation.

⁵³An example, pointed out by Ken Judd, is that of employer's pension contributions which can account for over 10% of yearly earnings.

Furthermore, they are likely to represent different proportions of total earnings in different education groups, and tend to be higher for college graduates. This might lead to an underestimate of the aggregate human capital for the higher education groups.

4.4 Combining CPS data and PSID estimates

The wage bills should reliably represent the distribution of US working population over education groups in each year. For this reason we use CPS data: the Current Population Survey (CPS) is a monthly survey of about 50,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics.⁵⁴ This monthly survey of households is conducted for BLS by the Bureau of the Census through a scientifically selected sample designed to represent the civilian noninstitutional population. Respondents are interviewed to obtain information about the employment status of each member of the household 15 years of age and older. Each month about 50,000 occupied units are eligible for interview. Some 3,200 of these households are contacted but interviews are not obtained because the occupants are not at home after repeated calls or are unavailable for other reasons. This represents a non-interview rate for the survey that ranges between 6 and 7 percent. In addition to the 50,000 occupied units, there are 9,000 sample units in an average month which are visited but found to be vacant or otherwise not eligible for enumeration. Part of the sample is changed each month. The rotation plan, as explained later, provides for three-fourths of the sample to be common from one month to the next, and one-half to be common with the same month a year earlier. The CPS has been used to collect annual income data since 1948, when only two supplementary questions were asked in April: "How much did ... earn in wages and salaries in 1947 ..." and "how much income from all sources did ... receive in 1947". Over the years, the number of income questions has expanded, questions on work experience and other characteristics have been added, and the month of interview relating to previous year income and earnings has moved to March. This yearly survey goes under the name of March CPS Supplement.⁵⁵ Age classification is based on the age of the person at his/her last birthday. The adult universe (i.e., population of marriageable age) is comprised of persons 15 years old and over for March supplement data

⁵⁴The 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).

⁵⁵Today, information is gathered on more than 50 different sources of income, including noncash income sources such as food stamps, school lunch program, employer-provided pension plan and personal health insurance. Comprehensive work experience information is given on the employment status, occupation, and industry of persons 15 years old and over.

and for CPS labor force data. Each household and person has a weight that should be used in producing population-level statistics. The weight reflects the probability sampling process and estimation procedures designed to account for nonresponse and undercoverage. Unweighted counts can be very misleading and should not be used in demographic or labor force analysis.

We use the CPI for all urban consumer (with base year 1992) to deflate the CPS earning 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.⁵⁶

Figure (5) reports the number of people working in each year by education group, as reported by the CPS. Employed workers are a rough measure of employed human capital.

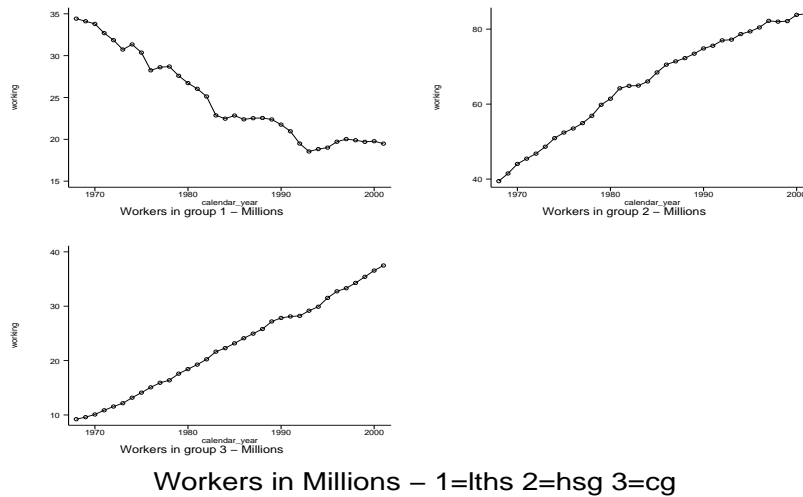


Figure 5: Employed workers in millions, by education and year

It is clear that some strong and persistent trends towards higher levels of education have characterized the sample period. This might seem all the more surprising since the return to higher education have increased substantially over the sample period.

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

⁵⁶This 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.

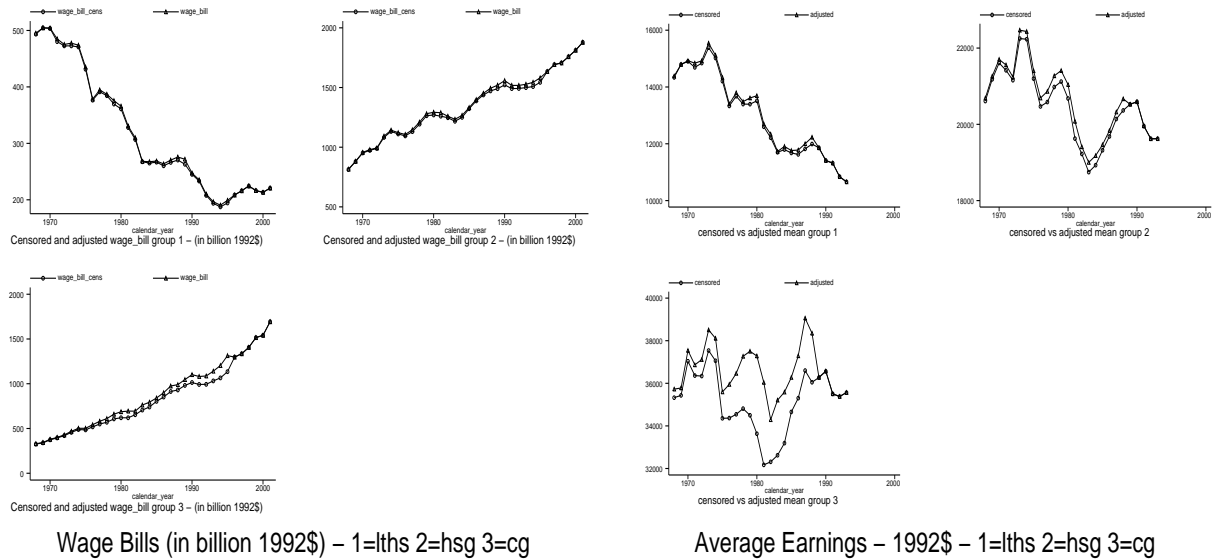


Figure 6: Total and average earned labor income, by education and year. Total in billions of 1992 dollars, average in units of 1992 dollars.

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 times much less than that).

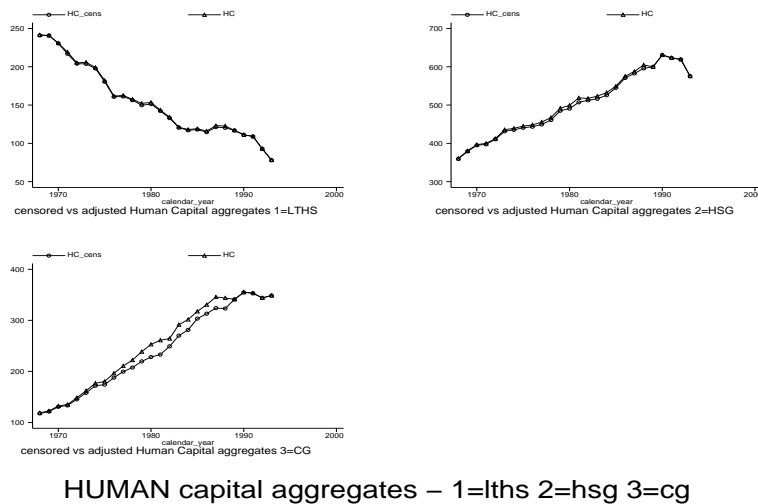


Figure 7: Value of efficiency weighted labor supply (HC) in billions of 1992 dollars, by education and year

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).

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. We take this as an indication of substantial changes in underlying technology parameters over that period.

4.5 Permanent characteristics and their distribution

We ‘residually’ identify permanent characteristics (fixed effects) θ_i , whose estimate is denoted as $\widehat{\theta}_i$, from the sequence of agent-specific residuals associated to the wage equation (12)⁵⁷. We resort to the fact that

$$\ln w_{eit} = \ln (w_t \exp^{\theta_i + g(\text{age}_{eit}) + u_{eit}}) = \ln w_{et} + g(\text{age}_{eit}) + \theta_i + u_{eit}$$

and compute

$$\widehat{\theta}_i = \frac{\sum_{t=1}^{T(i)} \ln w_{it} - \widehat{\ln w_t} - \widehat{g(\text{age}_{it})}}{T(i)}$$

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.

Under this specification the individual fixed effects θ_i capture omitted sources of permanent heterogeneity which have some effect on individual earnings: they range from observable characteristics such as gender and race to non-observable characteristics such as cognitive ability and family background. In this sense, the resulting distribution of estimated fixed effects can be thought of as a single-index summary of multi-dimensional heterogeneity.

These forms of heterogeneity constitute an essential part of the individual ability to earn that is not due to age or price effects. We include idiosyncratic permanent characteristics in the numerical simulations using a discretized distribution based on the shape of the estimated empirical density.⁵⁸

⁵⁷The so-called incidental parameters problem affects the FE estimates: limited panel length is responsible for FE estimation bias. The fact that our estimates rely on people who are observed at least for 8 years partially reduces the bias.

⁵⁸It must be noticed that, under our invariance assumption, such estimates provide an approximation to the unconditional

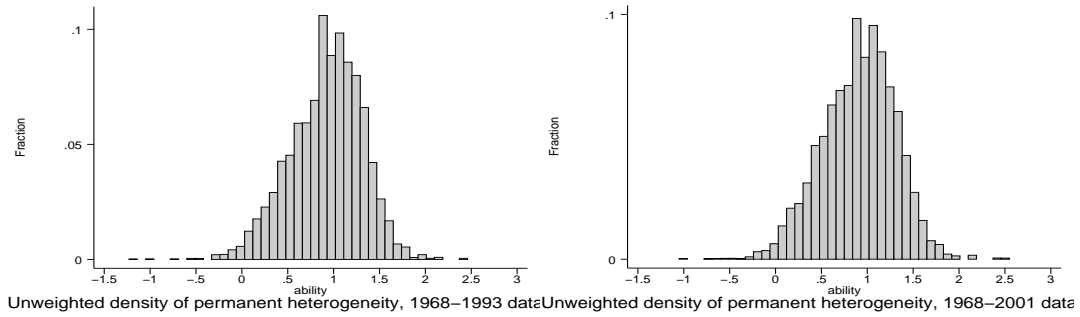


Figure 8: Estimated density of log fixed effects for small (67-93) and large (67-00) samples

In figure (8) we report the empirical frequencies of $\hat{\theta}$ obtained by aggregating both cross-sectionally and longitudinally. Changing the length of the sample and using weights does not introduce any substantial variation on the shape of the density. Therefore, we don't report the empirical density obtained by using longitudinal weights.

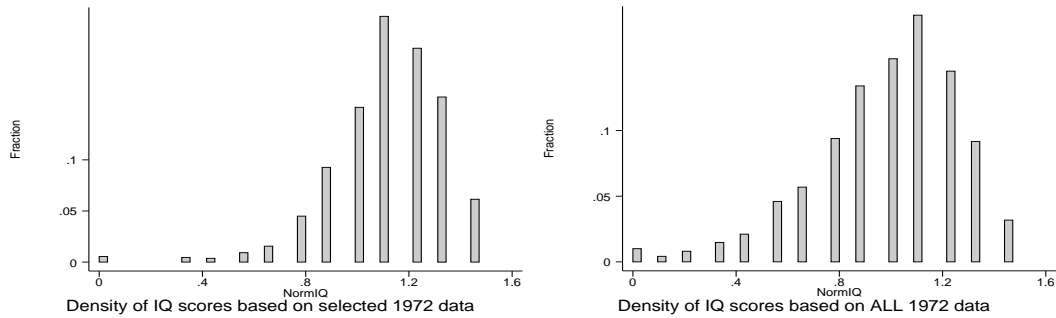


Figure 9: Density of IQ measurement from 1972 PSID wave, for the whole sample and a comparable sub-sample

Since the 1972 wave of the PSID contains a one-off IQ test for people who took part, we also report the normalized empirical density of the test scores for the whole 1972 sample and for a subsample selected by using our sampling criteria. This offers a relatively easy way to compare our inferred approximation to fixed effects with an actual measurement based on IQ. Figure (9) reports the measured IQ densities for the whole 1972 sample and the selected sub-sample. It seems that both the IQ density and the estimated fixed effect density exhibit a long left tail, indicating a larger downward dispersion. They are both strikingly non-normal.

distribution of the ability to earn, but cannot be used to infer information about the conditional distributions of ability in each education group because the longitudinal PSID sample that we select fails to represent the marginal densities of the US population over age and education. This problem becomes more severe the further away we go from the original sampling date. We have estimated this distribution using both the large sample covering the period between 1967 and 2000 and a smaller sample for covering the period 1968-1993. Furthermore, we have checked whether weighting the estimated $\hat{\theta}$ using PSID longitudinal weights would change the estimated empirical frequencies of ability.

4.6 Analysis of labor efficiency shocks

The wage equation residual, rescaled by removing the permanent characteristics, varies by construction around zero along time and across individuals and is defined as

$$\hat{u}_{eit} = \ln w_{eit} - \hat{g}(age_{nit}) - \ln \hat{w}_{et} - \hat{\theta}_i$$

We assume that the error term \hat{u}_{eit} can be decomposed into two components

$$\hat{u}_{eit} = z_{eit} + m_{eit}$$

where z_{eit} is an autocorrelated error process and m_{eit} is classical measurement error $iid(0, \sigma_{em}^2)$.

If we assume that $\{z_{eit}\}_t$ is a autocorrelated process⁵⁹

$$z_{eit} = \rho_e z_{eit-1} + \varepsilon_{eit}$$

with $\varepsilon_{eit} \sim iid(0, \sigma_\varepsilon^2)$, we can achieve identification of the autoregressive parameters in one of several ways. A first possibility is to use the following second moments

$$VAR(\hat{u}_{eit}) = VAR(z_{eit}) + VAR(m_{eit})$$

$$COV(\hat{u}_{eit}, \hat{u}_{eit-1}) = COV(z_{eit}, z_{eit-1})$$

where

$$VAR(z_{eit}) = \frac{\sigma_{\varepsilon_e}^2}{1 - \rho_e^2}$$

$$COV(z_{eit}, z_{eit-1}) = \frac{\rho_e \sigma_{\varepsilon_e}^2}{1 - \rho_e^2}$$

and compute⁶⁰

$$\rho = \frac{COV(z_{eit}, z_{eit-1})}{VAR(z_{eit})} = \frac{COV(\hat{u}_{eit}, \hat{u}_{eit-1})}{VAR(\hat{u}_{eit}) - VAR(m_{eit})} \quad (16)$$

Of course an external estimate of $VAR(m_{eit})$ is necessary in this case.

⁵⁹We could assume either a unique autoregressive coefficient ρ for all education groups and a set of group specific ρ_e . We choose the second option, which can provide a measure of insurability of shocks by education group.

⁶⁰In fact if the above specification is correct then we have that

$$VAR(\hat{u}_{eit}) = \frac{\sigma_{\varepsilon_e}^2}{1 - \rho_e^2} + \sigma_m^2$$

$$COV_j(\hat{u}_{eit}) = \frac{\rho_e^j \sigma_{\varepsilon_e}^2}{1 - \rho_e^2} \quad j \geq 1$$

An alternative way to identify the autoregressive coefficient without resorting to an external estimate of $VAR(m_{it})$ is also available, given classical measurement error.

In fact, in this case we can write⁶¹

$$\rho_e = \frac{COV(z_{eit}, z_{eit-2})}{COV(z_{eit}, z_{eit-1})}$$

Furthermore, we can compute

$$\tilde{u}_{eit} - \hat{\rho}_e \tilde{u}_{eit-1} = \hat{\varepsilon}_{eit} = (z_{eit} - \hat{\rho}_e z_{eit-1}) + (m_{it} - \hat{\rho}_e m_{it-1}) = \varepsilon_{eit} + (m_{it} - \hat{\rho}_e m_{it-1}) \quad (17)$$

The moments of the constructed residual $\hat{\varepsilon}_{eit}$ are

$$\begin{aligned} VAR(\hat{\varepsilon}_{eit}) &= \sigma_{\varepsilon,e}^2 + (1 + \hat{\rho}_e^2) \sigma_m^2 \\ COV_1(\hat{\varepsilon}_{eit}) &= -\hat{\rho}_e^2 \sigma_m^2 \\ COV_j(\hat{\varepsilon}_{eit}) &= 0 \quad j \geq 2 \end{aligned}$$

and can be used to test the goodness of the specification we assume for the z process.

4.7 Estimation and Testing of Labor Shock Processes

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 parenthesis. Higher persistence is associated with higher values of $\hat{\rho}_e$: higher persistence implies as a less insurable kind of shock and corresponds to a more volatile lifecycle pattern for earnings.

Table 2: Estimates of the autoregressive coefficient $\hat{\rho}$, by education group and pooled. Bootstrapped S.E. in parenthesis

Group 1	Group 2	Group 3	Pooled
0.651 (0.130)	0.557 (0.042)	0.608 (0.058)	0.584 (0.034)

The estimated values for $\hat{\rho}_e$ seem indicate that group 2 (High school graduates) experience the lowest earnings' risk.⁶² These findings are apparently in contrast with some of

⁶¹Estimates are based on non weighted residuals, as weighting would not add any information, since heterogeneity is factored out of the errors by construction.

⁶²The point estimates of σ_ε^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.

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. Moreover, alternative specifications of measurement error (such as non-classical autocorrelated noise) can push up these values. In the numerical simulations we experiment with various levels of persistence.

4.8 Aggregate Technology parameters

As we mentioned before, the estimation of an aggregate production function can only rely on constructed data for the aggregate labor inputs H_{et} .

In identifying and estimating technology parameters, we start from the relatively easier case of Cobb-Douglas technology

Let aggregate output Y be produced through the following technology

$$Y = \left(H_3^A H_2^{(1-A)B} H_1^{(1-A)(1-B)} \right)^{1-\alpha} K^\alpha$$

Using NIPA data we find the share of capital α 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. Moreover, we can apply the delta method to approximate the standard errors of such functions. If we apply this procedure separately for each year we can pinpoint the evolution of these functions over the sample period.

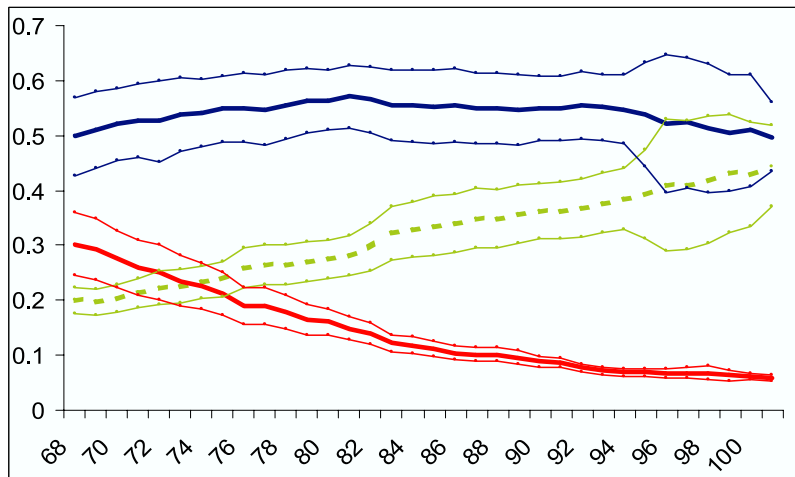


Figure 10: Labor shares in human capital input of technology, computed using Cobb-Douglas specification (with bounds equal to ± 2 standard errors). Period: 1968-2000. Larger bounds after 1996 are due to changes in top-coding of income in the CPS.

Figure (10) 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 (10) 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 confirms 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 performin 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 capital⁶³ and then use a 2-step GMM method which controls for endogeneity and serial correlation of TFP to estimate the parameters. In what follows we present some results obtained by applying the above method to the log-linearized version of the production function in which we set the elasticity parameters of a general CES specification to zero.

In fact, we find that a 2-step GMM procedure applied to the unrestricted CES specification provides scarcely robust and highly insignificant estimates for all technology parameters. On the other hand, a restricted Cobb-Douglas specification of the form $H = H_3^A (H_2^B H_1^{1-B})^{1-A} \exp^f$ can be easily log-linearized as

$$\ln F(H_t) = A \ln H_{3t} + (1 - A) [B \ln H_{2t} + (1 - B) \ln H_{1t}] + f_t$$

and given the small sample dimension (30 observations) this linearization makes the GMM procedure more robust and reliable. Moreover, in a C-D specification it does not matter whether H_2 is nested with H_1 or H_3 . Such distinction would matter only in a more general CES specification.⁶⁴

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

⁶³Details of the model can be provided by the authors.

⁶⁴We also run further check with a standard Powell-Newton algorithm using calculus conditions to identify a minimum.

To minimize the objective function we use a simplex method algorithm first proposed by Nelder and Meade (1965). This method has the advantage to check whether a candidate set of estimates is a real minimizing set by using a quadratic expansion in the neighborhood of such set and verifying that the minimum of such quadratic form corresponds to the minimum found by the Simplex.

alternative moment weighting matrix choices (the identity matrix and the optimal matrix

Table 3: Point estimates of (long-term, 1967-1997) labor shares in technology (S.E. in parenthesis)

	First Step Weighting: Identity Matrix	First Step Weighting: Optimal Matrix
A	0.260 (0.200)	0.284 (0.207)
B	0.783 (0.115)	0.790 (0.123)

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 say that we find an average yearly 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. Also the standard errors are very close to the ones we have obtained by using the delta method. Remarkably, the labor shares roughly sum up to one, even though we don't impose this restriction in the estimation procedure.

5 Simulations

The parameter estimates from the previous sections are used to set model parameters (technology, labor shocks, permanent characteristics, age-earnings profiles) for our equilibrium analysis. In the simulations each time unit represents a year and the parameters are based on yearly estimates.

5.1 Preferences parameters

We use utility CRRA preferences

$$u(c_j, l_j \mid d_j = 0) = \frac{[c_j^\nu l_j^{1-\nu}]^{(1-\lambda)}}{1-\lambda}$$

$$u(c_j \mid d_j = 1) = \frac{[c_j^\nu f^e(\theta)^{1-\nu}]^{(1-\lambda)}}{1-\lambda}$$

The parameters ν and λ of the period utility jointly pin down the intertemporal elasticity of substitution of consumption, that is $\frac{1}{1-\nu(1-\lambda)}$. With $\nu = 0.33$ and $\lambda = 2.00$ we have that such elasticity is roughly 0.75.

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 $\bar{j} = 50$ years, after which, at the real age of 65, death is certain (retirement is not modelled, so that agents die at the end of their working life). The sequence of conditional survival probabilities $\{s\}_{j=1}^{50}$ is based on mortality tables for the US.

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.⁶⁵

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 β 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 tax rates (on capital income and labor earnings) are both set to 15%. The depreciation rate is set to 0.07, which we compute from NIPA data. No assets borrowing is allowed in the benchmark economy.

Notice that we report experiments in which the deterministic leisure function $f^e(\theta)$ is set to match enrolment rates: this presents however a problem, because there are potentially many functions which can satisfy this condition for the “marginal student”. We ignore this problem in the following simulations, although we are currently working on two alternatives to circumvent it. These experiments should therefore be considered as preliminary evidence only.

The initial wealth distribution is endogenously determined in this simulations: the accidental bequests are distributed to the new borns following the asset density of those individuals who have died. Nobody is born with negative assets by assumption, but some people are born with zero assets and others with different, positive amounts. In the simula-

⁶⁵Source: Education digest, NCES, National Center for Education Statistics.

Table 4: Tuition experiments both for fixed prices (P.E.) and market clearing prices (G.E.) - Comparison to Benchmark with no subsidy.

	Tuit. \$92	Subs. \$92	% workers by edu			Month. Salary \$92, pretax			0 assets % of pop.	r %
			LTHS	HS	C	LTHS	HS	C		
Benchmark	5826	0	25.2	58.5	16.3	1111	2023	2899	12.2	3.92
50% subs. (PE)	4831	2415	5.3	16.2	78.5	921	1677	2403	5.7	3.92
50% subs. (GE)	5821	2910	25.2	58.4	16.4	1110	2021	2890	12.1	3.90
150% subs. (PE)	4868	7302	5.3	16.4	78.2	928	1690	2422	6.0	3.92
150% subs. (GE)	5840	8761	25.1	58.3	16.5	1121	2027	2889	11.9	3.87

tions we are not correlating initial assets and permanent characteristics, although we plan to do it in the future.

The most obvious 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 following table reports results for such experiments. No additional taxes are levied on people in order to finance the additional subsidy costs: resources are obtained from existing tax revenues. This simplifying assumption is admittedly more realistic for the P.E. case than for the G.E. case.

Results are reported in table 5.3. The tuition cost in dollars changes in the experiments because tuition costs are computed as a fixed percentage of average earnings, which change in different experiments. This allows however to maintain consistency in relative prices. We report the percentage of workers in the different education groups, the pre-tax monthly salary by education group, the proportion of people in the economy with zero assets and the pre-tax interest rate.

It is immediately evident that big differences exist between the fixed price P.E. case and the G.E. case: in P.E. tuition subsidies seem to have big effects on overall output and earnings inequality (which go down), on college enrolment which jumps up by three times and on the proportion of people with zero assets, which also goes down.

All these effects are cancelled in G.E., when prices are let free to adapt the the new supply of labor skills. In this case we find almost no change with respect to the benchmark

Table 5: Tax experiments both for fixed prices (P.E.) and market clearing prices (G.E.) - Comparison to Benchmark with no subsidy.

	Tuit. \$92	Subs. \$92	% workers by edu			Month. Salary \$92, pretax			0 assets % of pop.	r %
			LTHS	HS	C	LTHS	HS	C		
Benchmark	5826	0	25.2	58.5	16.3	1111	2023	2899	12.2	3.92
cut K tax 2/3 (PE)	10411	0	36.8	29.9	33.2	1985	3615	5180	0.08	3.92
cut K tax 2/3 (GE)	5334	0	25.7	54.0	18.6	1015	1852	3079	15.9	4.14

economy.

If we think of P.E. as a policy intervention targeting a small part of the population, it is evident that micro-interventions yield massive effect, which dissipate as the scale of the intervention increases.

In this economy there is no asset borrowing, so that poor

5.4 Some tax experiments

We also run some simple tax experiments, which again we compare to the benchmark economy. In this case we cut the tax rates on capital income by 2/3.

Cutting the physical capital tax gives an insight on the effects of changes in taxation of interest income on education decisions. Heckman (1976) argues that increased taxation of interest income encourages investment in human capital because of a simple substitution of a cheaper method of transferring consumption across periods.

Results for the tax experiments are reported in table 5.4.

The result of the capital tax experiment seems to indicate that in P.E. there is less HC accumulation when we reduce capital income tax, as we might expect from the simple substitution argument described above. However, there seems to be a positive effect on the type of HC accumulated, with relatively higher levels of college graduates which is quite interesting. The incentives are profoundly changed by this kind of policy. In GE we have relatively small changes with respect to the benchmark, but the switch from HS to College persists and is significantly larger than any tuition policy intervention in G.E..

6 Conclusions

In this paper we investigate the relationship between individual education decisions and life cycle earnings, considering the case of alternative policy interventions changing the parameters of the model.

We use the model to impose structure on U.S. data and directly estimate some relevant parameters of the economy we want to model. This framework allows us to numerically simulate the model economy in order to assess the effects of policy interventions under both partial and general equilibrium assumptions.

We find that the general equilibrium implications of interventions almost completely cancel out the effects that are visible under partial equilibrium assumptions. This is in line with results proposed 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

- ABRAHAM, A. (2001): “Wage Inequality and Education Policy with Skill-biased Technological Change in OG Setting,” Ph.D. thesis, Universitat Pompeu Fabra.
- ACEMOGLU, D. (2000): “Technical change, inequality, and the Labor Market,” Mimeo.
- 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.
- BLACKORBY, C., AND R. R. RUSSELL (1981): “The Morishima Elasticity of Substitution; Symmetry, Constancy, Separability, and its Relationship to the Hicks and Allen Elasticities,” *Review of Economic Studies*, 48(1), 147–158.
- BLINDER, A. S., AND Y. WEISS (1976): “Human Capital and Labor Supply: A Synthesis,” *Journal of Political Economy*, 84(3), 449–472.
- CUNHA, F., J. HECKMAN, AND S. NAVARRO (2004): “Counterfactual Analysis of Inequality and Social Mobility,” 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.
- FRENCH, E. (2000): “The Labor Supply Response To (Mismeasured but) Predictable Wage Changes,” Mimeo, Federal Reserve Bank of Chicago.
- GALE, W., AND J. SCHOLZ (1994): “Intergenerational Transfers and the Accumulation of Wealth,” *Journal of Economic Perspectives*, 8(4), 145–160.
- GALLIPOLI, G. (2004): “The problem of non convexities in life-cycle models,” Mimeo.
- HAMERMESH, D. S. (1993): *Labor Demand*. Princeton, N. J.: Princeton University Press.
- HEATHCOTE, J., K. STORESLETTEN, AND G. VIOLANTE (2004): “The Macroeconomic Implications of Rising Wage Inequality in the US,” Mimeo.

- 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. (1976): “A Life-Cycle Model of Earnings, Learning, and Consumption,” *Journal of Political Economy*, 84(4), S11–S44, Part 2: Essays in Labor Economics in Honor of H. Gregg Lewis.
- 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.
- KEANE, M., AND K. WOLPIN (1997): “The Career Decisions of Young Men,” *Journal of Political Economy*, 105(3), 473–522.
- LEE, D. (2001): “An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice,” Ph.D. thesis, University of Pennsylvania.
- LEVHARI, D., AND Y. WEISS (1974): “The effect of risk on the investment in human capital,” *American Economic Review*, 64, 950–63.
- LILLARD, L. A., AND R. J. WILLIS (1978): “Dynamic aspects of income mobility,” *Econometrica*, 46, 387–422.
- LUCAS, R. E. (1980): “Equilibria in a Pure Currency Economy,” in *Models of Monetary Economics*, ed. by J. Kareken, and N. Wallace. Federal Reserve Bank of Minneapolis, Minneapolis.
- MEGHIR, C., AND L. PISTAFERRI (2004): “Income Variance Dynamics and Heterogeneity,” *Econometrica*, 72(1), 1–32.

- 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.
- NELDER, J. A., AND R. MEADE (1965): “The Role of Pre-market Factors in Black-White Wage Differences,” *Computer Journal*, 7(5), 308–313.
- POLIVKA, A. E. (2000): “Using Earnings Data from the Monthly Current Population Survey,” mimeo, Bureau of Labor Statistics.
- 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.
- STORESLETTEN, K., C. TELMER, AND A. YARON (2002): “Consumption and Risk Sharing over the Life Cycle,” Forthcoming, *Journal of Monetary Economics*.
- TOWNSEND, R. M., AND K. UEDA (2003): “Financial Deepening, Inequality, and Growth: A Model-Based Quantitative Evaluation,” Discussion Paper 238, International Monetary Fund, IMF Working Papers 03/193.
- 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.
- YAARI, M. E. (1965): “Uncertain Lifetime, Life Insurance and the Theory of the Consumer,” *Review of Economic Studies*, 32(2), 137–150.

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 households⁶⁶. 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 1994-2001⁶⁷.

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

⁶⁶A 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.

⁶⁷We 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 weight⁶⁸. 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,

⁶⁸Beginning 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:

$$W_{i,1993} = W_{i,sel} \times \prod_{j=1969}^T [W_{i,NR(j)} \times W_{i,M(j)}] \quad (18)$$

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; $W_{i,NR(j)}$ 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 ⁶⁹.

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

⁶⁹Of course, non sample people have a zero weight because $W_{i,sel} = 0$ for them.

Table 6: Distribution of observations for the 1967-1992 sample, by year

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

Table 7: Distribution of observations for the 1967-1992 sample, by education group

years of education	Number of Individuals	Number of Observations
less than 12	330	4,804
12 to 15	1,354	19,902
16 or more	687	10,487

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 5559. 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 8: Distribution of observations for the 1967-2000 sample, by year

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

Table 9: Distribution of observations for the 1967-2000 sample, by education group

years of education	Number of Individuals	Number of Observations
less than 12	364	5,358
12 to 15	1,621	25,358
16 or more	806	13,587