

# Post schooling human capital investments and the life cycle variance of earnings

Thierry Magnac\*, Nicolas Pistoiesi†, Sébastien Roux‡

First version: April 2011

This version: October 2014

## Abstract

We propose an original model of human capital investments after leaving school in which individuals differ in their initial human capital obtained at school, their rates of return and costs of human capital investments and their terminal values of human capital at an arbitrary date in the future. We derive a tractable reduced form Mincerian model of log earnings profiles along the life cycle which is written as a linear factor model in which levels, growth and curvature of earnings profiles are individual-specific. Using panel data from a single cohort of French male wage earners observed over a long span of 30 years starting at their entry in the labor market, we estimate this model by random and fixed effect methods, test restrictions, decompose variance of earnings into permanent and transitory components and evaluate how earnings inequality over the life-cycle is affected by changes in structural parameters. In particular, increases in life expectancy bring about sizeable increases in earnings inequality.

**JEL Codes: C33, D91, I24, J24, J31**

**Keywords: human capital investment, earnings dynamics, post-schooling earnings growth, dynamic panel data**

---

\*Toulouse School of Economics (Université Toulouse Capitole, GREMAQ, IDEI)

†Toulouse School of Economics (Université Toulouse Capitole & GREMAQ)

‡Banque de France, Ined & Crest, Paris

# 1 Introduction<sup>1</sup>

Since the seminal work of Friedman and Kuznets (1954), a large literature studying the change in earnings inequality has emerged. Landmark articles such as Lillard and Willis (1978), Hause (1980) or MaCurdy (1982) introduced parsimonious statistical representations of earnings processes to assess the relative importance of permanent and transitory components in the variance of earnings. This decomposition builds on the observation that permanent changes in earnings have a larger impact on individual welfare than transitory ones (Blundell, 2014). However, not much attention is devoted to its economic foundation and how economically interpretable permanent differences between individuals may contribute to explain earnings dispersion and its evolution over the life-cycle.

Another more structural strand of the literature focuses on the estimation of post-school human capital investment models derived from Ben Porath (1967). Heckman, Lochner and Taber (1998) were among the first to estimate such a human capital investment model at school and later in life in a dynamic and stochastic general equilibrium set-up. Guvenen and Kuruscu (2012) analyze as well the effect of skill biased technical change on inequality in an equilibrium set-up with heterogeneous agents investing in human capital. This is also the object of Huggett, Ventura and Yaron (2011) who use such a microeconomic model calibrated on the US PSID data to decompose inequality into their long-run individual determinants and short-run shocks. Polachek, Das and Thamma-Apiroam (2013) estimate individual-specific parameters that govern human capital investment behavior and describe these heterogeneity terms as functions of cognitive ability, personality traits and family background. Except in the latter paper, unobserved heterogeneity in the human capital production technology is restricted. Furthermore, predicted earnings equations are non linear. Both issues raise econometric questions about identification of parameters and estimation biases due to incidental parameters.

Our paper aims at bridging the gap between these two strands and we highlight three of its contributions. First, we propose an original human capital investment model from which we derive that the earnings process is given by a linear factor model whose factor loadings are functions of individual specific parameters. Second we estimate these parameters using a panel

---

<sup>1</sup>This is a substantially revised version of an earlier paper of ours presented at the 1st French Econometrics Workshop in 2009. We thank Christian Belzil, Richard Blundell, Laurent Gobillon, Jim Heckman, Bernard Salanié and three anonymous referees for helpful comments and discussions and seminar participants at Yale, Tilburg, Toulouse, CREST, Pompeu Fabra, Leuven University, Panel Data'12 in Paris, ESEM'12 in Malaga, Jean-Pierre Florens' Festschrift in Toulouse, Duke, Hong Kong University, Mc Gill, SOLE-EALE ' 13, Ecole Polytechnique, University College of London, Mannheim and the Cowles conference'14 for their helpful comments. All errors remain ours.

of nearly 7,500 male French wage earners who enter the labor market in 1977 and have been followed for 30 years. These data provide an interesting contrast with the US and the UK experience since labor earnings inequality has been stable in France over the last forty years. Third we develop an original empirical strategy that combines the use of random effect and fixed effect methods. This structural and empirical set-up permits counterfactual evaluation in new dimensions. It also provides a more traditional decomposition of earnings inequality into permanent and transitory components. We now return to each of these contributions in more detail.

In our human capital investment model, inspired by Ben Porath (1967), the life-cycle profiles of individual earnings are summarized by a limited number of individual-specific permanent components in a linear factor model. An appealing feature of our model is that the factor loadings are functions of the individual specific structural parameters. In a sense, we are extending to post-school investments described by Mincer (1974) what has been developed times ago by Heckman (see Heckman, Lochner and Todd, 2006, for a survey) and Card (2001) for schooling investments in human capital. By providing a simple setting, this analysis stands apart from Polachek, Das and Thamma-Apiroam (2013) in which the estimated earnings equation is non linear (Haley, 1976) and results from an approximation by the truncation of a series.

In our model, agents differ in four dimensions. Firstly, they have different initial human capital levels when they enter the labor market. Secondly, they differ in their returns to skill investments, that is, some are more productive in transforming invested time in productive skills. We also assume that the marginal utility cost of invested time is heterogenous within the population. Finally, we allow the terminal value of human capital to vary across individuals and infer from the curvature of the earnings profile, the implicit horizon of investment that agents consider. This follows a suggestion by Lillard and Reville (1999) insisting on this crucial aspect of earnings growth.

As a result, this model predicts a linear factor model for the earnings equation in which factor loadings are functions of the individual specific structural parameters. Some structural restrictions are testable and some structural parameters can be point-identified while others are only partially identified. Furthermore, this set-up is fully compatible with the view that human capital stocks are perfectly substitutable within education or skill groups while they are imperfect substitutes between groups as discussed in Browning, Hansen and Heckman (1999).

By construction, the model delivers the well known predictions of a human capital setting (Ben Porath, 1967). Earnings profile are increasing and concave and this reflects the shortening

of the investment horizon. Second, the variance of earnings is U-shaped along the life-cycle because high-return investors have a steeper earnings profile than low-return individuals experiencing a flatter profile and these profiles cross after a few years (Mincer, 1974). Third, because investments in human capital are more intensive at the beginning of the life cycle for the high return investors, the cross-section correlation, at the beginning of the life cycle, between earnings growth and level, is negative although this correlation increases along the life-cycle and becomes positive (Rubinstein and Weiss, 2006).

Adopting a highly stylized human capital model comes at the price of simplifying other elements that might drive earnings dynamics in a structural way. We first take as given past investments in schooling although this is an important heterogeneity dimension in our model. We treat search and job mobility as frictions under the form of exogenous shocks (see e.g. Postel-Vinay and Turon, 2010) that contribute to the transitory part of the income process. We neglect taxes because we cannot reconstruct their value from our data and we find a simple way of modeling the interactions between investments and uncertainty which partially neutralize the importance of risk (see e.g. Huggett, Ventura and Yaron, 2011). We do not model general equilibrium effects as in Heckman et al. (1998) or uncertainty or learning of individual effects as in Cunha, Heckman and Navarro (2007) or Guvenen (2007). Furthermore, the linear factor representation of the earnings equation is obtained by assuming away consumption smoothing although we provide a more elaborate structural form in which consumption is smoothed. From an empirical perspective, some reduced-form specifications estimating also other characteristics of the earnings distribution such as Browning, Erjnaes and Alvarez (2012) are richer in terms of heterogeneity while our model focusses on the profiles of means and variances of earnings over time that condition the main diagnostics about life-cycle earnings inequality.

Nonetheless, the linear factor representation remains attractive as a simple model of human capital investments and as a bridge between structural models and earnings dynamics models. First, the conditions for identification of the distribution of heterogeneity are clearer than in more complicated models in which identification is difficult to prove and in which parametric assumptions are often made in order to make the estimation tractable. Second, the estimation of this linear model is simpler than alternatives such as the one developed by Polachek et al. (2013) in which non-linearities might be difficult to deal with and in which the issue of incidental parameters might be a problem. Third, this model provides a structural interpretation to specifications used in the literature on the dynamics of earnings such as the heterogeneous income processes (HIP) examined by Baker (1997) and Guvenen (2009). We provide foundations

for these specifications and make explicit their underlying assumptions.

As a second contribution, we estimate the model on a very long panel for a single cohort of male French wage earners working in the private sector and observed from 1977 to 2007. DADS social security data is an administrative dataset collecting earnings in the private sector. The first key advantage for our purposes is that it includes enough observations so that we can study a large single cohort of males (nearly 7,500). As in the structural model of human capital investment, individuals simultaneously enter the labor market and face the same economic environment over their life-cycle in contrast with most studies of earnings dynamics that must pool different cohorts to collect samples large enough (Meghir and Pistaferri, 2010). Using a single entry cohort makes the assumption of common priors among agents about future price or depreciation processes more credible. Furthermore, France is characterized by a stable labor earnings distribution along those years in strong contrast to the US and the UK (Atkinson, 2008). Thirdly, as the data come from social security records, we expect fewer measurement errors than in usual surveys or other administrative data although this is not entirely convincing in our application. Finally, the DADS data are long and homogeneous enough to study the dynamics of earnings over a long period of time. In particular we find much longer dependence for transitory earnings than what is usually found in the literature.

Our model is a linear factor model whose use was pioneered by Jim Heckman through a series of papers with diverse coauthors (Aakvik, Heckman and Vytlačil, 2005, Carneiro, Hansen and Heckman, 2003). Factors consist of a level term, a linear trend and exponential term which captures the curvature of earnings profiles. We propose as a third contribution an original empirical strategy to estimate such models. We first decompose the estimation of earnings equations into aggregate and individual specific components. The identification of aggregate components is obtained by imposing restrictions on human capital prices and depreciation rates and assessing the robustness of our results to changes in those restrictions. Residuals of log earnings are then analyzed through a sequence of random and fixed effect methods to control serial dependence and to show that results are robust to incidental parameter issues. We begin with estimating a random effect model by pseudo maximum likelihood (Alvarez and Arellano, 2004) that lets us estimate serial dependence. Controlling for serial dependence, we then derive fixed effect estimates of the individual factors in a second step. To evaluate the importance of incidental parameter issues, we compare random and fixed effect estimates of the covariance matrix of individual factor loadings and we show that the bias becomes second-order when the number of observed periods is roughly above 20.

We then evaluate structural restrictions and compute estimates of the structural unobserved heterogeneity terms. This allows us to construct counterfactuals by measuring the impact of changes in the economic environment. The alternative strategy of estimating distributions of individual-specific effects as in Cunha, Heckman and Schennach (2010) turns out to be difficult to implement here because of the presence of structural constraints on individual specific effects while direct fixed effect estimation is performed at a reasonable cost.

Our main results can be summarized as follows. Levels and growth of earnings are positively correlated in the long run while they are negatively correlated initially. This corroborates one of the predictions of the human capital setting as seen above. Moreover, the larger the level and the slope of earnings profiles, the more concave they are and this stems from the horizon effect. Structural restrictions are satisfied in most of the sample although a small fraction of earnings profiles do not agree with our set-up.

The human capital component of earnings account for 65% of the variance, on average during the observation period. This share grows over time from 3% at entry in the labor market to 90%, 30 years after entry. Moreover, a counterfactual experiment shows that an increase in the horizon of investment or life expectancy by two years increases means and variances of earnings, above all at the end of the observation period. The longer the working period, the more high-return investors reap benefits from investing. Cross-section inequality increases by around 20% at the end of the period although this estimate has quite a large standard error.

In the next Section, we present the model of human capital accumulation and derive the predicted life-cycle profile of earnings. In Section 3, we state the economic and identifying restrictions that yield an identified linear factor model of life-cycle earnings. Data are described in Section 4 and this is followed in Section 5 by a presentation of our empirical strategy. We also detail the econometric estimation methods that we use and results are presented in Section 6. After the presentation of inequality decompositions and counterfactuals in Section 7, a final Section concludes.

## 2 The Model

We present an original model of human capital investment in discrete time which shares common features with Ben Porath (1967) but not all. Specifically, we characterize the optimal sequence of post-schooling human capital investments over the life cycle of agents who maximize their utility over their lifetime. Agents start with an individual specific level of human capital obtained

at school and have individual specific costs, individual specific rates of return for investments and individual specific terminal values of human capital stocks. We are considering a single cohort of agents who enter the labor market at the same time and face the same economic environment. Time and potential experience are confounded. Our structural assumptions on the decision problem lead to a closed form solution for the life-cycle profile of earnings whose dynamics depend on individual-specific abilities to earn and to learn (Browning, Hansen and Heckman, 1999).

## 2.1 The set up

Individuals enter the labor market at a period which is normalized to  $t = 1$ . Schooling and the entry decision in the labour market are considered as given. We follow Heckman, Lochner and Taber (1998) by assuming that the post-school human capital production process differs from the one affecting school investments although both are interdependent. Schooling as the main element of previous human capital accumulation and as a determinant of labor market entry is likely to be correlated with individual specific characteristics affecting post-school investments in human capital.

From period 1 onwards, agents can acquire human capital by devoting time or effort to training. Human capital is assumed to be of one type only, skills are general and costs are borne by the workers. Labor supply is inelastic and potential individuals earnings,  $y_i^P(t)$ , is the product of the individual-specific stock of human capital,  $H_i(t)$  by its individual specific price,  $\exp(\delta_i(t))$  that yields  $y_i^P(t) = \exp(\delta_i(t))H_i(t)$ . Individuals face uncertainty through the variability of human capital (log) prices  $\delta_i(t)$  which are mainly affected by aggregate shocks but also by individual ones when there are frictions (e.g search, information asymmetry or learning as in Rubinstein and Weiss, 2006). Firms might temporarily value individual specific human capital in a way that differs from the market in order to attract, retain or discourage specific individuals, or because information is imperfect. The human capital (log) price,  $\delta_i(t)$ , is assumed to follow a stochastic process and is fully revealed at period  $t$  to the agent. We do not provide a market analysis of the wage equilibrium process and take it as given (in terms of its distribution). We defer the presentation of the stochastic properties of this process until the end of this section and of its statistical properties until Section 3.2.

Current individual earnings are assumed to be given by:

$$y_i(t) = \exp(\delta_i(t))H_i(t) \exp(-\tau_i(t)),$$

in which  $1 - \exp(-\tau_i(t))$  can be interpreted as the fraction of working time, or alternatively the part of working effort, devoted to investing in human capital. This fraction is increasing in  $\tau_i(t)$ , equal to zero when  $\tau_i(t) = 0$  and equal to one when  $\tau_i(t) = +\infty$  and in this sense, full time learning is a limit case. We call,  $\tau_i(t) \geq 0$ , the level of investment in human capital at time  $t$ .

The technology of production of human capital is described by

$$H_i(t+1) = H_i(t) \exp[\rho_i \tau_i(t) - \lambda_i(t)] \quad (1)$$

in which  $H_i(t)$  is the stock of human capital,  $\rho_i$  is the individual specific rate of return of human capital investments and  $\lambda_i(t)$  is the depreciation of human capital in period  $t$ . Depreciation  $\lambda_i(t)$  embeds individual-specific or aggregate shocks that depreciate previous vintages of human capital. Individual-specific shocks can be negative because of unemployment periods or of layoffs followed by mobility across sectors. These shocks can also be positive when certain components of human capital acquire more value because of voluntary moves across firms or sectors. As the (log) price  $\delta_i(t)$ , the variable  $\lambda_i(t)$  is assumed to be revealed at period  $t$  to the agent and we treat the distribution of  $\lambda_i(t)$  as given. We also defer the presentation of its stochastic and statistical properties.

The human capital technology differs from Ben Porath (1967) in two ways. First, returns  $\rho_i$  to investments are constant in the level of human capital,  $H_i(t)$  although Magnac, Pistoletti and Roux (2013) shows how the extension to non constant returns leads to a more general factor model. Second and more importantly, agents could stop investing in human capital before the end of the horizon unlike in Ben Porath (1967). In the latter model, returns to investments  $\tau_i(t)$  are equal to  $+\infty$  at  $\tau_i(t) = 0$  and then decrease with  $\tau_i(t)$  while these returns are constant in log terms in the current model ( $\frac{\partial \ln H_i(t+1)}{\partial \tau_i(t)} = \rho_i$ ). This means that in Ben Porath (1967), the last marginal unit of investment today is infinitely less productive than the first marginal unit of investment tomorrow. Equalizing marginal productivities of investments today and tomorrow is what uniquely determines investments and those investments are never equal to zero.

Our model relies on a different rationale. Investments are as productive today and tomorrow and the agent decides to stop investing or learning today because effort is costly in utility terms, as specified below. Agents can stop investing before the end of the horizon because costs are too high and this justifies in an exact way the notion of flat spots that Heckman et al. (1998) have used as an approximation in an otherwise standard Ben Porath model. Furthermore, our specification will avoid the "regression to the mean" effect emphasized by Huggett, Ventura and Yaron (2011) that makes individual profiles closer and closer at the end of the working life.



The next step is to formulate a utility flow and the way individuals move assets across time. In order to generate the popular log-linear specification for the earnings equation (e.g. Guvenen, 2009), we assume that period  $t$  utility is equal to current log earnings net of investment costs and that there is no consumption smoothing over time. We return to this assumption at the end of Section 2.2. Period- $t$  utility is written as:

$$\ln y_i(t) - c_i \frac{\tau_i(t)^2}{2},$$

in which the cost of investment is individual specific and quadratic in utility terms. We neglect the linear component of the cost in terms of  $\tau_i(t)$  because it cannot be identified as current log-earnings are:

$$\ln y_i(t) = \delta_i(t) + \ln H_i(t) - \tau_i(t), \quad (2)$$

and the unit in which  $\tau_i(t)$  is expressed, is unobserved. Increasing marginal costs fits well with the interpretation of  $\tau_i(t)$  as an exerted effort which decreases current earnings and provides future returns. This is what makes unique the solution  $\tau_i(t)$  in the dynamic programming.

The decision program of individuals maximizing their discounted expected utility stream over the present and future is given by the following Bellman equation:

$$V_i(H_i(t), \tau_i(t)) = \delta_i(t) + \ln H_i(t) - \left( \tau_i(t) + c_i \frac{\tau_i(t)^2}{2} \right) + \beta_i E_t [W_{t+1}(H_i(t+1))] \quad (3)$$

in which  $\beta_i$  is the individual-specific discount rate and:

$$W_{t+1}(H_i(t+1)) = \max_{\tau_i(t+1)} V_{t+1}(H_i(t+1), \tau_i(t+1)).$$

The terminal condition of this decision program could be written by specifying an individual specific date at which investing in human capital stops as in Ben Porath (1967). We proceed differently by using the dual formulation that the value of human capital stocks at an arbitrary date  $T$  in the future is individual specific.<sup>2</sup> Specifically, we write that at the future date  $T+1$  the value function or the discounted value of utility stream from  $T+1$  onwards is given by:

$$W_{T+1}(H_i(T+1)) = \delta_i^* + \kappa_i \ln H_i(T+1). \quad (4)$$

In this expression,  $\kappa_i$  can be interpreted as the capitalized value of one consumption unit over the remaining period of life after  $T+1$  and:

$$\kappa_i = 1 + \beta_{i,T+2}(1 + \beta_{i,T+3}(1 + \dots))$$

---

<sup>2</sup>This will be the last date of observation in our empirical analysis further on.

in which discount rates  $\beta_{i,t}$  vary with period  $t$  and embody heterogeneous survival probabilities after  $T + 1$ . If we assume that discount factors  $\beta_{i,t>T+1} \leq \beta_i$  e.g.  $\beta_{i,t>T+1} = \beta_i \Pr(\text{Survival at } t)$  then for all  $i$ :

$$\kappa_i \leq \frac{1}{1 - \beta_i}. \quad (5)$$

This suggests that a general interpretation of period  $T + 1$  is as a separating date between a span of periods before  $T$  in which the probability of survival is equal to 1 and a span of periods after  $T + 1$  in which the survival probability is less than one. As human capital investments are embodied, a smaller discount rate is a source of decreasing returns to investment as the original argument by Mincer put it and this explains the concavity of earnings profiles.

In summary, investments are driven by individual specific parameters describing the abilities of agents to earn and to learn. The initial human capital level at time  $t = 1$  is an ability to earn parameters while the returns to investments,  $\rho_i$ , and costs of learning,  $c_i$ , describe the ability to learn since they affect the accumulation process in human capital. Finally, parameter  $\kappa_i$  is the implicit value that individuals place on human capital at the horizon  $T$  and as such, can also be considered as an ability to earn parameter although such an interpretation is less straightforward.

## 2.2 Investment profiles

As time  $t$  denotes the time elapsed since labor market entry or potential experience, we call the sequence of investments between  $t = 1$  and  $t = T$  a life cycle profile of investments. When human capital investments are always positive, this profile is summarized in:

**Proposition 1** *Suppose that :*

$$\beta_i \rho_i \kappa_i > 1, \quad (6)$$

*then:*

$$\tau_i(t) = \frac{1}{c_i} \left\{ \rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i^{T+1-t} \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right] - 1 \right\} > 0, \quad \forall t \leq T \quad (7)$$

**Proof.** See Appendix A.1 ■

Equation (7) expresses the well known result that human capital investments decrease with time. The term in  $\beta_i^{-t}$  indeed means that it is always better to invest earlier than later because the horizon over which investments are valuable is becoming shorter and shorter (Mincer, 1974 and Lillard and Reville, 1999). This is the negative value of  $\kappa_i - \frac{1}{1 - \beta_i}$  (condition (5)) that

commands the intensity of the decrease. In addition, levels of investments increase with returns,  $\rho_i$ , and decrease with costs,  $c_i$ .

Condition (6) ensures that investments in human capital are positive until period  $T$ . Nonetheless, investments could stop before period  $T$ . Because investments are decreasing, the absence of investments in a period  $t$ ,  $\tau_i(t) = 0$ , means that no investments would take place later on,  $\tau_i(t') = 0, \forall t' \geq t$ . In consequence, we can proceed backwards and analyze the conditions under which human capital investments stop before the last period.

**Proposition 2** *There exists an optimal stopping period for human capital investments denoted  $T_i \in \{1, \dots, T\}$  such that:*

$$\forall t \geq T_i, t \leq T, \tau_i(t) = 0, \text{ and } \tau_i(T_i - 1) > 0$$

*if and only if:*

$$\frac{1}{\kappa_{i,T_i}} < \beta_i \rho_i \leq \frac{1}{\kappa_{i,T_i+1}}, \quad (8)$$

where  $\kappa_{i,T} = \kappa_i$  and  $\kappa_{it} = 1 + \beta_i \kappa_{i,t+1}$  for all  $t \leq T$  (and by convention  $\frac{1}{\kappa_{i,T+1}} = +\infty, \frac{1}{\kappa_{i,1}} = 0$ ). Additionally, for all  $t < T_i$ , investments are given by replacing  $\kappa_i$  in equation (7) by  $\kappa_{i,T_i}$ :

$$\tau_i(t) = \frac{1}{c_i} \left\{ \rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i^{T_i-t} (\kappa_{i,T_i} - \frac{1}{1 - \beta_i}) \right] - 1 \right\} > 0, \quad \forall t < T_i. \quad (9)$$

**Proof.** See Appendix A.2 ■

Even if life-cycle investments can stop before  $T$ , the shape of the profile before stopping remains similar. This Proposition also shows that if we had information about the stopping time of human capital investments, we could tie in this information with parameters  $\rho_i$  and  $\kappa_i$ . The cost parameter,  $c_i$ , does not affect the duration of investments but their level only. This is a strong prediction of our set-up and this is because costs do not depend on human capital stocks.

We can now return to the assumption that consumption is not smoothed over time through financial assets. Allowing for consumption smoothing would allow two channels of inter-temporal transfers through financial assets and human capital accumulation. We show in appendix A.3 that the investment equation (7) would include an additional simple function of the savings rate  $s_i(t)$ . When current income is less than (respectively greater than) consumption, human capital investments would be larger (resp. smaller) than in the absence of consumption smoothing. This illustrates the reaction of investments to a change in their opportunity costs (Browning et al., 1999). We cannot use this investment equation however in our empirical application

since we do not have access to consumption data and since the reduced form of endogenous savings depends on other unknown parameters. Moreover, we could not find any specification allowing for consumption smoothing and ensuring that the earnings dynamics equation takes a linear factor format as shown in the next section. Our conjecture is that there does not exist a dynamic model with financial and human capital accumulation that would generate a log-earnings equation of the type we find. In other words, the micro-funded factor model for log-earnings that we derive next and which embeds most equations used in the literature about earnings dynamics is not robust to the presence of consumption smoothing. Nonetheless, we now continue with the setting in which consumption tracks income exactly as could be justified by the evidence gathered by Carroll and Summers (1991).

### 2.3 The Life-cycle Profile of Earnings

We deduce from the investment profile the life-cycle profile of earnings:

**Proposition 3** *If  $T_i$  is the optimal stopping period defined in Proposition 2, log earnings are:*

$$\ln y_i(t) = \eta_{i1} + \eta_{i2}t + \eta_{i3}\beta_i^{-t} + v_{it} \text{ if } t \leq T_i, \quad (10)$$

$$\ln y_i(t) = \ln y_i(T_i) + v_{it} - v_{iT_i} \text{ if } t > T_i \quad (11)$$

in which:

$$\eta_{i1} = \ln H_i(1) - \frac{\rho_i^2}{c_i} \left( \kappa_{i,T_i} - \frac{1}{1-\beta_i} \right) \frac{\beta_i^{T_i+2}}{1-\beta_i} - \frac{\rho_i+1}{c_i} \left( \rho_i \frac{\beta_i}{1-\beta_i} - 1 \right), \quad (12)$$

$$\eta_{i2} = \frac{\rho_i^2}{c_i} \frac{\beta_i}{1-\beta_i} - \frac{\rho_i}{c_i}, \quad (13)$$

$$\eta_{i3} = \frac{\rho_i}{c_i} \beta^{T_i+1} \left( \kappa_i - \frac{1}{1-\beta} \right) \left( \rho_i \frac{\beta}{1-\beta} - 1 \right). \quad (14)$$

and  $v_{it}$  is defined by:

$$v_{it} = \delta_i(t) - \sum_{l=1}^{t-1} \lambda_i(t) = \delta_i(t) - \Lambda_i(t), \quad (15)$$

**Proof.** See Appendix A.4 ■

Proposition 3 shows that the life-cycle profile of earnings can be decomposed sequentially into a first period in which human capital investments are positive and the earnings equation (10) has a non linear factor structure and a second period in which investments stopped and earnings are a function of price and depreciation shocks only.

Initially, earnings given by (10) are the sum of a deterministic and permanent component and a stochastic one. The first component is fully deterministic for the agent because it depends

on individual specific parameters,  $\eta_{i1}$ ,  $\eta_{i2}$ ,  $\eta_{i3}$ ,  $\beta_i$  and potential experience,  $t$ , only. Furthermore, the reduced-form parameters  $\eta_{i1}$ ,  $\eta_{i2}$ ,  $\eta_{i3}$  are functions of the structural parameters which are the primitives in our post-schooling human capital investment model. Firstly, the level of log earnings  $\eta_{i1}$  is affected one to one by the initial human capital stock,  $H_i(1)$  with some correction factors. Secondly, the individual specific growth rate  $\eta_{i2}$  depends positively on the return  $\rho_i$  and negatively on the cost  $c_i$ . Finally, parameter  $\eta_{i3}$  (which depends on  $\left(\kappa_{i,T_i} - \frac{1}{1-\beta_i}\right)$ ) and the discount rate  $\beta_i$  control the degree of curvature of the profile and the effect of the horizon of investment. The closer to zero parameter  $\eta_{i3}$  is, or the closer to 1  $\beta_i$  is, the less curved the profile is.

The stochastic component term  $v_{it}$  in earnings equation (10) as defined in equation (15) is the sum of the (log)-price of human capital,  $\delta_i(t)$  net of the cumulative human capital (log)-depreciations,  $\Lambda_i(t) = \sum_{l=1}^{t-1} \lambda_i(l)$  since labor market entry. We will refer to it thereafter as the net (log)-price of human capital. This component is the source of stochastic dynamics that affects the life-cycle of earnings. Not much structure is needed in the dynamic model on this stochastic component except that it is not under the control of the agent. The developments in this section and the proofs in the appendices are valid (for instance Stokey and Lucas, 1989) under general assumptions like:<sup>3</sup>

$$v_{it} \perp (H_{it}, \cdot, H_{i2}) \mid v_{it-1}, \cdot, v_{i1}, \rho_i, c_i, \kappa_i, \beta_i, H_i(1).$$

In this sense, the derivations of the model above are robust to quite general assumptions on the expectational side of the model as seen from the developments in Appendix A.

Nonetheless, more restrictive assumptions are needed to identify the parameters of the deterministic component in the earnings equation (10) and this is the issue that we now analyze.

### 3 Economic Restrictions and Identification

We present in this section how we restrict equation (10) that describe the life-cycle profile of earnings to a linear factor model that is comparable to specifications used in the literature on earnings dynamics (Meghir and Pistaferri, 2010). We also state the identifying restrictions on the stochastic process of the log-prices of human capital net of its depreciation,  $v_{it}$ , that we will need at the estimation stage later. Finally, we derive the economic structural restrictions that bear on the individual specific reduced form parameters of equation (10).

---

<sup>3</sup>We shall assume additional technical assumptions such as  $E_{t-h}(|v_{it}|) < \infty$  that makes the dynamic program well defined. For the sake of readability these standard assumptions are not fully stated here (see Stokey and Lucas, 1989).

### 3.1 A linear factor specification

We could estimate the individual-specific parameters appearing in the deterministic component of equation (10) as Polachek et al. (2013) do, by specifying orthogonality conditions on  $v_{it}$  and by using a nonlinear earnings function. We did not pursue this path because the estimation of such non-linear expressions is fragile and sometimes difficult to achieve.<sup>4</sup>

This is the foremost reason why we assume that the discount rate  $\beta_i$  is homogenous. The estimation of individual discount rates requires more information from the data than ours can supply (see for instance in an experimental setting, Andersen, Harrison, Lau and Rutström, 2008) or additional restrictions in an observational set-up (see for instance Alan and Browning, 2010). Indeed, the expression of the deterministic component of earnings in equation (10) could be approximated by:

$$\eta_{i1} + \eta_{i2}t + \eta_{i3}\beta^{-t} + \eta_{i4}t\beta^{-t},$$

in which  $\eta_{i4} = -\eta_{i3}(\ln \beta_i - \ln \beta)$  and  $\ln \beta$  is the population average of  $\ln \beta_i$ . Identification of  $\eta_{i4}$  however would rely on the interaction between a linear trend and a curvature term and would probably be fragile.

Furthermore, we conducted experiments that showed that the random effect likelihood function derived in the next section, is flat with respect to the value of the discount rate. This is why we fixed the value of  $\beta$  at 0.95 (as in many other studies, see for instance Heckman et al., 1998). Note however that parameter  $\kappa_i$  indexing the terminal value of human capital partly captures the heterogeneity of discount factors after period  $T$ .

Under these conditions and as long as human capital investments remain strictly positive, the (log) earnings equation (10) can be written as a linear factor model where the three factors are  $f_t = (1, t, \beta^{-t})$  and  $\eta_{i1}$ ,  $\eta_{i2}$  and  $\eta_{i3}$  are individual specific effects or factor loadings:

$$\ln y_{it} = \eta_{i1} + \eta_{i2}t + \eta_{i3}\beta^{-t} + v_{it}. \quad (16)$$

We now turn to the identifying restrictions on the stochastic component  $v_{it}$ .

### 3.2 Human capital prices and depreciation rate

We decompose the net log price of human capital,  $v_{it}$ , into aggregate components and individual specific components. The aggregate components are constructed using groupings in the data

---

<sup>4</sup>In particular, when the number of observations for each individual is limited. Polachek et al. (2013) report that their estimation method did not converge for around 3% of the individuals.

according to skill and age of labor market entry and can be interpreted as market prices net of depreciation for these types of human capital. In contrast, individual specific components are interpreted as individual specific frictions or depreciations. The mechanisms that underlie the specific dynamics of aggregate and individual specific components are allowed to differ and are left unrelated. In our empirical application, we handle the human capital group aggregates and individual specific effects separately and recombine them to recover the full effects.

### 3.2.1 Aggregate components

At the aggregate level of human capital groups, equation (16) can be linearly aggregated into:

$$\overline{\ln y_{gt}} = \bar{\eta}_{g1} + \bar{\eta}_{g2}t + \bar{\eta}_{g3}\beta^{-t} + v_{gt}, \quad (17)$$

in which  $g$  denotes a group defined by skill and age of entry in the empirical application and  $\bar{\eta}_{gk} = E(\eta_{ik} \mid i \in g)$  for  $k = 1, 2$  or  $3$ ,  $v_{gt} = E(v_{it} \mid i \in g)$ . The term  $v_{gt}$  stands for the market log-prices of human capital of group  $g$  at time  $t$  since macro shocks in log prices  $\delta_i(t)$  and depreciation  $\Lambda_i(t)$  are its underlying components. There are no constraints across groups in these dynamics if human capital stocks owned by these groups are imperfect substitutes while perfect substitution holds within groups (Heckman et al., 1998).

We restrict the stochastics of aggregate components by:

$$E(v_{gt} \mid f_t = (1, t, \beta^{-t})) = 0. \quad (18)$$

Condition (18) requires that the net log price dynamics is driven by factors orthogonal to the ones which govern average human capital accumulation and is the key restriction that separates quantities from prices. First, orthogonality with respect to the level is a normalization. Regarding mean independence with respect to the trend, Heckman et al. (1998) and Bowlus and Robinson (2012) use a "flat spot" condition by which (18) is satisfied only for a restricted window of periods close to the end of the working life (around 50) and at which investments and depreciation shocks exactly balance each other. In our empirical application, we have limited data that would allow the full application of such a technique. This is why we will resort to the assessment of robustness of our results to condition (18) by using various earnings deflators among which a series derived from a flat spot procedure. We will also argue that the stability of the between and within group distributions of earnings in France over this period, in contrast to the US and the UK, make assumption (18) more credible.

### 3.2.2 Individual specific components

Turning to the within group dimension, we define centered individual effects by their deviations to their means,  $\eta_{ik}^c = \eta_{ik} - \bar{\eta}_{gk}$ , for  $k = 1, 2$ , or  $3$  and  $v_{it}^c = v_{it} - v_{gt}$ . The earnings equation becomes:

$$u_{it} = \ln y_{it} - \overline{\ln y_{gt}} = \eta_{i1}^c + \eta_{i2}^c t + \eta_{i3}^c \beta^{-t} + v_{it}^c \quad (19)$$

in which  $u_{it}$  is the deviation of individual log-earnings to their group averages ( $\overline{\ln y_{gt}}$ ). Individual specific deviations,  $v_{it}^c$ , stands for frictions in a model of search and mobility. Indeed what Postel-Vinay and Turon (2010) nicely explicit in their presentation is that the dynamics of the earnings process is partly controlled by two other processes which are individual productivity in the current match and outside offers that the agent receives while on the job. In this setting, three things can happen: either earnings remain in the band within the two bounds defined by these processes; or earnings are equal to the productivity process because adverse shocks on that process make employee and employer renegotiate the wage contract; or alternatively, labor earnings are equal to the outside offer in the case the employee can either renegotiate with his employer or take the outside offer if the productivity is lower than the outside option.

We do not model these frictions and posit that they are mean independent of factors and factor loadings:<sup>5</sup>

$$E(v_{it}^c | f_t = (1, t, \beta^{-t}), \eta_i^c) = 0. \quad (20)$$

Note that it lets other moments of  $v_{it}^c$  depend freely on factors and individual effects  $\eta_i^c$ . Nonetheless, this assumption requires for instance that if the depreciation rate has a fixed component,  $\lambda_i$ , it has to be homogeneous within group  $g$ . Otherwise, the individual deviation of prices would exhibit a linear trend with a slope equal to  $\lambda_i - \lambda_g$ , since

$$\Lambda_i(t) - \Lambda_g(t) \propto \sum_{l=1}^{t-1} (\lambda_i - \lambda_g) = (\lambda_i - \lambda_g)(t-1),$$

and this would modify equation (13) relating the growth effect  $\eta_{i2}$  to the structural parameters. Note that if condition (20) does not hold, it does not necessarily affect the linearity of the factor model but invalidates our structural interpretation of factor loadings in terms of returns, costs and terminal values that we pursue here.<sup>6</sup>

---

<sup>5</sup>We slightly relax the assumption of mean independence of frictions with respect to individual effects in the empirical application below by authorizing general initial conditions in the panel data model that we estimate.

<sup>6</sup>The mean independence of frictions,  $v_{it}$ , with respect to factors,  $f_t$ , is the main driver for the identification of structural parameters. The mean independence with respect to individual effects,  $\eta_i^c$ , allows us to justify our random effect procedure but is not necessary in the fixed effect method (see below).



A brief comparison to the empirical literature on earnings dynamics might be useful at this point. This literature aims at fitting the empirical covariance structure of (log) earnings over the life-cycle, i.e.  $u_{it}$  in our notation, using competing specifications like the one described as heterogeneous income profiles (HIP) or restricted income profiles (RIP). Up to now, there is no consensus in the literature about which specification fits the data best (see Baker, 1997, Guvenen, 2007 and Hryshko, 2012). Our linear factor structure embeds both models since the permanent component includes individual specific levels and growth rates of earnings as HIP does and the stochastic component can be any mixture of permanent and transitory shocks like in RIP. Nonetheless, the three factor structure might blur the key identifying assumption about the correlations between first differences of within shocks,  $u$  (for instance Blundell, 2014) because of the presence of the geometric term. In our setting, as long as investments in human capital are positive (i.e.  $t \leq T_i$ ), both HIP and RIP can coexist. Yet, when human capital investments stop, the deterministic individual specific terms should vanish and RIP would prevail.

### 3.3 From the Reduced to the Structural Form

We shall assume from now on, for want of better identification, that investments in human capital are positive until the last period of observation  $T$ , or  $T_i \geq T$ :

$$\tau_i(t) > 0 \text{ for all } t \leq T. \quad (21)$$

so that the econometric model is given by equation (16). This section shows that condition (21) is testable. If this condition did not hold, the life-cycle profile of earnings would be the mixture of two different processes: (1) a generalized heterogeneous growth model driven by human capital investments in periods before  $T$  and affected by the dynamics of human capital prices net of depreciation and frictions; (2) a process driven by the latter dynamics of human capital prices and frictions only. Identification would have to rely on specific distributional assumptions.

We consider from now on that the reduced-form parameters,  $(\eta_{i1}, \eta_{i2}, \eta_{i3})$  are identified although we will return to this issue in Section 5.1. The structural model not only imposes a linear factor structure on the reduced form but also restrictions on these reduced-form parameters.

The non linear system of three equations (12), (13) and (14) have four unknown parameters  $\ln H_i(1)$ ,  $\rho_i$ ,  $c_i$  and  $\kappa_i$  that are by consequence underidentified although structural restrictions can be binding. First, there is no restriction on  $\eta_{i1}$  since equation (12) is the only source of identification of the level of initial human capital  $\ln H_i(1)$ . Second, structural restrictions

consist in statements about the terminal value  $\kappa_i$  or about costs and returns i.e.:

$$\kappa_i \in [0, \frac{1}{1-\beta}], c_i > 0, \rho_i > 0, \quad (22)$$

as well as condition (21). Their implications for the reduced form and the identification of structural parameters are summarized as:

**Proposition 4** *Condition (21) and structural restrictions (22) imply the following restrictions on the individual fixed effects  $\eta_{i2}$  and  $\eta_{i3}$  :*

$$\eta_{i2} > 0, \frac{\eta_{i3}}{\eta_{i2}} \in [-\frac{\beta^{T+1}}{1-\beta}, 0].$$

*Parameter  $\kappa_i$  is identified and:*

$$\kappa_i = \frac{1}{1-\beta} + \beta^{-(T+1)} \frac{\eta_{i3}}{\eta_{i2}}.$$

*Furthermore, parameters  $(\rho_i, c_i)$  are partially identified in the sense that there exist values  $(\rho_i^L, c_i^L)$  such that*

$$\rho_i \geq \rho_i^L, c_i \geq c_i^L.$$

*and a one-to-one relationship:*

$$c_i = c(\rho_i, \eta_{i2}).$$

**Proof.** See Appendix A.5 ■

These results are intuitive. The growth parameter  $\eta_{i2}$  is positive because human capital investments are productive and the curvature term  $\eta_{i3}$  is negative because the horizon is finite and profiles are concave. It is also this curvature relative to the growth term, and therefore the implicit horizon over which investments are valued, which identifies the capitalized value of future returns to human capital after period  $T + 1$ .

Section 5 describes how we deal with estimation and inference in this model of the earnings formation process. We first describe the data with which we estimate such a model.

## 4 A Brief Description of the Data

We first explain how our sample is constructed as well as our measurement for earnings. We propose a brief summary of earnings inequality in France over the last 40 years and present stylized facts about means, variances and autocorrelations of log earnings in our sample.

## 4.1 Sample Selection

Our panel dataset on earnings is extracted from a French administrative source named *Déclarations Annuelles de Données Sociales* (DADS). DADS data is collected through a mandatory data requirement (by French law) for social security and tax verification purposes. All employers must send to the social security and tax administrations the list of all persons who have been employed in their establishments during the year. Firms report the full earnings they have paid to each person but this does not include other labor costs borne by the firm. Each person is identified by a unique individual social security number which facilitates the follow-up of individuals through time although we cannot reconstruct taxes they pay. The tax system is household-based in France and the linking of this dataset with fiscal records is not authorized yet.

The French National Statistical Institute (INSEE) has been drawing a sample from this dataset at a sampling rate of around 4% since 1976. Regarding the sampling device, all individuals who were born in October of even years are included in this sample. Using administrative data is an important advantage since these data are less subject to attrition or measurement errors.<sup>7</sup> Unlike survey data, the collection of information does not rely on individual response behavior and individuals are better followed over time. Moreover, the large sample size enables us to use a single large cohort of individuals who entered the labor market in the same year. For simplicity, we shall, in the following, use "cohort" as standing for a labor market entry cohort.

Observations can yet be missing for different reasons. First, data were not collected in three years (1981, 1983 and 1990) for reasons specific to INSEE. Second, this dataset is restricted to individuals employed in the private sector or in publicly-owned companies. As a consequence, this analysis is restricted to individuals who have been employed at least one year between 1976 and 2007 in the private sector or in a publicly-owned company. Third, the coding of individual identifiers (on 13 digit positions) was at times manually entered at the beginning of the data collection and this causes errors.

We now describe our sample selection and the construction of earnings. Further details can be found in a supplementary Appendix available upon request (specifically in Section A.IV.1 of it).

As in Le Minez and Roux (2002), we consider individuals right from their entry into the labor market and onwards. Labor market entry is defined as being employed for more than 6 months and being still employed the following year, possibly in different firms. For the entry

---

<sup>7</sup>To our knowledge, the only other contributions in the earnings literature that use administrative data is Hofmann (2013) and Daly, Hryshko and Manovskii (2014) (German and Danish data).

cohort of interest which starts in 1977, this leads us to select from the administrative data 36,883 individuals who were employed more than 6 months in 1977 and at least one day in 1978. Among them, 53% have worked but not permanently before. Conversely, individuals who have worked in 1977 are not considered as entrants if their jobs are not permanent enough. They may however enter with a subsequent cohort.

In addition, we aim at keeping employees with a permanent full-time attachment to the private sector only. Firstly, we consider workers employed full time only and we censor information about part-time jobs. In addition to the condition which requires workers to work in the private sector during the year of entry and the following one, we further restrict the sample to men also working in 1982 and 1984. This is because we want to avoid dealing with non participation issues for females and with too many exits from the sample since the bulk of entries into public service occurs at the beginning of the working life. These restrictions lead us to retain in the 1977 entry cohort 16,091 men who entered the labor market in 1977 in a full-time position for more than 6 months and who were also full-time employed in 1978. Adding the condition on the presence in a full-time position in 1982 and 1984 further restricts the sample to 8,288 individuals. Finally, we keep only workers who were aged between 16 and 30 at their entry in the labor market and this restricts the sample to 7,446 workers.

We impose these restrictions in order to concentrate on a relatively homogeneous sample of workers with a long term attachment to the private firms' labor market. Admittedly, it does not represent the full working population. Because of the lack of a credible identification strategy to correct for selection, we shall assume that selection is at random or can be conditioned on individual-specific effects only. The distribution functions of unobserved factor loadings or idiosyncratic components that we estimate in the following refer to this subpopulation.

The empirical analysis uses "annualized" earnings. It is defined as the sum of all earnings during the year divided by the number of days worked and remultiplied by 360 (total number of days during the year in the administrative data). Accounting only for total yearly earnings would miss other earnings from employment in the public sector, self-employment income or unemployment benefits that are not observed in the data. Considering annualized earnings instead limits this problem, although it may lead to overestimating yearly income. In order to weaken the possible impact of measurement error, we coded as missing the first and last percentiles of the earnings distribution in every period.

A shortcoming of administrative data is that few observable characteristics are available apart from age at labor market entry and the skill level of the first job. We consider the interaction

between the age at entry and the skill level of the entry job as defining our human capital groups. Skills are grouped into three categories based on the codification in 2 digits provided in the data: high-skill (managers, professionals), medium-skill (blue-collar or white-collar skilled workers) and low-skill jobs. Due to the specific sampling of the dataset as explained above, the age at entry of workers (in 1977) can only take odd integer values from 17 to 29, i.e seven different values. It is likely that workers delaying entry have a higher education level than the ones who entered earlier. Interacting skill levels and ages of entry, we end up with 20 groups, one being too small to be considered on its own (see Table 4 for the definition and size of each group). Since these groups are defined according to characteristics recorded at the entry on the labor market, individuals are attached to the same group during their whole working life.

## 4.2 Earnings Inequality in France

The sharp increase in earnings inequality in the UK and in the US over the last thirty years is a well known empirical fact (see for example Autor, Katz and Kearney, 2008, or Moffitt and Gottschalk, 2011, for the US and Blundell & Etheridge, 2010, for the UK). Yet, the picture is more balanced in other OECD countries and while some European countries have experienced an increasing dispersion in earnings, others have not been affected by this trend and have had stable or decreasing dispersion. Atkinson and Morelli (2012) compute international earnings inequality comparisons over the second half of the twentieth century for 25 countries. As regards European economies, they conclude that earnings inequality has increased in Germany, Italy, Portugal, Sweden, Switzerland while in Finland, France, Netherlands, Norway, and Spain earnings dispersion has stayed constant or decreased over this period.

In France, earnings inequality in 2010 is broadly comparable to its level in the sixties and if anything has decreased. Atkinson and Morelli (2012) report an unchanged Gini coefficient for earnings over the period. Using Labour Force Surveys (LFS), they also compute yearly measures of inequality and show a very stable inequality level. Using two different datasets, the DADS and the French LFS Verdugo (2014) concludes that the two data sets provide strikingly similar figures of constant or decreasing earnings dispersion between 1964 and 2005. Verdugo (2014) decomposes the total earnings dispersion into upper and lower-tail earnings inequality. The dispersion in the top of the distribution has remained constant since the P90/P50 index in earnings fluctuates around 2, while the dispersion at the bottom measured by the P50/P10 index has decreased from 1.9 to 1.5. Charnoz, Coudin and Gaini (2011) also use the DADS data to reach the same conclusion that earnings inequality in France has been rather stable from 1976

to 1992 and has been slightly decreasing from 1995 to 2004.<sup>8</sup> This stability has been attributed, at least partly, to a strong policy driven increase in education at the end of the 1980s and labor market policy regulations at the end of the 1990s (Charnoz, Coudin and Gaini, 2014).

A note of caution is in order. While these studies consider changes in the cross-sectional earnings distributions, changes in the structure of the population that has been given a large role by the previous studies are neutralized in this paper. We adopt a different perspective by following a single cohort of individuals entering the labor market in 1977.

### 4.3 Description of earnings dynamics

Table 1 reports descriptive statistics of the sample. The sample size is 7446 observations in 1977 and 4670 in 2007. The human capital groups defined above are of unequal size, the groups with an early age of entry being the largest ones and the late age of entry groups the smallest. Attrition follows a somewhat irregular pattern which is partly due in the first years to our sampling design since we require that wage earners be present in 1977, 1978, 1982 and 1984 (see appendix, Table A.i). Some years are also completely missing (1981, 1983 and 1990). There are also more surprising features for instance in 1994 (or 2003 at a lesser degree) a year in which many observations are missing. This is due to the way INSEE reconstructed the data from the information in the original files.

We report in Figure 1, the increase of average log-earnings over the period in 2007 euros for three age of entry groups ( $< 20$ ,  $\geq 20$  and  $\leq 24$  and  $> 24$ ) since the evolution for smaller groups is similar.<sup>9</sup> The covariance matrix of log-earnings that we consider from now on are computed by taking deviations of (log) earnings with respect to their means in the 20 groups defined by age of entry, skills and periods.

The left panel of Figure 2 represents the change in the cross-sectional variance of (log) earnings residuals for the full sample, while the right panel represents the variance by age of entry groups.<sup>10</sup> The first few years witness a strong variability of earnings. Until the sixth year

---

<sup>8</sup>These diagnostics are relative to labor earnings inequalities and not to income inequalities. In particular, income inequality has increased in France over the period as in many other countries, albeit at a lesser extent, and the growth of incomes at the very top is the main reason for this increase (Atkinson, Piketty and Saez, 2011).

<sup>9</sup>Inflation, as measured by consumer prices, leads to a subtracting factor for current log-earnings over the whole period which is equal to 1.17. This can be roughly subdivided into two sub-periods between 1977 and 1986 in which this factor is equal to .77 and between 1986 and 2007 during which inflation levelled off and this factor is equal to .40.

<sup>10</sup>Choosing the variance as a description of the process is adapted to the random effect specification that we estimate. Using other inequality indices (Gini, Theil or Atkinson) does not change the qualitative features of our descriptions.

of observation, 1982 (respectively the fourth, 1980), the variance of log earnings drops for the low skill groups (resp. for the other groups) whereas it increases gradually over the rest of the sample period till around 1995. The variance profile is flat afterwards in contrast to the US (PSID) where it continues to grow (Rubinstein and Weiss, 2006). From the right panel one can notice that late entrants in the labor market experience higher levels and larger rates of growth for the variance of (log) earnings over the life-cycle. The complete covariance matrix of log earnings residuals is reported in Table 2 although this is easier to use graphs to describe the main features of earnings autocorrelations.

Figure 3 displays for the full sample the autocorrelation of residuals of log earnings with residuals in an early (resp. late) year, 1986 (resp. 2007). This Figure reveals an asymmetric pattern over time which is quite robust to the choice of these specific years (1986 and 2007). The correlation between earnings in year  $t$  and in 1986 is swiftly increasing when  $t$  is before 1986 and this is also true for 2007 albeit at a lesser degree. Meanwhile the correlation between earnings in 1986 and in year  $t$  is only slowly decaying in  $t$  if  $t$  is after 1986. Figure 4 takes a different view that confirms the previous diagnostic by plotting the autocorrelations of order 1 and 6. Note that their shape are very similar and increase uniformly over time although at different levels. The closer to the end of the working life, the larger the autocorrelation coefficients are.<sup>11</sup>

## 5 Econometric Modeling of Earnings Dynamics

In this section we summarize our empirical strategy. Most of the technical details are relegated to appendices. Our first objective is to recover estimates of the individual effects  $(\eta_{i1}, \eta_{i2}, \eta_{i3})$  in the linear factor structure (16) using the identifying restrictions presented in Section 3.2. In a second stage, we turn to the test of structural restrictions and the estimation of structural parameters. We finally use those estimates to compute counterfactuals.

We start from the linear factor model given by equation (16). Stacking log earnings  $\ln y_{it}$  and the stochastic component  $v_{it}$  into  $T$ -vectors  $\ln y_i$  and  $v_i$  as well as  $\eta_{ik}$  into a 3-vector  $\eta_i$ , this equation writes:

$$\ln y_i = M(\beta) \eta_i + v_i$$

in which  $M(\beta)$  is a  $T, 3$  matrix in which a constant, a linear and a geometric term are stacked.

Estimating individual effects  $\eta_i$  raises two difficulties. The first one, which is standard in the

---

<sup>11</sup>In the supplementary appendix available upon request, we provide graphs showing that this cohort has nothing specific when it is compared to younger cohorts entering later the labor market (Appendix A.IV.2).

panel data literature, is that these parameters can be estimated using the temporal variability of the observations for every individual only (at most 28 observations in our data). Let  $N_i$  the number of observations for individual  $i$ .<sup>12</sup> Even if OLS estimates of parameters  $\eta_i$  for every individual profile are consistent when  $N_i$  becomes large, the bias in  $1/N_i$  and lack of precision if serial correlation is sizeable might lead to estimates of poor quality. The second issue is that we made a case in Section 3 for distinguishing aggregate and idiosyncratic components, both having their own dynamics.

Hence, we split the estimation in two stages. First, we estimate aggregate equation (17) group by group using standard time series methods. At this aggregate level, we have 28 observations per group. Second, within-group centered individual specific parameters,  $\eta_i^c$ , are parameters in the rewriting of equation (19) as:

$$u_i = M(\beta) \eta_i^c + v_i^c. \quad (23)$$

in which the matrix of factors  $M(\beta)$  is defined above. Because some data are missing and serial correlation is likely, inference might be poor if we estimate this equation individual by individual.

To overcome this,  $\eta_i^c$  are estimated using a two-step strategy which consists in estimating a flexible random effect model using the whole sample first and second, in estimating equation (23) by FGLS, individual by individual. In the second step fixed effect procedure, the FGLS weight is the inverse of the population covariance matrix of  $v_i^c$  over time estimated in the first random effect step. The final individual specific estimates are obtained by adding aggregate estimates to these fixed effect estimates of the centered factor loadings.

On the one hand, using random effect in a first step allows us to control for general serial correlation in the fixed effect estimation so that the latter estimates are presumably more precise. Even if the random effect specification is only an approximation of a more complicated data generating process, for instance obtained by aggregating individual specific processes, we argue that using this approximation enhances the precision of the fixed effect estimations that we implement in the second stage provided that these shocks are mean independent of the permanent individual effects as stated in condition (20). Furthermore, random effect estimation provides a benchmark against which we can assess the amount of bias in the fixed effect estimation due to the finite length of the observation period for each profile.

On the other hand, looking at covariance restrictions only and estimating by random effect methods is overly restrictive. In particular, structural constraints on parameters are not easily

---

<sup>12</sup>By construction,  $N_i$  is less or equal to  $T$ .



imposed.<sup>13</sup> This is what fixed effect estimation delivers in the second step. Structural restrictions of Proposition 4 become easily testable and structural estimates are easily computed.

We continue with presenting random effect, fixed effect and aggregate effect estimation methods in more detail.

## 5.1 Random Effect Estimation

For expository purposes, we neglect in the text initial conditions that are rigorously dealt with in Appendix B, and in the supplementary appendix. We can use equation (23) and mean independence restrictions (20) to show that:

$$\begin{aligned} E(u_i | \eta_i^c) &= M(\beta)\eta_i^c, \\ V(u_i | \eta_i^c) &= V(v_i^c | \eta_i^c) \equiv \Omega(\eta_i^c), \end{aligned}$$

and therefore:

$$V(u_i) = V(E(u_i | \eta_i^c)) + E(V(u_i | \eta_i^c)) = M(\beta)V(\eta_i^c)M(\beta)' + E(\Omega(\eta_i^c)). \quad (24)$$

Our parameter of interest in this equation is the covariance matrix of the individual effects,  $V(\eta_i^c)$  and specifically, the correlations between level, growth and curvature effects. Identifying the covariance matrix requires restrictions on the variance of the idiosyncratic errors,  $E(\Omega(\eta_i^c))$ . An ARMA specification is most common in the earnings dynamic literature and generally low orders are used (see Guvenen, 2009, or Hryshko, 2012) whereas alternatives could be the composition of permanent and transitory shocks with specific structures (Bonhomme and Robin, 2009, Lochner and Shin, 2013) or general factor models (Bai, 2009). Furthermore, Arellano and Bonhomme (2012) show that a finite lag ARMA specification is sufficient to get identification of  $V(\eta_i^c)$ . We use this result and proceed by specifying that the processes  $v_{it}^c$  belong in the family of time-heteroskedastic ARMA processes although we limit the orders of the AR and MA to vary between 1 and 3. This allows us to assess the robustness of our results about the covariance of individual effects,  $V(\eta_i^c)$ , to the orders of the ARMA process. Moreover, we allow for time heteroskedasticity of the innovations whose importance is shown by Alvarez and Arellano (2004).

What the decomposition (24) exhibits in addition, is that a restricted form of individual heterogeneity, possibly dependent on parameters,  $\eta_i$ , could be allowed in the ARMA process

---

<sup>13</sup>One route would be to use deconvolution techniques as in Bonhomme and Robin (2010) or Cunha et al. (2010) although this would require the development of estimation methods for distributions under structural constraints.

provided that the expected value,  $E(\Omega(\eta_i^c))$ , remains in the ARMA family that we consider.<sup>14</sup>

Appendix B presents the full specification of the process for  $v_{it}^c$  in which we also deal with initial conditions in the most general way used in the dynamic panel data literature. The covariance matrix of initial conditions is free as well as the covariance between those initial conditions and the individual specific parameters  $\eta_i^c$ . Further details are given in the supplementary Appendix A.II.

The most commonly used minimum distance method for estimating equation (24) as in Abowd and Card (1989) is severely small-sample biased since the range of moments involved when the time dimension becomes large makes first order asymptotics a poor guide in empirical research (see Arellano, 2004 for a review). Okui (2009) derives the small sample biases not only in the mean but also in the variance of GMM estimates due to the presence of too many moments and he suggests some moment selection mechanism. This is why some researchers proposed to return to an OLS set up adding a bias correction step (Hahn and Kuersteiner, 2002) or to maximum or pseudo-maximum likelihood methods that reduce the number of moments available (Hsiao, Pesaran and Tahmiscioglu, 2002, Alvarez and Arellano, 2004).

Specifically, the estimation method proposed by Alvarez and Arellano (2004) seems to dominate in Monte Carlo experiments other fixed  $T$  consistent estimators such as the maximum likelihood estimator using differenced data and the corrected within group estimator. This method is particularly well adapted to the case in which there are missing data in earnings dynamics. For any missing data configuration, it consists in deleting the rows and columns of the covariance matrix corresponding to missing data and write the likelihood function accordingly. Random effect estimates remain consistent if data are missing at random.

Under a normality assumption, the implicit moment selection underlying this estimation method is optimal and though the method loses optimality in the non-normal cases, it is still useful for moment selection and for small-sample bias reduction (Okui, 2009).

## 5.2 Fixed effect estimation of $\eta_i^c$

Random effect estimates can now be exploited to construct individual specific estimates of parameters  $\eta_i^c$ . First, if the ARMA model that we retained above is the correct specification at the individual level for all individual profiles, we obtain fixed effect estimates as linear combi-

---

<sup>14</sup>We also neglect other non linearities. Contributions to the non linear analysis of earnings dynamics include among others Geweke and Keane (2000), Hirano (2002), Bonhomme and Robin (2009), Browning et al. (2012), Mehri and Pistaferri (2004) or Hospido (2012).

nations of residuals  $u_{it}$  and those FGLS estimates are optimally weighted to account for serial dependence. Supplementary appendix A.III.1 develops the analytic computations that lead to the following individual effect estimates:

$$\hat{\eta}_i^c = \hat{B}u_i, \quad (25)$$

in which matrix  $B$  is a function of random effect parameters, in particular the covariance between the individual parameters and the initial conditions, and  $\hat{B}$  its plug-in estimate.

Even if the ARMA model is incorrect, those estimates are still consistent when  $N_i \rightarrow \infty$  because what matters is the mean independence of the factors with respect to the residuals, and not the specific form of serial dependence. Nonetheless, their standard errors should be corrected. We use Newey-West robust standard errors in the empirical section. Taking serial dependence as given by random effect estimates into account yet exploits the information that we have about "aggregate" serial dependence as opposed to a simple OLS or non-linear least square estimate (Polachek et al, 2013). It enhances the quality of the estimates if the term  $\Omega(\eta_i^c)$  is not too heterogeneous and this will be checked after estimation.

Consistency properties could nonetheless be misleading since  $N_i$  varies in our sample between 4 and 28. To assess the magnitude of the bias, we shall compare the estimates of the covariance matrix of  $\eta^c$  that we obtained by random effect and by fixed effect methods by grouping individual profiles according to the length of the observation period. The bias for the estimated variance of earnings can also be computed as in Arellano and Bonhomme (2012) and corrected. Abstracting first from sampling errors, an unfeasible estimate is defined as:

$$\tilde{\eta}_i^c = Bu_i = \eta_i^c + Bw_i,$$

in which random vector  $w_i$  has mean zero conditionally on  $\eta_i^c$  and covariance matrix,  $\Omega_w$ . These objects and this expression are defined and derived in the supplementary appendix A.III.1.<sup>15</sup>

We have:

$$\begin{aligned} V(\tilde{\eta}_i^c) &= EV(\tilde{\eta}_i^c | \eta_i^c) + VE(\tilde{\eta}_i^c | \eta_i^c) \\ \implies V(\tilde{\eta}_i^c) &= B\Omega_w B' + V(\eta_i^c). \end{aligned}$$

The bias term is given by  $B\Omega_w B'$  and it is easy to show that the dominating term is of order  $1/N_i$ .<sup>16</sup>

<sup>15</sup>The new notation  $w_i$  is introduced since it differs from  $v_i^c$  in equation (23) because of the correlation of initial conditions with  $\eta_i^c$ .

<sup>16</sup>Because our factors are a constant, a linear trend and a geometric one, there are also bias terms in  $1/(N_i)^2$  and  $\beta^{N_i}$  that are dominated by the leading one,  $1/N_i$ .

Our estimate has an additional bias term which is given by the measurement equation:

$$\hat{\eta}_i^c = \hat{B}u_i = \tilde{\eta}_i^c + (\hat{B} - B)w_i,$$

although this term is in  $1/\sqrt{N}$  and thus dominated in large  $N$  and moderate  $N_i$  samples by the bias in  $1/N_i$ . These biases can also be estimated and bias-corrected estimates of those variance terms can be constructed.

### 5.3 Estimation of average human capital effects, $\bar{\eta}_g$

We use simple time series techniques to derive estimates of average effects for each group  $g$  defined by equation (17):

$$E(\log(y_{it}) \mid i \in g) = \bar{\eta}_{g1} + \bar{\eta}_{g2}t + \bar{\eta}_{g3}\beta^{-t} + v_{gt}$$

As explained in section 3.2, the time series properties of  $v_{gt}$  have nothing in common with the time series properties of  $v_{it}^c - v_{gt}$ . Furthermore, groups do not have necessarily things in common since substitution is imperfect between human capital stocks across group in the aggregate production function. This is why we estimate parameters in each group by simple OLS as justified by condition (18). This provides consistent estimates of  $\bar{\eta}_g$ , say  $\widehat{\bar{\eta}}_g$ , and standard errors are computed using a Newey West procedure.

We then recompose estimates of the individual effects  $\hat{\eta}_i$  by adding estimates of aggregate components and of individual effects defined in the previous section :

$$\hat{\eta}_i = \widehat{\bar{\eta}}_g + \hat{\eta}_i^c.$$

### 5.4 Constraints and Structural Parameters

We now show how to impose structural constraints on individual-specific estimates as derived in Proposition 4. Indeed, estimates  $\hat{\eta}_i$  do not necessarily satisfy the constraints:

$$\eta_{i2} > 0 \text{ and } \frac{\eta_{i3}}{\eta_{i2}} \in \left[-\frac{\beta^{T+1}}{1-\beta}, 0\right].$$

We let  $\pi_T = \frac{\beta^{T+1}}{1-\beta}$  and write these constraints as:

$$H_{0i} : \eta_{i2} > 0, \eta_{i3} < 0 \text{ and } \eta_{i3} + \pi_T\eta_{i2} > 0.$$

As we know the asymptotic distribution of each factor loadings, we can test each single restriction (e.g.  $H_{0i}^{(1)} : \eta_{i2} > 0$ ) at the individual level. We can also jointly test for the three constraints and this leads to the construction of constrained estimates.

### 5.4.1 Constrained estimates

We construct constrained estimates by projecting unconstrained estimates on the set of restrictions using the distance defined by the (log)-likelihood function criteria as explained in Appendix A.III.3. We can then construct the distribution of the distance in the data between the unconstrained and the constrained estimates,  $\hat{\eta}_i^R$ :

$$d(\hat{\eta}_i^R, \hat{\eta}_i) = (\hat{\eta}_i^R - \hat{\eta}_i)' \hat{\Omega}_{\eta_i}^{-1} (\hat{\eta}_i^R - \hat{\eta}_i),$$

in which  $\hat{\Omega}_{\eta_i}$  is the estimate of the covariance matrix of the individual fixed effects  $\Omega_{\eta_i}$ . This distance is the basis for a Quasi-Likelihood Ratio test of all structural restrictions (e.g. Silvapulle and Sen, 2005). The distribution of this statistic under the null hypothesis  $H_{0i}$  is a mixture of chi-square distributions that can be simulated.

### 5.4.2 Structural estimates

Constrained estimates are by construction on the frontier of structural restrictions when the unconstrained estimates are outside the set of structural constraints. This happens quite often even when the null hypothesis is true because the number of observed periods  $N_i$  is not large enough. For instance, it could be that constrained estimates verify the constraint,  $\eta_{i3} + \pi_T \eta_{i2} = 0$ , which would mean that the estimate of parameter  $\kappa_i$  is equal to 0. Because  $\rho_i > 1/\beta\kappa_i$ , this would generate an infinitely large estimate for  $\rho_i$ .

This is why we use simulation to sample into the asymptotic distribution of constrained estimates. Under the additional assumption that  $v_i^c$  are normally distributed, we use that the likelihood function of an individual earnings profile is given by:

$$L(\eta_i^c | u_i) = H(u_i) \exp \left( -\frac{1}{2} (\eta_i^c - Bu_i)' \Omega_{\eta}^{-1} (\eta_i^c - Bu_i) \right) L_0(\eta_i^c),$$

in which structural restrictions are implicitly stated in the prior distribution  $L_0(\eta_i^c)$ . We draw into this posterior distribution to construct simulated constrained estimates,  $\hat{\eta}_i^s$ , of  $\eta_i$  using the developments in Appendix A.III.4. We then take the average of these simulated values to construct our simulation estimates of factor loadings which by construction are interior points of the constrained set.

## 5.5 Counterfactual analysis

We assume that there is a "technological" improvement in survival probabilities in such a way that there are additional  $K$  years after period  $T$  during which the survival probability remains

equal to 1 (instead of starting declining). This amounts to the transformation of  $\kappa_i$  into  $\kappa_i^*$ :

$$\kappa_i^* - \frac{1}{1-\beta} = \beta^K \left( \kappa_i - \frac{1}{1-\beta} \right)$$

as if we were prolonging, all of a sudden, life expectancy by  $K$  years. Other parameters like returns,  $\rho_i$ , and costs,  $c_i$ , are held fixed. We evaluate the consequences on the earnings profiles of these changes as if these news had been revealed at time  $t = 1$  so that the initial level of human capital would also remain the same. We assume that human capital price dynamics, net of human capital depreciation, are not changed by these news, so that the stochastic process of shocks  $v_i$  also remains the same.

Evaluating individual parameters (12) to (14) at the new values  $(\kappa_i^*, \rho_i, c_i, H_i(1))$  yields that the new values  $(\eta_{1i}^*, \eta_{2i}^*, \eta_{3i}^*)$  are such that  $\eta_{2i}^* = \eta_{2i}$ ,  $\eta_{i3}^* = \beta^K \eta_{i3}$  and that:

$$\eta_{i1}^* - \eta_{i1} = -\frac{\rho_i^2}{c_i} \left( \kappa_i - \frac{1}{1-\beta} \right) \frac{\beta^{T+2}}{1-\beta} (\beta^K - 1). \quad (26)$$

In order to abstract from the idiosyncratic noise of transitory earnings which remains fixed by assumption, we shall then compare the earnings variance profile  $V(M(\beta)\eta_i^*)$  with the original profile of  $V(M(\beta)\eta_i)$  in which  $M(\beta)\eta_i$  is the permanent component of the earnings profile.

Nonetheless, parameters  $\rho_i$  and  $c_i$  in equation (26) are only partially identified. A lower bound  $(\rho_i^L, c_i^L)$  on their values can be estimated and used to construct the counterfactual. We shall then proceed by making different assumptions like  $\rho_i = \rho_i^L$ ,  $\rho_i = 1.20\rho_i^L$  etc to assess the robustness of this construction.

## 6 Results

We comment the estimation results of the within groups earnings equation by random effects in Section 6.1 and report the estimates of the aggregate human capital group effects in Section 6.2. In Section 6.3 we detail the procedure we implement to estimate unconstrained individual factor loadings or fixed effects and present descriptive statistics of these estimates. Next, we test and impose structural constraints on estimates. This leads us in Subsection 6.5 to the estimation of structural parameters which are identified (the terminal value coefficient) or partially identified (rates of return). We wind up the section with robustness checks and other diagnostics.

### 6.1 Random effect estimates

We focus on the estimation of the covariance matrix of centered individual effects,  $\eta_i^c$ , and the goodness of fit of estimates. All other results are detailed in the supplementary appendix.

The estimated covariance matrix of the centered individual effects is quite stable across the different specifications of ARMA processes (see Table 3). Their standard deviations are very precisely estimated at around .30 for the fixed level factor,  $\eta_1^c$ , and .25 for the geometric factor,  $\eta_3^c$ , and at around .04 for the linear trend factor,  $\eta_2^c$ . The correlation between the linear trend and geometric factors is very strongly negative and equal to  $-.95$  consistently across ARMA specifications. The magnitude of this correlation and its sign are consistent with the structural model that ties in  $\eta_{i2}$  and  $\eta_{i3}$ :  $\eta_{i3} = (\kappa_i - \beta) \beta^{T+1} \eta_{i2}$  (Proposition 4). More simply, the higher the growth in earnings is, the more curved the earnings profile is.

The correlation coefficient between the geometric,  $\eta_3^c$ , and the level,  $\eta_1^c$ , factor loadings is also significantly negative – around  $-0.6$  – and the one between the level and linear trend factor loadings,  $\eta_1^c$  and  $\eta_2^c$ , is positive and around .4. The sign of the latter correlation coefficient is to be expected if the level of human capital at the entry date is positively correlated with the returns to human capital which govern the factor loading of the linear trend. This correlation has been examined in the previous literature (see table 4 in Guvenen, 2007, for a summary of the results) and has not always been found positive (for instance, Hause, 1980).

Both the introduction of the curvature factor and the control of initial conditions seems to contribute to this finding. Irrespective of the order of the ARMA process, the initial conditions are negatively correlated with the fixed level factor  $\eta_1^c$ , positively with the linear trend factor  $\eta_2^c$  and negatively with the geometric factor  $\eta_3^c$ .<sup>17</sup> These initial conditions account for the strong transitory conditions that seem to affect the earnings process at the beginning of the working life (as well as our data selection process)<sup>18</sup>. Even if the earnings process is asymptotically stationary, initial conditions are for this reason not necessarily set on the stationary path that corresponds to this process (see Magnac and Roux, 2010).

Goodness-of-fit is examined in different graphs. In Figure 2, we report how the estimated variances as well as the observed variances evolve over time. They fit very nicely in the first part of the sample (until 1994) but this breaks down after 1994 after which the shape of the evolution of variances is similar albeit at a level which is higher than the observed level. It confirms that 1994 is an abnormal year even if the goodness-of-fit for autocorrelations is good as reproduced in Figures 3 and 4.

---

<sup>17</sup>Estimates of the covariances between the factor loadings and the initial conditions are presented in Table A.iv in the supplementary Appendix.

<sup>18</sup>The strong decrease of the variance observed during the first years might partly be due to the very stringent selection made in the 1977 entry cohort. The very flexible initial conditions as they are accounted for in the random effect estimation control partly for this selection.

We tried different mechanisms in order to understand better the discrepancy between observed and predicted variance profiles. One possibility is to allow for an additional measurement error term in 1994 for instance, like in Guvenen (2009) or to drop this year altogether. These attempts did not affect goodness-of-fit. A more disturbing explanation for those discrepancies is that it reflects a failure in the missing at random hypothesis. When one represents the evolution of earnings variance over the life-cycle using fixed effect estimates (see below), it clearly appears that the level of these profiles negatively depends on the number of periods we observe each person. Variances are larger for individuals who are present in the panel during shorter spells and random effect estimates overestimate the sample variance. Nevertheless, correcting for non random attrition seems out of the scope of this paper and we leave it for further research.<sup>19</sup>

## 6.2 Average effects estimation

Table 4 present for each of the 20 groups the estimations of the average effects by OLS. We neglect the possible impact of initial conditions and assume that the transitory conditions at the beginning of the working life are pure idiosyncratic frictions and do not affect market prices for the different human capital groups that we observe. Results exhibit the expected patterns. The first factor loading average  $\bar{\eta}_{g1}$  ranges from 2.4 for the lowest skill groups to 3.4 for the highest skill groups, i.e. the workers who have their first durable high-skilled job at the age of 27. The estimated coefficient of the linear trend average,  $\bar{\eta}_{g2}$  ranges from 0.017 to 0.07. As the previous average, it is larger for the high-skilled groups than for the low-skilled ones although the evidence is weaker. The geometric factor loading average  $\bar{\eta}_{g3}$  is negative as expected or non significantly different from zero.

Interestingly the pattern of correlations of the average effects across human capital groups (weighted by group size) is very close to the correlation pattern of centered factor loadings estimated by random effects as presented in Table 3. The coefficient of correlation between  $\bar{\eta}_{g1}$  and  $\bar{\eta}_{g2}$  (i.e.  $g$  varying) is equal to 0.66 and close to the random effect estimate of the correlation between  $\eta_1^c$  and  $\eta_2^c$ , which is equal to 0.5. The estimated correlation between  $\bar{\eta}_{g1}$  and  $\bar{\eta}_{g3}$  is negative,  $-0.64$  and very close to the random effect estimate,  $-0.636$ . Finally, the estimated correlation between  $\bar{\eta}_{g2}$  and  $\bar{\eta}_{g3}$  of  $-0.96$ , is very close to the random effect estimate. We interpret this similarity of patterns as evidence that human capital investment patterns between and within groups are similar in France in contrast to what was found in the US (Heckman et

---

<sup>19</sup>Furthermore, the conditions for consistency for the fixed effect estimates described below are less stringent since the missing at random assumption can be weakened and taken as conditional on individual effects.



al., 1998). This might be due to the stability of relative human capital prices over this period.

### 6.3 Fixed effect estimation

We now turn to the results of FGLS estimation of the three individual factor loadings which uses random effect estimates to correct for aggregate serial dependence as derived in equation (25). Technical details are given in the Supplementary Appendix. Estimated group averages as described in the previous section are added to within group estimates to reconstruct the final estimates of individual effects  $\hat{\eta}_i$  as described in Section 5.3.

As said, fixed effect estimates are consistent when the number of observed periods  $N_i$  tends to infinity. Table 5 presents the estimates of quantiles of their distributions in subsamples constructed according to the number of periods individuals are observed (between 4 and 28). The bias in  $1/N_i$  is noticeable as the larger the number of observed periods is, the lower the inter-quartile ratio for all three factors is. Notwithstanding those differences, the median of the coefficient attached to the level factor is of the order of magnitude of the mean (log-)earnings at around 2.5 and the range between the 20th and 80th percentile is .5 if the number of observed periods is maximal ( $N_i = 28$ ). The median of the coefficient of the linear trend factor is of the order of 3 or 4% while its 20-80 quantile range is about 6-8%.

Finally, the median of the coefficient of the geometric factor lies around -.17 and its inter-quantile range is about .40. This coefficient enters multiplicatively in the curvature of the earnings profiles over time since the second derivative of the latter with respect to potential experience is this coefficient multiplied by  $(\log \beta)^2 = 2.5 \cdot 10^{-3}$ . This fits well with the usual estimates of earnings equations predicting the maximal value of earnings at a time  $t$  close to  $\log(\log(\beta)\eta_2/\eta_3)/\log \beta$  which is equal to 31.2 at the median estimates. Furthermore, the return to potential experience at labour market entry is  $\eta_2 - \log(\beta)\eta_3$  and the median estimate is in the range of 2 to 3%.

Table 6 presents estimates of the covariance matrix of centered individual effects or factor loadings obtained by fixed and random effect methods. Standard errors for any function of fixed effects are computed using sampling variability to which is added the effects of parameter uncertainty due to random effect estimation. We use Monte Carlo simulations to compute the latter by sampling 1000 times in the asymptotic distribution of random effects estimates.

We find that for individuals observed over a small number of periods (less than 22) the estimates are severely biased upwards and this affects the fixed effect estimates for the complete sample when compared to the consistent random effect estimates. For the two remaining group-

ing of observed periods ( (22,26] and (27,28]), random effect estimates lie between or close to these two fixed-effect estimates. This might be due not only to a remaining  $1/N_i$  bias but also to different underlying stochastics which characterize these two sub-populations. Random effect estimates would reflect the mixture of these two groups.

This interpretation finds some confirmation when plotting the profile of variances of earnings along the life-cycle in Figure 5. This sets more clearly the question whether these fixed effect estimates are able to reproduce the pattern of earnings variances over time. In both panels of this figure, we graph the life-cycle profile of variances due to the factor part of the model only (i.e. the permanent effects due to factors and factor loadings  $V(M(\beta)\eta_i^c)$ ) in which matrix  $M(\beta)$  is composed by a constant, a trend and the geometric rate  $\beta^{-t}$ , as explained in Section 5. Stochastic earnings,  $v_i^c$ , are fixed and their passive role obscures these comparisons so that we prefer not to include them in this comparison. Figure 5 graphs the prediction of the variance profiles that can be computed using random or fixed effect estimates. We use the subsample in which the number of observed periods is larger than 22 because Table 6 shows that the bias is much less severe for such observations. Earnings profiles using fixed effect estimates reproduce the random effect variance profile at a higher level during the first years of working life in Figure 5. Discrepancies with random effect estimates seem nevertheless second order and this validates the use of this selected sample in the counterfactual experiment below.

Finally, the comparison between random effect and fixed effect estimates implicitly relies on an homogeneity assumption of the residuals,  $\hat{v}_i^c$  as a function of  $\eta_i$ . When plotting the variance profiles of these residuals in groups defined either by human capital or by the length of the observation periods, we find very little differences between those groups (see Figure A.iii in the supplementary Appendix). The three factor structure seems to be sufficient to describe the individual permanent heterogeneity in our data and this partly justifies ex-post the homogeneity assumption on the covariance matrix of transitory terms in the random effect specification.

## 6.4 Structural restrictions

With these estimates in hand, we can directly evaluate the relevance of economic restrictions. We have three restrictions, the coefficient of the linear trend should be non negative ( $\eta_2 \geq 0$ ), the coefficient of the geometric factor should be non positive ( $\eta_3 \leq 0$ ) and a weighted sum of these two coefficients should be non negative ( $\eta_3 + \pi_T \eta_2 \geq 0$ ). Parameter  $\pi_T > 0$  is fixed in the population and a function of  $\beta$  (see Section 5.4).

Figure 6 represents those restrictions in an informal way. The cloud of estimates of  $\eta_2$  and

$\eta_3$  is scattered around a downward sloping line and this reflects the strong negative correlation between these two factor loadings that was found using random effect estimates and confirmed by average effect estimates. From an econometric perspective, this is attributable with no doubt to the very different asymptotic behaviour of the two factors, one being a linear trend and the other being geometric. In economic terms, it means that the larger the slope of the profile, the larger the curvature. Second, points in orange (or light) referring to individuals for which the number of observed periods are few (less than 20) are more scattered than the blue (or dark) points which refer to more continuously observed individuals. Finally, constraints are represented by the triangle in red (or dark). This Figure makes clear that checking these constraints is very sensitive to two key elements. The first one is the relative position of the origin point (0,0) with respect to the earnings group averages described in Section 6.2.<sup>20</sup> This is the object of robustness checks below. Second, the  $\pi_T$  parameter which determines the slope of the bottom-left side of the triangle.

More formally, Tables 7 and 8 report frequencies of restriction violations using previous estimates and the same presentation regarding the number of observed periods. In Table 7, we report the sample frequency of rejections at level 5% of each of the three single restrictions using a standard one sided t-test. Should restrictions hold true in the sample, we would expect the rejection frequency to be around 5%, i.e. the level of the test. This frequency tends to decrease with the number of observed periods (at least in the first two columns) and this may be partly due to the quality of the normal asymptotic approximation that we use for testing.

Concentrating on the two groups for which the number of observed periods belongs to (22,26] and (26,28], we see that the restriction that the random growth parameter,  $\eta_{i2}$ , is non negative is plausible.

The second restriction that the the curvature parameter,  $\eta_{i3}$ , is non positive is less acceptable since the rejection rate rises up to 10% in the group (26,28]. The last restriction involving both parameters is even less acceptable, with a rejection rate of 18% in the group (26,28]. This restriction is related to the assumption that investments are positive until period  $T$ . This means that some people stop investing before the end of the period of observation and this agrees with hours of formal learning decreasing with age as emphasized by Mincer (1997).

Table 8 reports testing the three restrictions globally. Following what was developed in Sec-

---

<sup>20</sup>These restrictions are only partially fulfilled by the point estimates of group averaged coefficients. If  $\bar{\eta}_{g2}$  is positive for every group, and  $\bar{\eta}_{g3}$  is negative for every group but one, the ratio  $\bar{\eta}_{g3}/\bar{\eta}_{g2}$  should be greater than  $-\pi_T$ , equal here to 4.08 since  $\beta = 0.95$  and  $T = 30$ . This last constraint is verified for only 9 groups that account for 2/3 of the sample.

tion 5.4.1 and further explained in Appendix A.III.3, we use the quasi likelihood ratio (QLR) statistic associated to the global restriction. P-values are obtained by simulation in the distribution of the statistic under the null (Silvapulle and Sen, 2005). Table 8 reports the sampling frequency of rejections using three different levels (0.01, 0.05 and 0.10).<sup>21</sup> Overall, these results are slightly less favourable for the specification that we use and the frequency of rejections is far larger than the level, in particular in the incomplete group ( $N_i \in (22, 26]$ ) and when the level is small (0.01, 0.05). There seems to exist a "tail" of individuals for whom we reject these restrictions and who would roughly account for 10% of the sample.

## 6.5 Structural parameter estimates

To estimate structural parameters, we adopt the simulation strategy under constraints described in subsection 5.4 and more precisely in section A.III.4 of the supplementary appendix. Quantiles of the simulated constrained estimates are presented in Table A.viii of the supplementary appendix which mirrors Table 5 which presents quantiles of the unconstrained estimates. Since restrictions apply to  $\eta_2$  and  $\eta_3$  only, it is their distributions which are primarily affected and quantiles of simulated constrained estimates are much less dispersed.

Table 9 presents the estimated quantiles of the structural parameters  $\kappa_i$ ,  $\rho_i^L$  and  $c_i^L$  based on the simulated constrained estimates of  $\eta_i$ . Recall that the true structural parameters  $\rho_i$  and  $c_i$  are not separately identified and  $\rho_i^L$  and  $c_i^L$  are lower bounds (Proposition 4). Quantiles differ according to the number of observed periods. The larger the number of periods of observation is, the smaller most quantiles of the terminal value of investment  $\kappa_i$ , and the larger most quantiles of the lower rate of return  $\rho_i^L$  and of the lower cost of investment  $c_i^L$ . If the distribution of  $\kappa_i$  seems to be symmetric, this is not the case of  $c_i^L$  and  $\rho_i^L$  which are characterized by thick tails on the right of their distributions. For individuals with the largest number of observation periods, the terminal value of investments range from the 5<sup>th</sup> percentile 1.4 to the 95<sup>th</sup> 13.9. The distribution of  $\kappa_i$  is by construction bounded between 0 and  $1/(1-\beta) = 20$  (Proposition 4). Furthermore, we can invert the relationship  $\kappa_i = 1/(1-\beta_i^{(T)})$  to construct the implicit individual specific discount rates  $\beta_i^{(T)}$  to which those estimates of  $\kappa_i$  correspond. The 5<sup>th</sup> (respectively 95<sup>th</sup>) percentile is 0.28 (resp. 0.93).

The restriction that human capital investments are non negative until the last period of observation is  $\beta\rho_i\kappa_i > 1$  (Proposition 1). This constraint is used to estimate the minimal value

---

<sup>21</sup>Because these computations are numerically intensive, the number of simulations in assessing standard errors due to parameter uncertainty is smaller than previously (see notes to Tables).

of  $\rho_i^L$ , i.e.  $1/(\beta\kappa_i)$ .<sup>22</sup> For the set with the highest number of periods of observation, between 26 and 28, 90% of individuals have  $\rho_i^L$  standing between 0.0866 and 6.83. This range is extremely broad with respect to the consequences it may have on the earnings profiles. To illustrate this, considering the values of the other parameters at their medians, a change of  $\rho$  from 0.3 to 0.4 multiplies by 1.9 the log human capital level accumulated by the individual. Clearly, this large range of values can be compatible with the data only if the other structural parameters are very correlated with each other, which is by construction the case of  $\rho_i^L$  and  $c_i^L$ . What it shows is that a key issue would be the identification of the date at which human capital accumulation stops. This information would allow the separate identification of  $\rho_i$  and  $c_i$ .

Figure 7 reports the graphs of the life-cycle profile of variance earnings using simulated estimates. The concavity pattern is somewhat more pronounced when we use these estimates than when we use the unconstrained fixed effect estimates as in Figure 5. The trough of the profile due to permanent effects (the Mincer "dip") is happening later in the life cycle ( $t = 12$ ) in this Figure with respect to  $t = 9$  using the unconstrained effect specification.

## 6.6 Robustness and other diagnostics

We review here various departures from our baseline estimates too check that our results are robust. We also comment on additional goodness-of-fit diagnostics.

The first and foremost issue is the validity of the identifying restriction (18) that commands the test of structural restrictions and the estimation of structural estimates. The relative positioning of the origin point in Figure 6 is sensitive to this restriction. We first use a simplified "flat spot" approach proposed by Heckman, et al. (1998) and developed by Bowlus and Robinson (2012). Details of the estimation of a single series of human capital prices are presented in the supplementary Appendix A.IV.8. In a nutshell, human capital prices are estimated using a population of older males whose potential experience ranges from 25 to 40 as in Bowlus and Robinson (2012). Those prices are used to deflate real earnings and the different procedures of testing and estimation are replicated. Results are provided in the supplementary Appendix A.IV.8 and change only marginally with respect to the results that were presented above.

More radically, the dynamics of human capital accumulation depends on whether the average earnings or productivity profile is attributed to human capital only or to other factors (physical

---

<sup>22</sup>From this relationship one might expect a one to one mapping between quantiles of  $\kappa_i$  and  $\rho_i^L$ . This is not the case because the relationship is not linear and estimates used in Table 9 are constructed as averages of 20 Monte-Carlo simulations (see Section 5.4.2).

capital for instance). To control for this issue, we also repeated our procedures by deflating real earnings by a series of average labor productivity. It is a more extreme experiment than the one using the flat spot condition above, yet our results remain qualitatively similar.

Another issue is related to the bias in the fixed effect estimates that we reported. Table A.vi in the supplementary Appendix reports the statistics using bias corrected estimates. First, bias-correction at the first order does not seem sufficiently precise to correct the bias for individuals which are observed less than 22 periods. Bias correction works much better for the other observations with a tendency to overcorrect in the group of individuals observed more than 26 periods. Overall, this approach does not qualitatively change our results.

Finally, we also attempted to use our model to assess its predictive power for the non linear dynamics of earnings. It is well known that linear methods as the ones we are using are poor when estimating transition matrices (e.g. Hirano, 2002). We report in Tables A.ix and A.x in the supplementary Appendix two estimated transition matrices across quintiles (i.e. 20% percentile groups) at 15 year intervals. As expected non linear predictions using unconstrained or simulated constrained estimates are not as good as linear predictions, specifically in the first sample period (from  $t = 1$  to  $t = 15$ ). The predictions are good in the middle quintiles but under or over-predict in the bottom and top quintiles. Predictions are much better for the second sample period (from  $t = 16$  to  $t = 31$ ). This is certainly because the transitory components are much less important in that period and because the permanent components or factor loadings are estimated without imposing that they are normally distributed as in the usual earnings dynamics set-up.

## 7 Life cycle inequalities and a counterfactual experiment

Table 10 provides a decomposition of the cross-sectional inequality into permanent and transitory components. The first column reports the variance of logs earnings every five years from 1977 to 2007 and the average variance over the sample period. The second column reports the share of the variance due to permanent factors and the third, the share due to transitory components. Permanent and transitory components are not orthogonal because of initial conditions. To make them orthogonal, initial conditions are first projected onto permanent components and this projection is aggregated to permanent components and the residuals to the transitory ones. We here report results obtained by random effects since results obtained by fixed effects or constrained fixed effects are very similar. This is because fixed and random effect estimates of

the variance profile are similar (e.g. Figure 7).

Firstly, on average, 65% of the variance is due to permanent factors. This magnitude is close to the one found by Huggett, Ventura and Yaron (2011), who find that the initial endowments related to human capital (initial human capital, learning ability and initial wealth) account for 60% of the variance of lifetime earnings. This share displays a sharp increase over the life cycle from 3% at entry on the labor market to 89% thirty years later. During the first years of the working life, the transitory component is the main contribution to the variance of earnings (see Figure 2).

The very low contribution of the permanent component in the early stage of the working life is partly related to our treatment of initial conditions. Initial conditions are negatively correlated to  $\eta_1$  and  $\eta_3$ , which are the effects that play an important role during the first years of the working life (Table A.iv of the supplementary appendix). As a consequence, these negative correlations lower the contribution of the permanent component to the variance of earnings as computed by the procedure above. The variance of earnings computed from fixed effects only without controlling for initial conditions would be equal to 0.06 a ten times increase and the permanent components would now account for 30% of the variance at the beginning of the working life. As both constructions are valid, it is fair to say that the uncertainty of the permanent component contribution is rather large at the beginning of the working life (3 to 30%). Those remarks do not apply at the other end of the profile since the influence of initial conditions becomes negligible.

The counterfactual exercise of prolonging life expectancy is easily implemented. Life expectancy is increased by two years ( $K = 2$ ) and we use the simulated structural estimates as derived in Section 5.5 to compute those counterfactuals. Nonetheless, this counterfactual is only partially identified because the rate of return is only partially identified. We first set the individual specific rate of return to the minimal estimated value and check the robustness of results by using larger and larger rates of return by multiplying them by fixed values. We report results for the minimal value and not the robustness checks that show that these estimates are quite robust to changes in the assumptions about  $\rho$ .<sup>23</sup> In Figure 8, the top panel reports the effect on mean earnings for those individuals who are observed more than 22 periods. Mean earnings increase and the more so the closer we are to the end of the observation period. This change has also an impact on the profile of earnings variance reported in Figure 8, bottom panel. Variances are increasing in particular at the end of the period. Because rates of return are heterogenous, a

---

<sup>23</sup>More rigorously, the true identified set would be obtained by making the parameters controlling partial identification individual specific. Analyzing results in this case is left for future research.

larger life expectancy magnifies individual differences in earnings and this implies more earnings inequalities. In the last period, this increases cross-section inequality by 20% although this figure is imprecise because standard errors are quite large.

The construction of counterfactuals for the human capital technology is more speculative. Since only a lower bound for rates of returns can be identified, experiments for constructing counterfactuals leads to very large bounds. It is thus fair to say that those specific counterfactuals are not identified. As mentioned earlier, one possible route would be to use parametric assumptions for structural parameters in order to identify rates of return and consequently counterfactuals involving those rates.

## 8 Conclusion

In this paper, we proposed a structural model of human capital investments that predicts that a linear factor model describes well the earnings profile over the working life and that its unobserved individual factor loadings have an economic interpretation. Using a long panel on a single cohort of private sector wage earners in France from 1977 to 2007, we used random effect and fixed effect methods to estimate factor loadings and assess the relevance, the bias and the accuracy of these estimates. This procedure enabled us to evaluate the relevance of structural restrictions and to construct estimates of structural individual components in the original model which are returns, costs and terminal values of human capital stocks. This let us compute richer counterfactuals than the ones that are directly available using random effect procedures. We also showed the importance of dealing properly with initial conditions at entry in the labor market.

Random effect estimation delivers empirical results which are close to what has been obtained in the literature and are easily interpretable in a human capital framework. Fixed effect estimation evinces that structural restrictions are not rejected for most of our sample observations. It remains to be seen if this is because of the low power of our testing procedure as in Baker (1997) in which heterogenous growth and RIP models are hard to discriminate. Regarding empirical results, a simple counterfactual analysis showed that increasing life expectancy has quite a large effect on earnings inequality even if this result is obtained in a partial analysis in which initial human capital investments are held constant. It seems dubious to us that making those initial investments vary as well in the counterfactual scenario would overturn our conclusion that inequality increases with increasing life expectancy, as individual specific rates of returns to schooling and post-schooling are strongly correlated.



There are many extensions worth exploring that we are leaving for future research. First, human capital investment profiles vary across different education groups. In particular, a pending conjecture would be that investments by the low skill group stop much earlier than those by the high skill group. Second, this model is versatile enough to accommodate consumption smoothing in the case in which consumption and income data are available. Estimating the general model under these conditions might help further understanding the role that human capital investments play in welfare inequalities. Third, goodness-of-fit measures seem to point out that the missing at random assumption might be invalid. Analyzing this condition using complete and incomplete samples might lead to a better correction of selection and small sample biases although this is a project on its own. Another theoretical issue in econometric modeling is the analysis of a mixture of the heterogeneous growth and restricted income processes in the specification of the earnings equation and specifically involve looking at identification issues.

## REFERENCES

- Aakvik, A. J.J., Heckman and E., Vytlacil**, 2005, "Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs", *Journal of Econometrics*, 125:15–51.
- Abowd, J. and Card, D.**, 1989, "On the covariance structure of earnings and hours changes", *Econometrica* 57(2): 411-445.
- Alan, S. and M. Browning**, 2010. "Estimating Intertemporal Allocation Parameters using Synthetic Residual Estimation", *Review of Economic Studies*, 77(4):1231-1261.
- Alvarez, J. and Arellano, M.**, 2004, "Robust likelihood estimation of dynamic panel data models", Working Paper 0421.
- Andersen, S. G. Harrison, M. Lau, and E. Rutstrom**, 2008, "Eliciting Risk and Time Preferences", *Econometrica*, 76(3), 583-618.
- Arellano, M.**, 2004, *Panel data*, Cambridge UP: Cambridge.
- Arellano M., and S., Bonhomme**, 2012, "Identifying Distributional Characteristics in Random Coefficient Panel Data Models", *The Review of Economic Studies*, 79(3): 987-1020.
- Atkinson, A.B.**, 2008, *The changing distribution of earnings in OECD countries*. Oxford University Press.
- Atkinson, A.B. and S. Morelli**, 2012, "Chartbook of Economic Inequality: 25 Countries 1911-2010". *INET Research Note* n 015.
- Atkinson, A.B., T. Piketty and E., Saez**, 2011, "Top Incomes in the Long Run of History", *Journal of Economic Literature*, 49(1), 3-71.
- Autor, D. Katz, L. and M. Kearney**, 2008, "Trends in the US wage inequality: Revisiting the revisionists", *Review of Economics and Statistics*, 90(2), 300-323.
- Bai, J.**, 2009, "Likelihood Approach to Small  $T$  Dynamic Panel Models with Interactive Effects", *Econometrica*, 77(4), 1229-1279.
- Baker, M.**, 1997, "Growth-rate heterogeneity and the covariance structure of life-cycle earnings", *Journal of Labor Economics* 15(2):338-375.
- Ben-Porath, Y.**, 1967, "The production of human capital and the life cycle of earnings", *Journal of Political Economy* 75(4, Part.1):352-365.
- Blundell, R.**, 2014, "Income Dynamics and Life-cycle inequality: Mechanisms and Controversies", *The Economic Journal*, 124:289-318.
- Blundell, R., & Etheridge, B.**, 2010, "Consumption, income and earnings inequality in Britain", *Review of Economic Dynamics*, 13(1):76-102.
- Bonhomme, S., and J.M., Robin**, 2009, "Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990-2000", *Review of Economic Studies*, 76(1), 63-92.
- Bonhomme, S., and J.M., Robin**, 2010, "Generalized Nonparametric Deconvolution with an Application to Earnings Dynamics", *Review of Economic Studies*, 77(2), 491-533.
- Bowlus, A. and C. Robinson**, 2012, "Human Capital Prices, Productivity, and Growth", *American Economic Review*, 102(7), 3483-3515.
- Browning, M., Ejrnaes, M. and Alvarez, J.**, 2012, "Modelling income processes with lots of heterogeneity", *Review of Economic Studies*, 77(4):1353-1381
- Browning, M., Hansen, L. P. and Heckman, J. J.**, 1999, "Micro data and general equilibrium models", in J.B. Taylor and M.Woodford (eds), *Handbook of Macroeconomics*, Elsevier Science, Amsterdam.
- Card, D.**, 2001, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, 69(5):1127-60.

- Caroll, C.D., and Lawrence H. Summers**, 1991, "Consumption Growth Parallels Income Growth: Some New Evidence", in B. Douglas Bernheim and John B. Shoven, eds., *National Saving and Economic Performance*, Chicago: Chicago University Press for NBER, 305-43.
- Carneiro, P., K. Hansen and J.J. Heckman**, 2003, "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice", *International Economic Review*, 44(2): 361-422.
- Charnoz, P. Coudin, E. and Gaini, M.**, 2011, "Wage inequalities in France 1976-2004: a quantile regression analysis", WP DESE n°G2011/06.
- Charnoz, P. Coudin, E. and Gaini, M.**, 2014, "Forty Years of Decreasing Wage Inequality in France: The Roles of Supply and of Hidden Skill-Biased Technical Change", WP CREST n° 2014-20.
- Cunha, F., J.J., Heckman and S. Navarro**, 2005, "Separating uncertainty from heterogeneity in life cycle earnings", *Oxford Economic Papers*, 57(2):191-261.
- Cunha, F., J.J. Heckman and S.M., Schennach**, 2010, "Estimating the Technology of Cognitive and Noncognitive Skill Formation", *Econometrica*, 78(3):883-931.
- Daly, M., D., Hryshko and I.Manovskii**, 2014, "Reconciling Estimates of Earnings Processes in Growth Rates and Levels", unpublished manuscript
- Friedman, M. and Kuznets, S.**, 1954, *Income from Independent Professional Practices*. National Bureau of Economic Research, Inc.
- Geweke, J. and Keane, M.**, 2000, "An empirical analysis of earnings dynamics among men in the PSID: 1968-1989", *Journal of Econometrics*, 96: 293-356.
- Guvenen, F.**, 2007, "Learning your earnings: Are labor income shocks really very persistent?", *American Economic Review* 97(3): 687-712.
- Guvenen, F.**, 2009, "An empirical investigation of labor income processes", *Review of Economic Dynamics*, 12:58-79.
- Guvenen, F. and B., Kuruscu**, 2012, "Understanding the Evolution of the US Wage Distribution: A Theoretical Analysis", *Journal of the European Economic Association*, 10(3):482-517.
- Haley, W.**, 1976, "Estimation of the Earnings Profile from Optimal Human Capital Accumulation", *Econometrica*, 44(6):1223-1238.
- Hahn, J., and G., Kuersteiner**, 2002, "Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T Are Large", *Econometrica*, 70(4), 1639-1657.
- Hause, J.**, 1980, "The fine structure of earnings and the on-the-job training hypothesis", *Econometrica*, 48(4):1013-1029.
- Heckman, J.J., L. Lochner and C. Taber**, 1998, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics*, 1, 1 - 58.
- Heckman, J. J., Lochner, L. L. and Todd, P. E.**, 2006, "Earnings functions, rates of return and treatment effects: the mincer equation and beyond", in E. A. Hanushek and F. Welch (eds), *Handbook of Economics of Education*, Elsevier Science, Amsterdam.
- Hirano, K.**, 2002, "Semiparametric bayesian inference in autoregressive panel data models", *Econometrica* 70(2): 35-45.
- Hofmann, F.**, 2013, "HIP, RIP and the robustness of empirical earnings processes", *Unpublished manuscript*.
- Hospido, L.**, 2012, "Modelling heterogeneity and dynamics in the volatility in individual wages", *Journal of Applied Econometrics*, 27(3):386-414.
- Hryshko, D.**, 2012, "Labor income profiles are not heterogeneous: Evidence from income growth rates", *Quantitative Economics*, 3(2)177-210.
- Hsiao, C., M. H. Pesaran, and A. K. Tahmiscioglu**, 2002, "Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods", *Journal*

of *Econometrics*, 109, 107–150.

**Huggett, M., G., Ventura and A., Yaron**, 2011, “Sources of Lifetime Inequality”, *American Economic Review*, 101:2923-2954.

**Le Minez, S. and Roux, S.**, 2002, “Les différences de carrières salariales à partir du premier emploi”, *Economie et Statistique* 351:31-63.

**Lillard, L. A. and Reville, R. T.**, 1999, “Life cycle human capital investment: New evidence on and old issue”, unpublished manuscript.

**Lillard, L. A. and Willis, R. J.**, 1978, “Dynamic aspects of earnings mobility”, *Econometrica* 46(5):985-1012.

**Lochner, L., and Y. Shin**, 2013, "Understanding Income Dynamics: Identifying and Estimating the Changing Roles of Unobserved Ability, Permanent and Transitory Shocks", unpublished manuscript.

**MaCurdy, T.**, 1982, “The use of time series processes to model the error structure of earnings in a longitudinal data analysis”, *Journal of Econometrics* 18: 83-114.

**Magnac, T. and S., Roux**, 2009, “Dynamique des salaires dans une cohorte”, *Economie et Prévision*, 187:1-24.

**Magnac, T., N., Pistolesi and S., Roux**, 2013, “Post Schooling human capital investments and the life cycle variance of earnings”, Toulouse School of Economics WP n ° 13-380.

**Meghir, C. and Pistaferri, L.**, 2004, “Income variance dynamics and heterogeneity”, *Econometrica*, 72:1-32.

**Meghir, C. and Pistaferri, L.**, 2010, “Earnings, consumption and lifecycle choices”, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 4b, Elsevier Science, Amsterdam, pp. 773-854.

**Mincer, J.**, 1974, *Schooling experience and earnings*, National Bureau of Economic Research, Inc.

**Mincer, J.**, 1997, “The Production of Human Capital and the Life Cycle of Earnings: Variations on a Theme”, *Journal of Labor Economics*, 15(1): S26-47

**Moffitt, R. A. and Gottschalk, P.**, 2011, "Trends in the Transitory Variance of Male Earnings in the U.S., 1970-2004," NBER Working Papers 16833, National Bureau of Economic Research, Inc.

**Okui, R.**, 2009, “The Optimal Choice of Moments in Dynamic Panel Data Models”, *Journal of Econometrics*, 151:1-16.

**Polachek, S. T. Das and R. Thamma-Apiroam**, 2013, “Heterogeneity in the Production of Human Capital”, IZA Discussion Paper 7335.

**Postel-Vinay, F. and Turon, H.**, 2010, “On the job search, productivity shocks and the individual earnings process”, *International Economic Review* 51(3):599-629.

**Rubinstein, Y. and Weiss, Y.**, 2006, “Post schooling wage growth: investment, search and learning”, in E. Hanushek and F. Welch (eds), *Handbook of the Economics of Education*, Elsevier, Amsterdam, pp. 1-67.

**Silvapulle, M.J. and P.K., Sen**, 2005, *Constrained Statistical Inference*, Wiley, Oxford.

**Stokey, N.L. and R.E., Lucas**, 1989, *Recursive methods in Economic Dynamics*, Harvard UP, Harvard.

**Verdugo, G.**, 2014, “The great compression of the wage structure in France 1969-2008”, *Labour Economics* 28:131–144.

# APPENDICES

## A Proofs of Propositions and Extensions

### A.1 Proof of Proposition 1

The first order condition of the maximization problem for  $t < T + 1$  is

$$-(1 + c_i \tau_i(t)) + \beta_i \rho_i H_i(t + 1) E_t \left[ \frac{\partial W_{t+1}}{\partial H_i(t + 1)} \right] = 0. \quad (\text{A.1})$$

The marginal value of human capital is the derivative of the Bellman equation so that by the envelope theorem:

$$\frac{\partial W_t}{\partial H_i(t)} = \frac{1}{H_i(t)} + \beta_i E_t \left[ \frac{\partial W_{t+1}}{\partial H_i(t + 1)} \right] \frac{H_i(t + 1)}{H_i(t)} \quad (\text{A.2})$$

For  $t = T + 1$ , condition (A.2) writes more simply as:

$$\frac{\partial W_{T+1}}{\partial H_i(T + 1)} = \frac{\kappa_i}{H_i(T + 1)} \implies H_i(T + 1) \frac{\partial W_{T+1}}{\partial H_i(T + 1)} = \kappa_i,$$

so that, by backward induction, we obtain:

$$H_i(T) \frac{\partial W_T}{\partial H_i(T)} = 1 + \beta_i \kappa_i, \quad H_i(T - 1) \frac{\partial W_{T-1}}{\partial H_i(T - 1)} = 1 + \beta_i (1 + \beta_i \kappa_i)$$

and so on. This yields:

$$H_i(t + 1) \frac{\partial W_{t+1}}{\partial H_i(t + 1)} = \frac{1 - \beta_i^{T-t}}{1 - \beta_i} + \beta_i^{T-t} \kappa_i.$$

Replacing in equation (A.1) yields:

$$(1 + c_i \tau_i(t)) = \beta_i \rho_i \left[ \frac{1}{1 - \beta_i} + \beta_i^{T-t} \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right] = \rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i^{T+1-t} \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right],$$

and equation (7) follows. Furthermore, as the second term in (A.1) is constant, the second order condition is satisfied if and only if  $c_i > 0$ .

Furthermore and given that  $c_i > 0$ , the condition that investments are always positive yields:

$$\rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i^{T+1-t} \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right] - 1 \geq 0. \quad \forall t < T + 1$$

As  $\kappa_i - \frac{1}{1 - \beta_i} < 0$  and  $\beta_i < 1$ ,  $\tau_i(t)$  is decreasing in  $t$  because of the term  $\beta_i^{-t}$  and the RHS attains its minimum at  $t = T$ . This yields condition (6) since:

$$\rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right] - 1 \geq 0 \iff \rho_i \geq \frac{1}{\beta_i \kappa_i}.$$

## A.2 Proof of Proposition 2

First, condition (8) is consistent since  $\kappa_{it} = 1 + \beta_i \kappa_{i,t+1} > \kappa_{i,t+1} \Leftrightarrow \kappa_{i,t+1} < \frac{1}{1-\beta_i} \Leftrightarrow \kappa_{i,t+2} < \frac{1}{1-\beta_i}$  and by repetition  $\kappa_{i,T+1-1} = \kappa_i < \frac{1}{1-\beta_i}$ .

We proceed by backward induction. By Proposition 1, we know that

$$\tau_i(T) > 0 \iff \frac{1}{\kappa_{i,T+1-1}} < \beta_i \rho_i \leq \frac{1}{\kappa_{i,T+1}},$$

and under this latter condition, that equation (7) is satisfied for all  $t + 1 \leq T$ .

Assume that for some  $t + 1 \leq T$ :

$$\forall t' \geq t + 2, t' < T + 1, \tau_i(t') = 0, \text{ and } \tau_i(t + 1) > 0 \iff \frac{1}{\kappa_{i,t+1}} < \beta_i \rho_i \leq \frac{1}{\kappa_{i,t+2}} \quad (\text{A.3})$$

and under this latter condition, that equation (7) is satisfied for all  $t' \leq t + 1$ . In a proof of Proposition 2 by backward induction, we thus shall prove that condition (A.3) is true at period  $t$ .

We analyze separately the condition  $\tau_i(t') = 0, \forall t' \geq t + 1$  and the condition  $\tau_i(t) > 0$ .

Assume first that  $\tau_i(t') = 0, \forall t' \geq t + 1$  so that the condition  $\tau_i(t') > 0$  is violated for any  $t' \geq t + 1$  and therefore by equation (A.3),  $\beta_i \rho_i \leq 1/\kappa_{i,t+1}$ . Conversely, if  $\beta_i \rho_i \leq 1/\kappa_{i,t+1}$  then  $\tau_i(t') = 0, \forall t' \geq t + 1$  because equation (A.3) is satisfied for  $t' \geq t + 1$ . Furthermore, conditions  $\tau_i(t') = 0$  implies simple forms for the Bellman equation (3):

$$W_t(H_i(t')) = \delta_i(t') + \log H_i(t') + \beta E_{t'} W_{t'+1}(H_i(t' + 1)),$$

and the accumulation equation (1):

$$\log H_i(t' + 1) = \log H_i(t') - \lambda_i(t').$$

Using equation (4) where we set  $\kappa_{i,T+1-1} = \kappa_i$  and the linearity of the previous two equations lead to the condition derived by induction again:

$$W_{t'}(H_i(t')) = \delta^*(t') + \kappa_{i,t'-1} \log H_i(t'). \quad (\text{A.4})$$

for any  $t' \geq t + 1$  and where  $\kappa_{it} = 1 + \beta_i \kappa_{i,t+1}$ .

Second, assume that  $\tau_i(t) > 0$ . Proposition 1 can be recast in a set-up where the last period becomes  $T_i = t + 1$  instead of  $T + 1$  since there are no further human capital investments after this date and since the value function can be written as in equation (A.4) evaluated at  $t' = t + 1$ . We rewrite equation (7) and obtain:

$$\tau_i(t) = \frac{1}{c_i} \left\{ \rho_i \left[ \frac{\beta_i}{1-\beta_i} + \beta_i \left( \kappa_{it} - \frac{1}{1-\beta_i} \right) \right] - 1 \right\} > 0, \quad (\text{A.5})$$

which is equivalent to  $\beta_i \rho_i > \frac{1}{\kappa_{it}}$ .

Therefore the equivalence stated in the Proposition is true at period  $t$ . Furthermore equation (7) applies for any  $t' \leq t$ . The statement under induction is therefore true at any date  $t \in \{0, \dots, T\}$ . By convention we set  $\frac{1}{\kappa_{i0}} = 0$  in order to cover all cases since  $\rho_i > 0$ . ■

### A.3 The Model with Consumption Smoothing

**Proposition A.1** *Under conditions stated in the proof and denoting  $s_i(t) = \frac{y_i(t) - C_i(t)}{y_i(t)}$ , the saving rate, we have:*

$$\tau_i(t) = \frac{1}{c_i} \left\{ \rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i^{T+1-t} \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right] - \frac{1}{1 - s_i(t)} \right\}.$$

**Proof.** Assume that consumption can be smoothed over time. The new dynamic program is written as:

$$\max_{C_i(t), \tau_i(t)} [\log(C_i(t)) - c\tau_i(t)^2/2 + \beta_i E_t W_{t+1}(A_i(t+1), H_i(t+1))]$$

under the constraints:

$$\begin{aligned} A_i(t+1) &= (1 + r_i(t))A_i(t) + y_i(t) - C_i(t), \\ y_i(t) &= \exp(\delta_i(t)) H_i(t) \exp(-\tau_i(t)), \\ H_i(t+1) &= H_i(t) \exp(\rho_i \tau_i(t) - \lambda_i(t)). \end{aligned}$$

The first order conditions write:

$$\begin{aligned} \frac{1}{C_i(t)} - \beta_i E_t \frac{\partial W_{t+1}(A_i(t+1), H_i(t+1))}{\partial A_i(t+1)} &= 0, \\ -c_i \tau_i(t) + \beta_i E_t \frac{\partial W_{t+1}(A_i(t+1), H_i(t+1))}{\partial A_i(t+1)} [-y_i(t)] \\ + \beta_i E_t \frac{\partial W_{t+1}(A_i(t+1), H_i(t+1))}{\partial H_i(t+1)} \rho H_i(t+1) &= 0, \end{aligned}$$

in which the second term comes from

$$\frac{\partial A_i(t+1)}{\partial \tau_i(t)} = \frac{\partial A_i(t+1)}{\partial y_i(t)} \frac{\partial y_i(t)}{\partial \tau_i(t)} = -y_i(t).$$

Replacing the first in the second first order conditions yields:

$$c_i \tau_i(t) + \frac{y_i(t)}{C_i(t)} = \beta E_t \frac{\partial W_{t+1}(A_i(t+1), H_i(t+1))}{\partial H_i(t+1)} \rho H_i(t+1).$$

If  $y_i(t) = C_i(t)$ , this is condition (A.1). Note that  $\frac{y_i(t)}{C_i(t)} = \frac{1}{1 - s_i(t)}$  in which  $s_i(t)$  is the savings rate.

We thus have to replace equation (A.5) in the above Section A.2 by:

$$\tau_i(t) = \frac{1}{c_i} \left\{ \rho_i \left[ \frac{\beta_i}{1 - \beta_i} + \beta_i \left( \kappa_i - \frac{1}{1 - \beta_i} \right) \right] - \frac{1}{1 - s_i(t)} \right\} > 0.$$

■

## A.4 Proof of Proposition 3

First, the stock of human capital in period  $t$  depends on previous investment choices and past depreciation that is

$$H_i(t) = H_i(1) \exp \left[ \sum_{l=1}^{t-1} \rho_i \tau_i(l) - \sum_{l=1}^{t-1} \lambda_i(l) \right] \text{ for } 2 \leq t.$$

We can write the logarithm of observed earnings in period  $t$  as

$$\ln y_i(t) = \delta_i(t) + \ln H_i(1) + \sum_{l=1}^{t-1} \rho_i \tau_i(l) - \sum_{l=1}^{t-1} \lambda_i(l) - \tau_i(t). \quad (\text{A.6})$$

If  $t \leq T_i$ , insert the structural expression for  $\tau_i(\cdot)$  given by the equation (9) of proposition 2 into the first sum of equation (A.6) to get:

$$\begin{aligned} \sum_{l=1}^{t-1} \rho_i \tau_i(l) &= \frac{\rho_i^2}{c_i} \sum_{l=1}^{t-1} \left[ \frac{\beta_i}{1-\beta_i} + \beta_i^{T_i+1-l} \left( \kappa_i - \frac{1}{1-\beta_i} \right) \right] - \frac{\rho_i}{c_i} (t-1), \\ &= \frac{\rho_i^2}{c_i} \frac{\beta_i}{1-\beta_i} (t-1) + \frac{\rho_i^2}{c_i} \left( \kappa_i - \frac{1}{1-\beta_i} \right) \beta_i^{T_i} \sum_{l=1}^{t-1} \beta_i^{1-l} - \frac{\rho_i}{c_i} (t-1) \\ &= \left( \frac{\rho_i^2}{c_i} \frac{\beta_i}{1-\beta_i} - \frac{\rho_i}{c_i} \right) (t-1) + \frac{\rho_i^2}{c_i} \left( \kappa_i - \frac{1}{1-\beta_i} \right) \beta_i^{T_i} \frac{1 - (1/\beta_i)^{t-1}}{1 - 1/\beta_i} \\ &= -\frac{\rho_i^2}{c_i} \left( \kappa_i - \frac{1}{1-\beta_i} \right) \frac{\beta_i^{T_i+1}}{1-\beta_i} + \left( \frac{\rho_i^2}{c_i} \frac{\beta_i}{1-\beta_i} - \frac{\rho_i}{c_i} \right) (t-1) \\ &\quad + \frac{\rho_i^2}{c_i} \left( \kappa_i - \frac{1}{1-\beta_i} \right) \frac{\beta_i^{T_i+2}}{1-\beta_i} \beta_i^{-t}, \end{aligned}$$

which writes as the sum of three factors whereas one factor is in levels, the second one is a linear trend and the last one is a geometric trend.

Using equation (9):

$$\tau_i(t) = \frac{1}{c_i} \left( \rho_i \frac{\beta_i}{1-\beta_i} - 1 \right) + \frac{\rho_i}{c_i} \beta_i^{T_i+1} \left( \kappa_i - \frac{1}{1-\beta_i} \right) \beta_i^{-t}$$

and rearranging expression (A.6), we obtain equation (10).

If  $t > T_i$ , which in our setting can only apply if  $T_i < T$ , the sequence of investment from the period  $T_i$  on have been nil. Hence,

$$\begin{aligned} H_i(t) &= H_i(1) \exp \left( \sum_{l=1}^{T_i} \rho_i \tau_i(l) - \sum_{l=1}^{t-1} \lambda_i(l) \right) \\ &= H_i(1) \exp \left( \left( \sum_{l=1}^{T_i} \rho_i \tau_i(l) - \lambda_i(l) - \sum_{l=T_i+1}^{t-1} \lambda_i(l) \right) \right) \\ &= H_i(T_i + 1) \exp (-\Lambda_i(t) + \Lambda_i(T_i + 1)) \end{aligned}$$



Finally,

$$\begin{aligned}\ln y_i(t) &= \ln H_i(T_i + 1) + \Lambda_i(T_i + 1) + \delta_i(t) - \Lambda_i(t) \\ &= \ln y_i(T_i + 1) + v_{it} - v_{iT_i+1}\end{aligned}$$

which corresponds to equation (11). ■

## A.5 Proof of Proposition 4

The two equations (13) and (14) simplify to:

$$\begin{cases} \eta_{i2} = \frac{\rho_i}{c_i} \left( \rho_i \frac{\beta}{1-\beta} - 1 \right), \\ \eta_{i3} = \frac{\rho_i}{c_i} \beta^{T+1} \left( \kappa_i - \frac{1}{1-\beta} \right) \left( \rho_i \frac{\beta}{1-\beta} - 1 \right). \end{cases} \quad (\text{A.7})$$

Taking the ratio of the second and the first equation yields:

$$\frac{\eta_{i3}}{\eta_{i2}} = \beta^{T+1} \left( \kappa_i - \frac{1}{1-\beta} \right)$$

we derive the restriction from  $\kappa_i \in [0, \frac{1}{1-\beta}]$  that:

$$\frac{\eta_{i3}}{\eta_{i2}} \in \left[ -\frac{\beta^{T+1}}{1-\beta}, 0 \right]. \quad (\text{A.8})$$

Conversely, if this restriction is valid, then  $\kappa_i$  is given by:

$$\kappa_i = \frac{1}{1-\beta} + \beta^{-(T+1)} \frac{\eta_{i3}}{\eta_{i2}} \in \left( 0, \frac{1}{1-\beta} \right).$$

Furthermore, Proposition 1 proved that investments remain positive until period  $T$  (inclusively) if and only if  $\beta \rho_i \kappa_i > 1$ . This yields that :

$$\rho_i > \rho_i^L = \frac{1}{\beta \kappa_i} = \frac{1}{\frac{1}{1-\beta} + \beta^{T+1} \frac{\eta_{i3}}{\eta_{i2}}} > 0,$$

by the above. The first equation of (A.7):

$$\eta_{i2} = \frac{\rho_i}{c_i} \left( \rho_i \frac{\beta}{1-\beta} - 1 \right) = \frac{\rho_i}{c_i \kappa_i} \left( \frac{\rho_i \beta \kappa_i}{1-\beta} - \kappa_i \right),$$

also implies that, given that all parameters are positive that

$$\eta_{i2} > \frac{\rho_i}{c_i \kappa_i} \left( \frac{1}{1-\beta} - \kappa_i \right) > 0.$$

Conversely, assume that  $\eta_{i2} > 0$  and  $\rho_i > \rho_i^L$ . By construction, the condition  $\beta \rho_i \kappa_i > 1$  is satisfied and investments are positive until  $T$ . Second, define

$$c_i = \frac{\rho_i}{\eta_{2i}} \left( \rho_i \frac{\beta}{1-\beta} - 1 \right),$$

and write

$$\frac{\partial c_i}{\partial \rho_i} = \frac{1}{\eta_{2i}} \left( 2\rho_i \frac{\beta}{1-\beta} - 1 \right)$$

which is positive since  $\rho_i \frac{\beta}{1-\beta} > 1$  because  $\beta\rho_i\kappa_i > 1$  and  $\kappa_i \leq \frac{1}{1-\beta}$ . Both expressions prove that

$$c(\rho_i, \eta_{2i}) = \frac{\rho_i}{\eta_{2i}} \left( \rho_i \frac{\beta}{1-\beta} - 1 \right)$$

is positive and increasing in  $\rho_i$ . Therefore  $c_i \geq c_L = c(\rho_L, \eta_{2i})$ . ■

## B Random effect specification

Redefining the time index accordingly, we shall assume that initial conditions of the process  $(u_{i(1-p)}, \dots, u_{i0})$  are observed. The dynamic process is thus a function of the random variables  $z_i = (v_{i(1-p)}, \dots, v_{i0}, \zeta_{i(1-q)}, \dots, \zeta_{iT})$  which collect initial conditions of the autoregressive process  $(v_{i(1-p)}, \dots, v_{i0})$ , initial conditions of the moving average process  $(\zeta_{i(1-q)}, \dots, \zeta_{i0})$  and the idiosyncratic shocks affecting random shocks between 1 and  $T$ . We write the quasi-likelihood of the sample using a multivariate normal distribution

$$z_i \rightsquigarrow N(0, \Omega_z)$$

We define  $v_{it}$  as

$$v_{it} = \alpha_1 v_{i(t-1)} + \dots + \alpha_p v_{i(t-p)} + \sigma_t w_{it},$$

where  $w_{it}$  is  $MA(q)$ :

$$w_{it} = \zeta_{it} - \psi_1 \zeta_{it-1} - \dots - \psi_q \zeta_{it-q}.$$

The construction of the structure of  $\Omega_z$  is detailed in the Supplementary Appendix A.II (Magnac et al, 2014) although it can be summarized easily. The correlations between initial conditions and individual effects are not constrained, while innovations  $\zeta_{it}$  are assumed orthogonal to any previous terms including initial conditions. However, the initial conditions  $(v_{i(1-p)}, \dots, v_{i0})$  can be correlated with previous shocks as  $\zeta_{i0}, \dots, \zeta_{i(1-q)}$ .

As for the individual effects  $(\eta_{i1}^c, \eta_{i2}^c, \eta_{i3}^c)$  we assume that they are independent of the idiosyncratic shocks  $\zeta_{i(1-q)}, \dots, \zeta_{iT}$  while they can be correlated with the initial conditions of the autoregressive process  $(v_{i(1-p)}, \dots, v_{i0})$  in an unrestricted way. From these restrictions it is possible to build the covariance matrix of the observed variables

$$Vu_i = (u_{i(1-p)}, \dots, u_{i0}, u_{i1}, \dots, u_{iT}) \equiv \Omega_u.$$

This covariance matrix,  $\Omega_u$ , is a function of the parameters of the model that are the autoregressive parameters  $\{\alpha_k\}_{k=1, \dots, p}$ , the moving average parameters  $\{\psi_k\}_{k=1, \dots, q}$ , the covariance matrix (conditional on groups) of  $\eta^c$ ,  $\Sigma_\eta$ , the heteroskedastic components  $\{\sigma_t\}_{t=1, \dots, T}$  and the covariance of fixed effects and initial conditions,  $\Gamma_{0\eta}$  (see Supplementary Appendix A.II).

A pseudo likelihood interpretation can always be given to this specification. As in Alvarez and Arellano (2004), the estimates remain consistent under the much weaker assumption that:

$$E(\zeta_{it} \mid \eta_i, u_i^{t-1}) = 0,$$

although optimality properties of such an estimation method are derived under the normality assumptions only.

Table 1: Sample size

	Age of Entry in 1977			All
	Below 20	Between 20 and 23	Above 23	
1977	4460	2112	874	7446
1978	4460	2112	874	7446
1979	3855	1923	787	6565
1980	3748	1930	785	6463
1982	4460	2112	874	7446
1984	4460	2112	874	7446
1985	3792	1808	724	6324
1986	3683	1800	726	6209
1987	3569	1741	678	5988
1988	3402	1654	637	5693
1989	3486	1657	644	5787
1991	3319	1598	613	5530
1992	3299	1581	603	5483
1993	3330	1620	627	5577
1994	2508	1316	503	4327
1995	3256	1566	578	5400
1996	3236	1557	579	5372
1997	3202	1529	556	5287
1998	3208	1521	543	5272
1999	3218	1503	547	5268
2000	3180	1506	536	5222
2001	3117	1480	517	5114
2002	3018	1463	511	4992
2003	2800	1323	467	4590
2004	2844	1387	463	4694
2005	2851	1399	467	4717
2006	2896	1382	442	4720
2007	2864	1377	429	4670

Table 2: Autocorrelation matrix of earnings residuals

1977	.438																											
1978	.280	.424																										
1979	.241	.367	.563																									
1980	.211	.343	.478	.539																								
1982	.223	.326	.439	.499	.733																							
1984	.221	.306	.401	.411	.665	.814																						
1985	.216	.301	.368	.430	.643	.785	.807																					
1986	.161	.266	.386	.441	.634	.767	.772	.853																				
1987	.156	.260	.401	.459	.634	.756	.744	.809	.871																			
1988	.134	.254	.368	.421	.617	.733	.730	.776	.830	.874																		
1989	.135	.239	.321	.383	.557	.682	.681	.726	.790	.824	.857																	
1991	.145	.221	.334	.370	.577	.685	.679	.721	.765	.798	.821	.887																
1992	.134	.193	.306	.333	.515	.619	.619	.667	.724	.738	.762	.831	.854															
1993	.111	.179	.274	.314	.482	.607	.606	.644	.695	.709	.723	.810	.803	.823														
1994	.102	.183	.280	.330	.480	.590	.580	.632	.696	.711	.735	.809	.815	.810	.792													
1995	.109	.197	.289	.319	.491	.589	.582	.624	.686	.711	.746	.802	.815	.804	.795	.836												
1996	.128	.192	.305	.315	.497	.623	.623	.653	.720	.741	.764	.826	.839	.827	.816	.854	.878											
1997	.129	.198	.308	.336	.507	.625	.614	.656	.716	.737	.761	.828	.842	.833	.816	.862	.883	.932										
1998	.108	.194	.294	.316	.496	.618	.610	.651	.707	.735	.756	.819	.835	.813	.797	.838	.859	.904	.939									
1999	.117	.160	.294	.291	.478	.600	.594	.638	.689	.714	.730	.791	.815	.799	.784	.812	.837	.881	.908	.904								
2000	.124	.179	.293	.310	.501	.619	.613	.635	.696	.715	.741	.808	.822	.802	.795	.820	.830	.885	.919	.913	.908							
2001	.122	.180	.294	.296	.463	.588	.591	.616	.656	.685	.707	.776	.787	.767	.751	.779	.798	.855	.884	.880	.874	.912						
2002	.122	.179	.257	.261	.415	.543	.558	.568	.577	.605	.622	.695	.720	.694	.697	.716	.720	.785	.810	.811	.811	.844	.875					
2003	.128	.168	.291	.299	.469	.589	.585	.616	.669	.697	.715	.780	.794	.770	.763	.787	.799	.858	.887	.883	.877	.916	.914	.862				
2004	.108	.170	.289	.296	.462	.593	.584	.610	.666	.691	.707	.773	.784	.763	.757	.781	.792	.849	.876	.877	.873	.905	.903	.854	.950			
2005	.103	.155	.291	.287	.470	.595	.587	.619	.671	.698	.709	.776	.794	.771	.770	.790	.800	.853	.878	.878	.875	.903	.901	.857	.942	.957		
2006	.106	.157	.286	.279	.449	.572	.558	.591	.638	.670	.677	.738	.754	.745	.732	.757	.770	.819	.840	.845	.841	.872	.874	.828	.909	.931	.952	

Note: In each cell, the correlation is computed using the individuals who are in the data both relevant years. Table A.i of the supplementary appendix presents the number of contributing individuals in each cell.

Table 3: Estimated standard errors and correlations of individual effects  $\eta_i^c$ : Random effect estimation

	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
$\sigma_{\eta_1}$	.302 (.001)	.302 (.003)	.301 (.003)	.310 (.003)	.306 (.003)	.304 (.003)	.306 (.003)	.300 (.003)	.298 (.004)
$\sigma_{\eta_2}$	.038 (.005)	.039 (.001)	.039 (.001)	.038 (.001)	.039 (.001)	.036 (.001)	.038 (.001)	.037 (.001)	.037 (.001)
$\sigma_{\eta_3}$	.255 (.005)	.259 (.006)	.256 (.006)	.263 (.004)	.260 (.005)	.248 (.005)	.258 (.005)	.247 (.006)	.242 (.007)
$\rho_{\eta_1, \eta_2}$	.473 (.016)	.413 (.021)	.454 .021	.571 (.013)	.486 (.017)	.610 (.013)	.505 (.017)	.485 (.020)	.365 (.030)
$\rho_{\eta_1, \eta_3}$	-.604 (.003)	-.548 (.020)	-.586 .019	-.694 (.011)	-.618 (.015)	-.729 (.012)	-.636 (.016)	-.620 (.019)	-.509 (.029)
$\rho_{\eta_2, \eta_3}$	-.946 (.023)	-.948 (.003)	-.947 .003	-.945 (.002)	-.946 (.002)	-.941 (.003)	-.946 (.002)	-.943 (.003)	-.944 (.004)

Note: The first line corresponds to the ARMA specification (AR-MA) used for the random effect estimation. Standard errors in parentheses.

Table 4: Group averages of individual factor loadings  $\eta_g$ 

Skill group	Age group	Nb Obs	$\bar{\eta}_{g1}$	$\bar{\eta}_{g2}$	$\bar{\eta}_{g3}$
2	17	1268	2.4 (0.032)	0.04 (0.0067)	-0.15 (0.051)
3	17	1224	2.4 (0.039)	0.039 (0.0056)	-0.15 (0.04)
1	19	41	2.7 (0.038)	0.07 (0.0057)	-0.33 (0.042)
2	19	934	2.6 (0.035)	0.044 (0.0046)	-0.17 (0.034)
3	19	994	2.5 (0.04)	0.042 (0.007)	-0.17 (0.051)
1	21	117	2.9 (0.086)	0.052 (0.0085)	-0.22 (0.068)
2	21	710	2.7 (0.014)	0.047 (0.0024)	-0.2 (0.018)
3	21	512	2.6 (0.015)	0.041 (0.0026)	-0.19 (0.019)
1	23	171	3.1 (0.018)	0.055 (0.0036)	-0.24 (0.027)
2	23	348	2.7 (0.026)	0.05 (0.0037)	-0.21 (0.028)
3	23	254	2.7 (0.046)	0.051 (0.0053)	-0.25 (0.04)
1	25	191	3.3 (0.056)	0.061 (0.0066)	-0.29 (0.05)
2	25	146	2.8 (0.059)	0.038 (0.0065)	-0.14 (0.046)
3	25	93	2.6 (0.018)	0.033 (0.0031)	-0.09 (0.024)
1	27	114	3.4 (0.019)	0.047 (0.0045)	-0.21 (0.034)
2	27	87	3 (0.02)	0.061 (0.0034)	-0.32 (0.025)
3	27	63	2.7 (0.036)	0.03 (0.005)	-0.079 (0.039)
1	29	58	3.2 (0.052)	0.041 (0.0058)	-0.14 (0.041)
2	29	67	2.8 (0.084)	0.038 (0.013)	-0.2 (0.11)
3	29	55	2.6 (0.048)	0.017 (0.0074)	0.0061 (0.059)

Note: Estimation of equation (17). Standard errors in parentheses.

Table 5: Quantiles of individual effects  $\eta_i$ : unconstrained estimates

Individual effects	Sample periods	Quantiles						
		0.05	0.2	0.35	0.5	0.65	0.8	0.95
$\eta_1$	(3,15]	-2.97 (2.33)	0.483 (0.591)	1.74 (0.191)	2.43 (0.057)	3.04 (0.135)	4.31 (0.485)	8.14 (1.98)
	(15,22]	1.52 (0.0703)	2.11 (0.0237)	2.32 (0.0195)	2.51 (0.0226)	2.7 (0.0211)	3 (0.0314)	3.8 (0.0527)
	(22,26]	2 (0.0232)	2.25 (0.019)	2.39 (0.0165)	2.52 (0.0162)	2.67 (0.0179)	2.87 (0.021)	3.46 (0.0346)
	(26,28]	2.13 (0.0245)	2.35 (0.0164)	2.47 (0.015)	2.58 (0.0138)	2.7 (0.014)	2.86 (0.0153)	3.26 (0.0318)
$\eta_2$	(3,15]	-0.435 (0.176)	-0.162 (0.0471)	-0.0538 (0.0178)	0.0203 (0.00933)	0.0937 (0.0132)	0.205 (0.042)	0.539 (0.152)
	(15,22]	-0.124 (0.00967)	-0.032 (0.00434)	0.00555 (0.00318)	0.033 (0.00315)	0.0601 (0.00329)	0.0973 (0.0042)	0.194 (0.00963)
	(22,26]	-0.0471 (0.00394)	-0.00103 (0.00265)	0.0203 (0.00246)	0.0388 (0.00218)	0.0567 (0.00247)	0.0834 (0.00271)	0.141 (0.00558)
	(26,28]	-0.0218 (0.00327)	0.00914 (0.0022)	0.0254 (0.00215)	0.0383 (0.00207)	0.0526 (0.00217)	0.073 (0.00252)	0.114 (0.00414)
$\eta_3$	(3,15]	-6.27 (2.14)	-2.14 (0.578)	-0.706 (0.15)	-0.00853 (0.0869)	0.751 (0.207)	2.24 (0.698)	5.71 (2.51)
	(15,22]	-1.44 (0.0845)	-0.622 (0.0337)	-0.324 (0.0255)	-0.125 (0.0224)	0.0874 (0.0286)	0.395 (0.0316)	1.29 (0.12)
	(22,26]	-0.907 (0.0474)	-0.443 (0.0194)	-0.275 (0.017)	-0.142 (0.0175)	-0.024 (0.0174)	0.131 (0.0173)	0.445 (0.0339)
	(26,28]	-0.632 (0.0238)	-0.36 (0.0166)	-0.242 (0.0141)	-0.147 (0.0139)	-0.0542 (0.0149)	0.0559 (0.0164)	0.269 (0.0268)

Note: Sample period: Number of observed periods. Standard errors (sampling and parameter uncertainty, 1000 MC simulations) in brackets.



Table 6: Covariance matrix of centered individual effects: fixed and random effect estimation

Sample periods	$Var(\eta_1^c)$	$Cov(\eta_1^c, \eta_2^c)$	$Cov(\eta_1^c, \eta_3^c)$	$Var(\eta_2^c)$	$Cov(\eta_2^c, \eta_3^c)$	$Var(\eta_3^c)$
(3,15]	11 (16)	0.93 (1.3)	-12 (18)	0.093 (0.1)	-1.1 (1.4)	14 (19)
(15,22]	0.5 (0.08)	0.057 (0.01)	-0.57 (0.11)	0.01 (0.0016)	-0.09 (0.015)	0.83 (0.15)
(22,26]	0.14 (0.0073)	0.011 (0.0011)	-0.099 (0.0091)	0.0038 (0.00032)	-0.027 (0.0024)	0.2 (0.018)
(26,28]	0.076 (0.0039)	0.0043 (0.00058)	-0.038 (0.0041)	0.002 (0.00015)	-0.013 (0.00098)	0.09 (0.0067)
Complete sample	2.6 (3.5)	0.22 (0.28)	-2.8 (3.8)	0.024 (0.023)	-0.27 (0.31)	3.3 (4.2)
Random effects	0.093 (0.0036)	0.0059 (0.00051)	-0.05 (0.004)	0.0015 (0.00011)	-0.0093 (0.00079)	0.066 (0.0059)

Notes: The first four lines are obtained using fixed effect estimates. Sample periods = number of observed periods. Standard errors (sampling and parameter uncertainty, 1000 MC simulations) between brackets.

Table 7: Frequencies of violations: single-dimensional restriction

Restrictions $\rightarrow$ Sample periods	$\eta_2 \geq 0$	$\eta_3 \leq 0$	$\eta_3 + \pi_T \eta_2 \geq 0$
(3,15]	0.13 (0.041)	0.16 (0.046)	0.15 (0.05)
(15,22]	0.12 (0.019)	0.18 (0.021)	0.19 (0.022)
(22,26]	0.068 (0.012)	0.13 (0.018)	0.21 (0.024)
(26,28]	0.039 (0.01)	0.1 (0.02)	0.18 (0.024)

Notes: Sample periods = number of observed periods. 5 per cent level rejection frequency of single-dimensional tests of restrictions. Standard errors (sampling and parameter uncertainty, 1000 MC simulations) between brackets.

Sample periods	P-values < 0.10	0.05	0.01
(3,15]	0.16 (0.012)	0.12 (0.011)	0.078 (0.0084)
(15,22]	0.21 (0.011)	0.17 (0.0096)	0.12 (0.0083)
(22,26]	0.21 (0.0088)	0.17 (0.0081)	0.12 (0.007)
(26,28]	0.18 (0.0085)	0.15 (0.0078)	0.1 (0.0066)
Complete sample	0.19 (0.005)	0.15 (0.0045)	0.1 (0.0038)

Notes: Sample periods = number of observed periods. Frequency of p-values of the test of restrictions satisfying the conditions. Standard errors (sampling and parameter uncertainty, 20 Monte Carlo simulations) between brackets. The distribution of the test statistic is obtained using 150 replications.

Table 8: Frequencies of violations: global restriction

Table 9: Quantiles of  $\kappa$ ,  $\rho_L$  and  $c_L$ : simulated estimates

Individual effects	Sample periods	Quantiles						
		0.05	0.2	0.35	0.5	0.65	0.8	0.95
$\kappa$	(3,15]	6.38 (0.284)	7.43 (0.292)	8.03 (0.253)	8.56 (0.284)	9.15 (0.298)	9.83 (0.239)	10.8 (0.344)
	(15,22]	4.51 (0.238)	6.22 (0.212)	7 (0.215)	7.65 (0.219)	8.29 (0.269)	9.07 (0.462)	11 (0.358)
	(22,26]	2.32 (0.131)	4.65 (0.228)	5.81 (0.265)	6.81 (0.272)	7.86 (0.256)	9.27 (0.405)	12.4 (0.318)
	(26,28]	1.4 (0.102)	3.25 (0.194)	4.85 (0.229)	6.27 (0.29)	7.76 (0.343)	9.91 (0.441)	13.9 (0.4)
$\rho_L$	(3,15]	0.16 (0.0792)	0.226 (0.0815)	0.285 (0.0921)	0.357 (0.102)	0.455 (0.148)	0.702 (0.199)	1.96 (0.525)
	(15,22]	0.14 (0.0947)	0.254 (0.076)	0.34 (0.0842)	0.451 (0.111)	0.621 (0.151)	0.996 (0.245)	2.99 (0.601)
	(22,26]	0.11 (0.0589)	0.227 (0.113)	0.364 (0.123)	0.546 (0.149)	0.797 (0.206)	1.39 (0.283)	5.13 (0.861)
	(26,28]	0.0866 (0.0603)	0.174 (0.139)	0.339 (0.17)	0.565 (0.209)	0.966 (0.256)	1.89 (0.356)	6.82 (0.894)
$c_L$	(3,15]	6.11 (1.19)	22.8 (1.75)	52 (2.52)	109 (3.61)	249 (5.77)	879 (12.8)	10994 (168)
	(15,22]	5.79 (1.6)	65.3 (2.13)	162 (3.28)	359 (5.11)	944 (11.4)	3380 (31.8)	52673 (678)
	(22,26]	2.42 (0.564)	50.2 (2.77)	200 (4.74)	528 (8.17)	1544 (14.6)	6160 (57.9)	128903 (1911)
	(26,28]	1.38 (0.475)	20.6 (2.74)	151 (6.02)	586 (10.5)	1954 (20.7)	8407 (44.5)	164666 (1131)

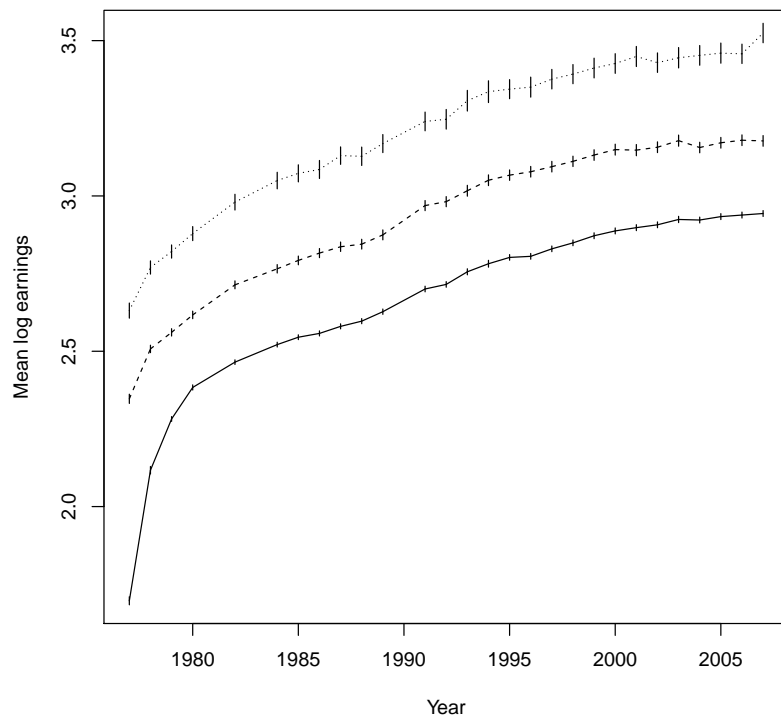
Notes: Sample period: Number of observed periods. Standard errors (sampling and parameter uncertainty, 100 MC simulations) in brackets.

Table 10: Short term inequalities and their decomposition

	Short term	Decomposition	
		Perm. (%)	Trans. (%)
1977	.167	.033	.966
1982	.086	.507	.492
1987	.102	.624	.375
1992	.126	.709	.290
1997	.146	.769	.230
2002	.152	.823	.176
2007	.151	.886	.113
Mean	.129	.648	.351

Notes: Inequality is measured with the variance of logs. Short term inequality: cross sectional inequality. Perm. stands for the share of cross sectional inequality due to the permanent heterogeneity components. Trans. stands for the share of cross-section inequality which is due to the transitory component.

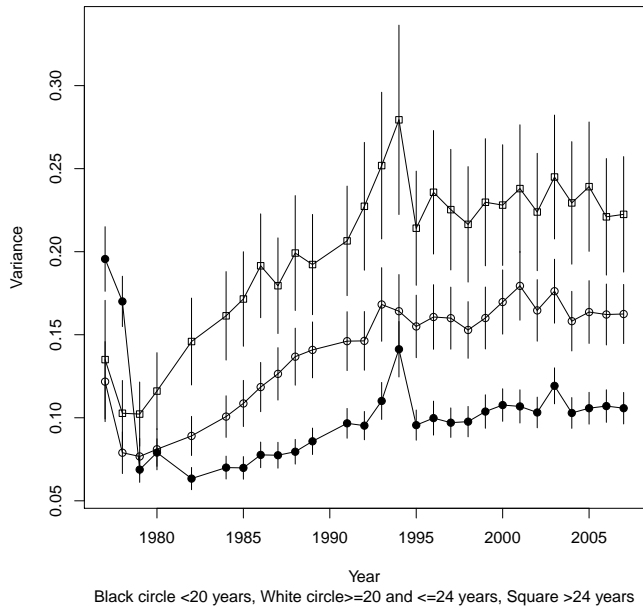
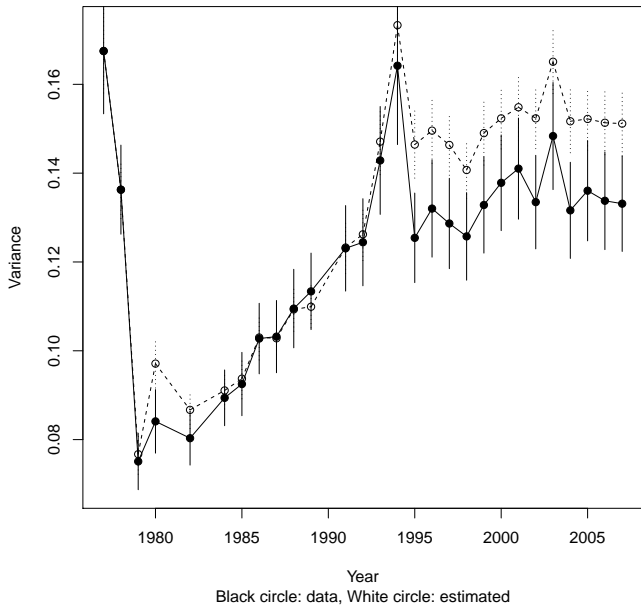
Figure 1: Mean log earnings by age at entry: 1977-2007



Black: <20 years, Dashed: >=20 and <24 years, Small dashed >24 years

Note: The small vertical lines represent the 95% confidence intervals.

Figure 2: Cross-sectional variance of earnings: 1977-2007



Note: The small vertical lines represent the 95% confidence intervals.

(A) full sample

(B) by age of entry

Figure 3: Autocorrelations with 1986 and 2007

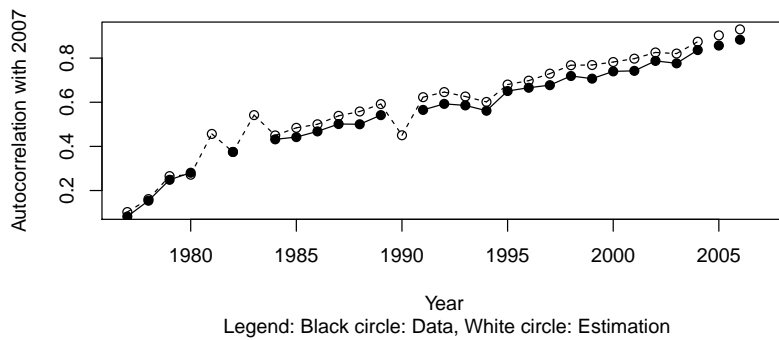
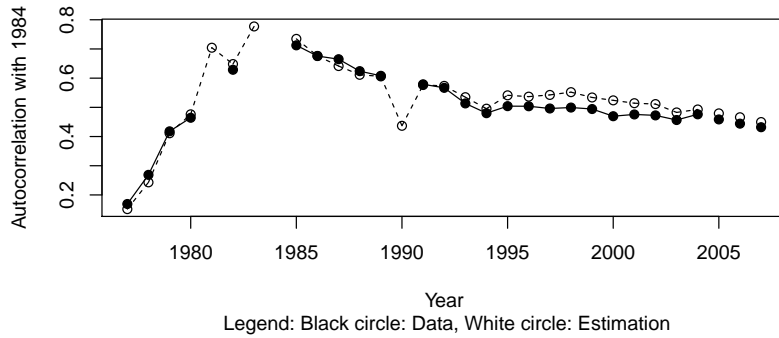


Figure 4: Forward autocorrelations of order 1 and of order 6

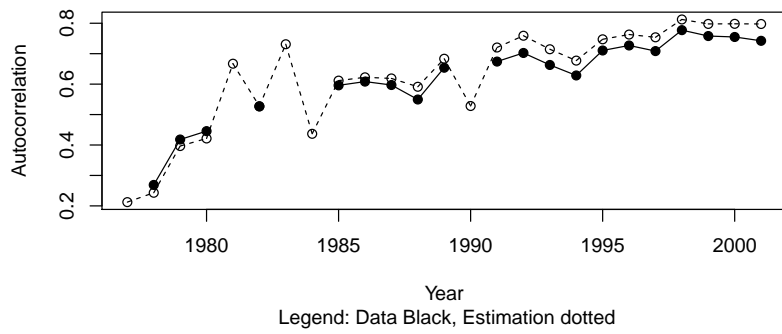
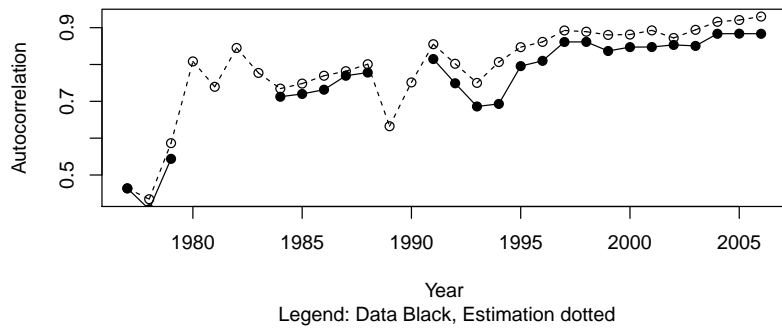
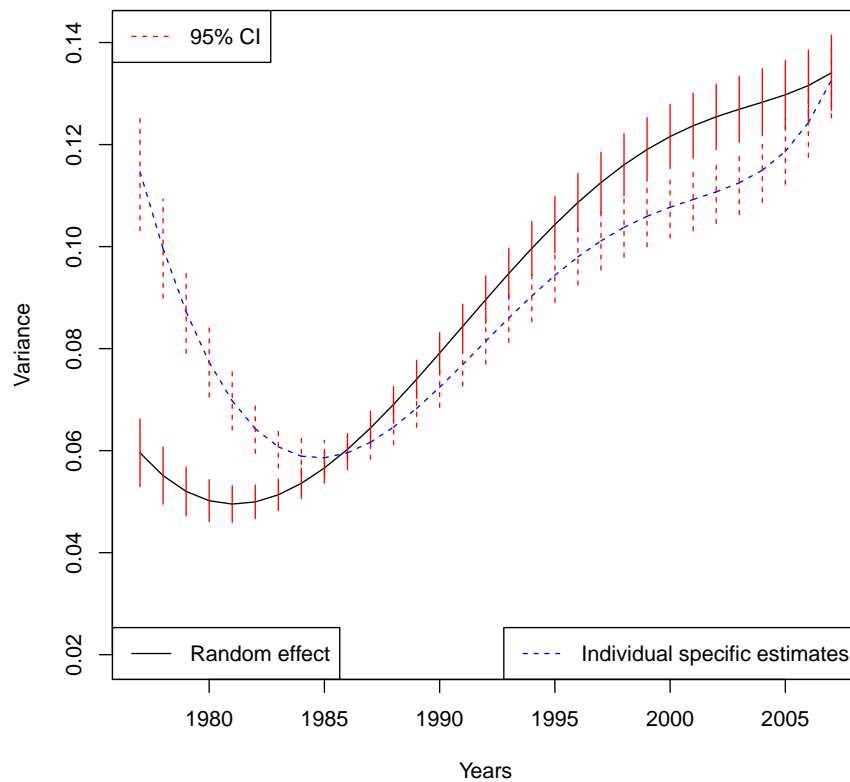


Figure 5: Variance of the permanent components



Note: The permanent component is  $M(\beta)\eta$  defined in equation (23). The sample is restricted to long history profiles (more than 22 periods). "Random effects" are using estimates derived from random effect estimation. "Individual specific effects" are using estimates derived from fixed effect estimation.



Figure 6: Scatter plot of  $\eta_2$  and  $\eta_3$  and the area describing the structural constraint

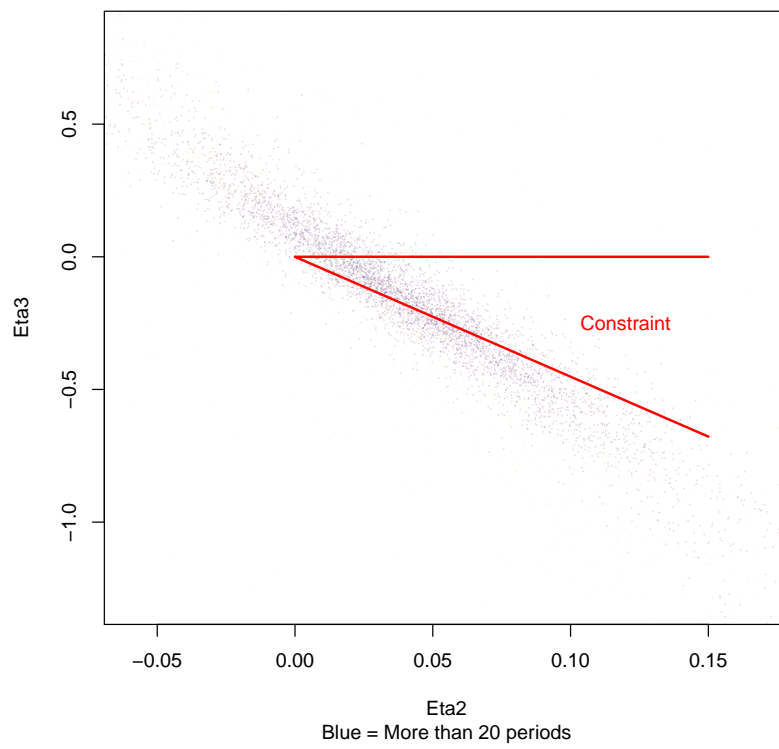
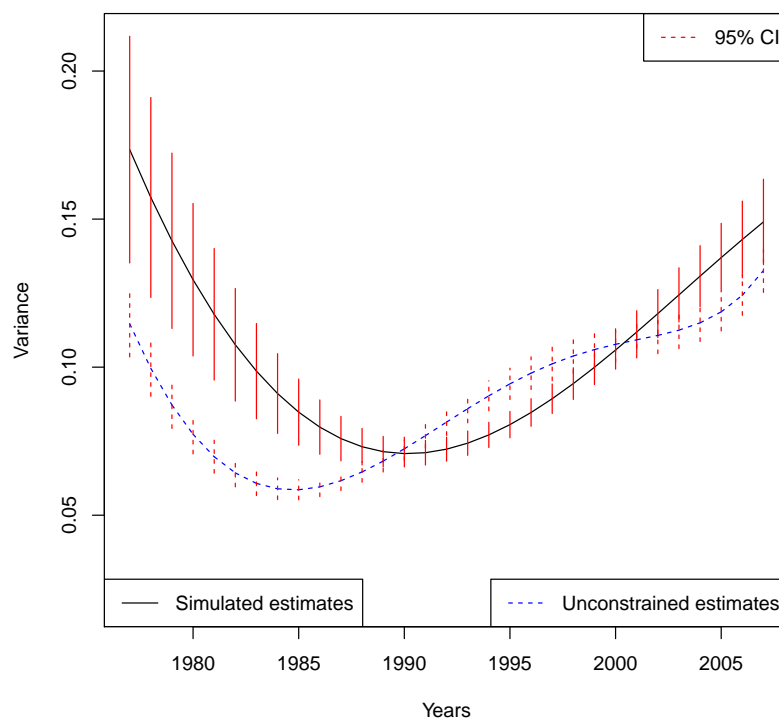
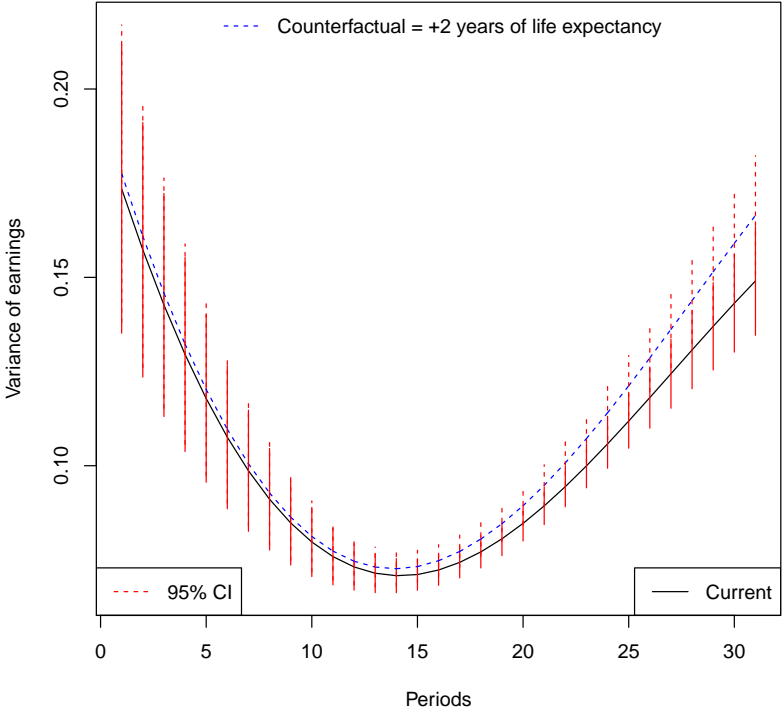
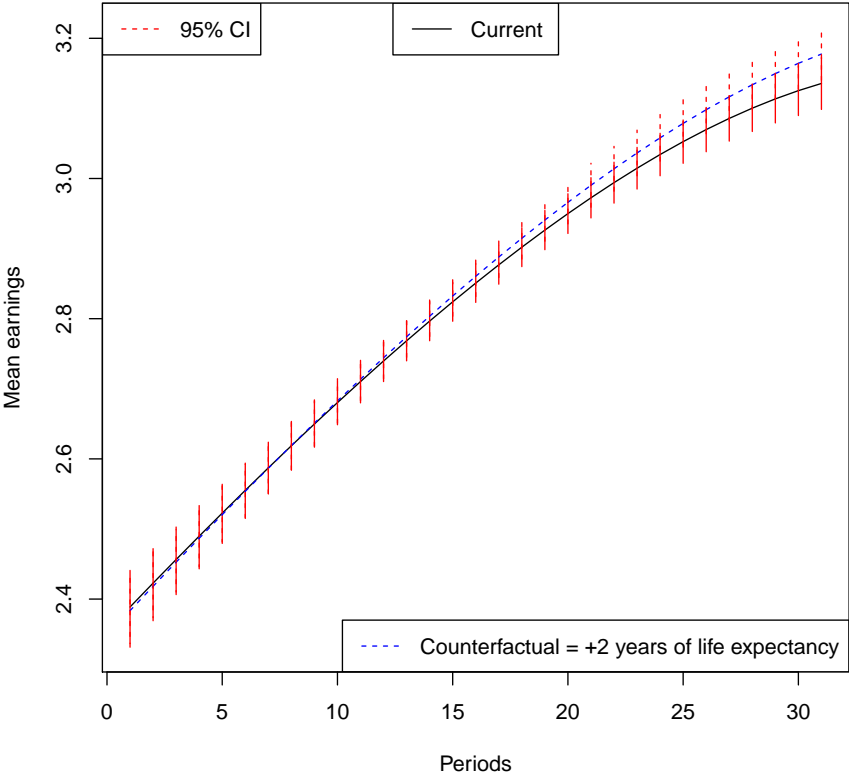


Figure 7: Earnings variances (permanent components): Unconstrained estimates and simulated constrained estimates



Note: The permanent component is  $M(\beta)\eta$ . Sample restricted to long history profiles (more than 22 periods). Simulated estimates = estimates drawn in the posterior constrained distribution. Unconstrained estimates derived from fixed effect estimation.

Figure 8: Counterfactual: Additional Years of Life Expectancy (K=2), Mean (Top panel) and Variance (Bottom Panel) Lower bound Impact



Note: Sample of 4292 long history profiles (more than 22 periods). Standard errors are due to sampling and parameter uncertainty (30 Monte Carlo replications) 67

# Supplementary Appendices

## A.I Notation

### A.I.1 The Model

- $t$ : time elapsed since the entry in the labor market.
- $i$ : index for individuals.
- $y_i^P(t)$ : potential individual earnings.
- $H_i(t)$ : individual-specific human capital.
- $\delta_i(t)$ : log rental rate of the human capital.
- $y_i(t)$ : current individual earnings.
- $\tau_i(t)$ : level of investment in human capital.
- $\rho_i$ : individual-specific rate of return of human capital investments (in log).
- $\lambda_i(t)$ : individual-specific depreciation rate of human capital, at time  $t$ .
- $c_i$ : individual-specific cost of human capital investments in utility terms.
- $V_t(H_i(t), \tau_i(t))$ : Sum of inter-temporal utilities, function of a state variable,  $H_i(t)$  and a control variable,  $\tau_i(t)$
- $\beta_i$ : Individual-specific discount rate.
- $W_t(H_i(t))$ : optimized sum of inter-temporal utilities, with respect to the control variable. In the consumption smoothing section of the appendix, it depends also on the accumulated savings or debt  $A_i(t)$  (see below).
- $T$ : Arbitrary date at which we examine whether individuals goes on investing in human capital, last date of observation in the empirical application.
- $T_i$ : Individual-specific date at which investing in human capital stops.
- $\tilde{T}_i = \min(T, T_i)$ : Individual-specific date at which investing in human capital stops, censored with  $T$ .
- $\delta_i^*$ : Sum of discounted prices of human capital after period  $T$ , which affects the utility in level.
- $\kappa_{i,t}$ : Discounted value of the stock of one unit of (log) human capital over the remaining period of life after  $t + 1$ , if the worker stops her investments.

- $\kappa_i$ : Discounted value of the stock of one unit of log(human capital) over the remaining period of life after  $T + 1$ ,  $\kappa_i = \kappa_{i,T}$ .
- $C_i(t)$ : consumption level at time  $t$ .
- $A_i(t)$ : accumulated savings or debt at time  $t$ .
- $s_i(t)$ : saving rate, share of earnings dedicated to savings or indebtedness.
- $r_{it}$ : individual-specific interest rate at time  $t$ .
- $HC_i(t)$ : Human capital contribution to log-earnings.
- $\eta_{i1}$ : individual-specific fixed level of log-earnings.
- $\eta_{i2}$ : individual-specific growth rate of log-earnings.
- $\eta_{i3}$ : individual-specific degree of curvature of log-earnings.
- $v_{it}$ : (log) price of human capital net of cumulative depreciation.
- $\Lambda_i(t)$ : Cumulative depreciation of human capital since the entry on the labor market.

### A.I.2 Identifying and economic restrictions

- $\eta_{i4}$ : individual-specific degree of curvature of log-earnings interacted with a linear trend (not estimated).
- $\beta$ : homogenous discount rate
- $g$ : group of workers, defined by their age at entry and their skills
- $\overline{\ln y_{gt}}$ : average of  $\ln y_{it}$  over the group  $g$
- $\eta_{gk}$ : average of  $\eta_{ik}$  over  $g$ , for  $k = 1, 2, 3$
- $v_{gt}$ : average of  $v_{it}$  over  $g$
- $\delta_g(t)$ : average of  $\delta_i(t)$  over  $g$
- $\lambda_g(t)$ : average of  $\lambda_i(t)$  over  $g$
- $\Lambda_g(t)$ : average of  $\Lambda_i(t)$  over  $g$
- $\eta_{ik}^c$ : centered individual effect of  $\eta_{ik}$ , for  $k = 1, 2, 3$
- $u_{it}$ : centered earnings, with respect to group  $g$
- $v_{it}^c$ : individual-specific variations of human capital prices

- $\rho_i^L$ : minimum value of  $\rho_i$
- $c_i^L$ : minimum value of  $c_i$

### A.I.3 Econometric Modeling of Earnings Dynamics

- $\ln y_i, u_i, v_i, v_i^c$ :  $T$ -vectors for  $\ln y_{it}, u_{it}, v_{it}, v_{it}^c$
- $\eta_i$ : vector of individual fixed effects  $(\eta_{i1}, \eta_{i2}, \eta_{i3})$
- $\eta_i^c$ : vector of centered individual fixed effects
- $M(\beta)$ :  $T, 3$  matrix of factors.
- $\Omega(\eta_i^c)$ : covariance matrix of centered individual fixed effects.
- $\hat{\eta}_i^c$ : estimate of the centered individual fixed effect.
- $B$ : matrix  $3, T$  establishing the relationship between the centered individual fixed effects and the earnings residuals.
- $\hat{B}$ : estimate of  $B$
- $\tilde{\eta}_i^c$ : unfeasible estimator of  $\eta_i^c$  using  $B$
- $w_i$ :  $T$ -vector of residuals, orthogonal to  $\eta_i^c$
- $\Omega_w$ : covariance matrix of  $w_i$
- $\pi_T$ :  $\frac{\beta^{T+1}}{1-\beta}$
- $\hat{\eta}_i^R$ : constrained estimate of  $\eta_i$
- $\Omega_\eta$ : covariance matrix of  $\eta_i$
- $\hat{\Omega}_\eta$ : estimate of  $\Omega_\eta$
- $\hat{\eta}_i^S$ : simulated constrained estimates of  $\eta_i$
- $\kappa_i^*$ : counterfactual value of the structural parameter  $\kappa_i$
- $\eta_{ik}^*$ : counterfactual value of  $\eta_{ik}$ , for  $k = 1, 2, 3$ , from  $\kappa_i^*$
- $N_i$ : number of actual observations for the individual  $i$ .

## A.II The random effect model : Model Specification and Likelihood function

The main difference with standard specifications lies in the introduction of three individual heterogeneity factors that interact in a specific way with factors depending on time. Equation (23) writes

$$u_i^{[1,T]} = M(\beta)^{[1,T]} \eta_i^c + v_i^{c[1,T]}$$

where  $u_i^{[1,T]} = (u_{i1}, \dots, u_{iT})'$ ,  $v_i^{c[1,T]} = (v_{i1}^c, \dots, v_{iT}^c)'$ ,  $\eta_i^c = (\eta_{i1}^c, \eta_{i2}^c, \eta_{i3}^c)$  are the centered versions of the  $\eta$ s and:

$$M(\beta)^{[1,T]} = \begin{bmatrix} 1 & 1 & 1/\beta \\ \vdots & \vdots & \vdots \\ 1 & T & 1/\beta^T \end{bmatrix},$$

is a  $[T, 3]$  matrix. The system is further completed by initial conditions, the number of which depends on the order of the autoregressive process. Denote  $p$  this order and write the initial conditions as:

$$u_i^{[1-p,0]} = v_i^{c[1-p,0]}$$

since unrestricted dependence between  $v_i^{c[1,T]}$ ,  $\eta_i^c$  and those initial conditions will be allowed for. We can rewrite the whole system as:

$$u_i^{[1-p,T]} = M(\beta)^{[1-p,T]} \eta_i^c + v_i^{c[1-p,T]}$$

in which the matrix  $M(\beta)^{[1-p,T]}$  is completed by  $p$  rows equal to zero,  $M(\beta)^{[1-p,0]} = 0$ .

We now go further and specify the correlation structure. A comment is in order. Usually, the autoregressive structure directly applies to earnings residuals  $u_{it}$  and in the absence of covariates, this is equivalent to specifying it through the residual part  $v_{it}^c$  because there is a single individual effect. This equivalence still holds when another heterogeneity factor interacted with a linear trend is present. Nevertheless, our specification includes a third factor interacted with a geometric term and this breaks the equivalence. To circumvent this problem, we posit that  $v_{it}^c$  is a (time heteroskedastic) ARMA process whose innovations are independent of the individual heterogeneity terms,  $\eta_i^c$ . As a consequence, our variable of interest,  $u_{it}$ , is the sum of two processes, the first one being related to fixed individual heterogeneity and the second one to the pure dynamic process. These processes are assumed to be independent between them although they are both correlated with initial conditions,  $u_i^{[1-p,0]}$ .

We now derive the covariance matrix of  $u_i^{[1-p,T]}$  as a function of the parameters of these processes in two steps. We first describe the ARMA process and then include the individual heterogeneity factors.

## A.II.1 Time heteroskedastic ARMA specification

Following Alvarez and Arellano (2004) or Guvenen (2009), we specify

$$v_{it}^c = \alpha_1 v_{it-1}^c + \dots + \alpha_p v_{it-p}^c + \sigma_t w_{it}$$

where  $w_{it}$  is  $MA(q)$ :

$$w_{it} = \zeta_{it} - \psi_1 \zeta_{it-1} - \dots - \psi_q \zeta_{it-q}.$$

Let  $\alpha = (\alpha_1, \dots, \alpha_p)$  and  $M_T(\alpha)$  a matrix of size  $[T, T + p]$  where  $p = \dim(\alpha)$ :

$$M_T(\alpha) = \begin{pmatrix} -\alpha_p & \dots & -\alpha_1 & 1 & 0 & \dots & 0 \\ 0 & -\alpha_p & \dots & -\alpha_1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\alpha_p & \dots & -\alpha_1 & 1 \end{pmatrix}.$$

As  $v_i^{c[1-p, T]} = (v_{i1-p}^c, \dots, v_{iT}^c)$ , we have:

$$\begin{pmatrix} \left( \begin{array}{cc} I_p & 0 \end{array} \right) \\ M_T(\alpha) \end{pmatrix} v_i^{c[1-p, T]} = \begin{pmatrix} v_i^{c[1-p, 0]} \\ \sigma_t w_i^{[1, T]} \end{pmatrix}$$

Since  $w_{it}$  is  $MA(q)$ , we have

$$w_i^{[1, T]} = M_T(\psi) \cdot \zeta_i^{[1-q, T]}$$

where  $\zeta_i^{[1-q, T]} = (\zeta_{i1-q}, \dots, \zeta_{iT})$ .

Denote  $\Lambda$  a diagonal matrix whose diagonal is  $(\sigma_1, \dots, \sigma_T)$  to get the following description of the stochastic process as a function of initial conditions and idiosyncratic errors:

$$\begin{pmatrix} I_p & 0 \\ M_T(\alpha) \end{pmatrix} \cdot v_i^{c[1-p, T]} = \begin{pmatrix} I_p & 0_{p, T+q} \\ 0_{T, p} & \Lambda \cdot M_T(\psi) \end{pmatrix} \begin{pmatrix} v_i^{c[1-p, 0]} \\ \zeta_i^{[1-q, T]} \end{pmatrix}. \quad (\text{A.II.1})$$

To compute the covariance of  $v_i^{c[1-p, T]}$ , we derive the covariance matrix of  $\begin{pmatrix} v_i^{c[1-p, 0]} & \zeta_i^{[1-q, T]} \end{pmatrix}$ . Since  $\zeta_i^{[1-q, T]}$  are i.i.d and are of variance 1, the South-East corner of the matrix is the identity matrix of size  $(1 + q + T)$ . The North West corner is assumed to be an unrestricted covariance matrix  $V u_i^{[1-p, 0]} = \Gamma_{00}$ . Assuming as usual that  $E(u_{i\tau} \zeta_{it}) = 0$  for any  $\tau < t$ , we have that  $E(v_i^{c[1-p, 0]} \cdot (\zeta_i^{[1, T]})') = 0$ . Only  $E(u_i^{[1-p, 0]} \cdot (\zeta_i^{[1-q, 0]})')$  remains to be defined:

$$E(v_i^{c[1-p, 0]} \cdot (\zeta_i^{[1-q, 0]})') = \Omega = [\omega_{rs}]$$

where  $r \in [1 - p, 0]$  and  $s \in [1 - q, 0]$  and where:

$$\begin{aligned} r < s &: & \omega_{rs} &= 0 \\ r \geq s &: & \omega_{rs} &\text{ is not constrained} \end{aligned}$$

because the innovation  $\zeta_{is}$  is drawn after  $r$  and is assumed to be not correlated with  $y_i^r$ .



Hence the covariance matrix of  $z_i = \begin{pmatrix} v_i^{c[1-p,0]} \\ \zeta_i^{[1-q,T]} \end{pmatrix}$  writes :

$$\Omega_z = V \begin{pmatrix} v_i^{c[1-p,0]} \\ \zeta_i^{[1-q,T]} \end{pmatrix} = V \begin{pmatrix} v_i^{c[1-p,0]} \\ \zeta_i^{[1-q,0]} \\ \zeta_i^{[1,T]} \end{pmatrix} = \begin{pmatrix} \Gamma_{00} & \Omega & 0 \\ \Omega' & I_q & 0 \\ 0 & 0 & I_T \end{pmatrix}.$$

## A.II.2 Individual heterogeneity

The covariance matrix of the individual heterogeneity factors is denoted  $\Sigma_\eta$ . as said above, we assume that the fixed heterogeneity terms are independent of the whole innovation process  $\zeta_i^{[1-q,T]}$ . As for the covariance structure between initial conditions and those factors, we assume that:

$$E \left( v_i^{c[1-p,0]} \eta_i^{c'} \right) = \Gamma_{0\eta}$$

Consider the covariance matrix of initial conditions  $\Sigma$  :

$$\Sigma = V \begin{pmatrix} v_i^{c[1-p,0]} \\ \eta_i^c \\ \zeta_i^{[1-q,0]} \end{pmatrix} = \begin{pmatrix} \Gamma_{00} & \Gamma_{0\eta} & \Omega \\ \Gamma_{0\eta}' & \Sigma_\eta & 0 \\ \Omega & 0 & I_q \end{pmatrix}.$$

and define,

$$R_T(\alpha) = \begin{pmatrix} (I_p & 0) \\ M_T(\alpha) \end{pmatrix}^{-1}$$

$$S_{T,p}(\psi, \Lambda) = \begin{pmatrix} I_p & 0_{p,T+q} \\ 0_{T,p} & \Lambda.M_T(\psi) \end{pmatrix}$$

Write the covariance matrix of vector  $u_i^{[1-p,T]}$ :

$$\begin{aligned} \Omega_u &= V \left( u_i^{[1-p,T]} \right) = V \left( v_i^{c[1-p,T]} + M(\beta)^{[1-p,T]} \eta_i^c \right) \\ &= V \left( \left[ M(\beta)^{[1-p,T]}, R_T(\alpha).S_{T,p}(\psi, \Lambda) \right] \begin{pmatrix} \eta_i^c \\ v_i^{c[1-p,0]} \\ \zeta_i^{[1-q,T]} \end{pmatrix} \right) \end{aligned}$$

Since  $v_i^{c[1-p,T]} = R_T(\alpha).S_{T,p}(\psi, \Lambda) \begin{pmatrix} v_i^{c[1-p,0]} \\ \zeta_i^{[1-q,T]} \end{pmatrix}$ , the matrix

$$V \left( v_i^{c[1-p,T]} \right) = R_T(\alpha).S_{T,p}(\psi, \Lambda).\Omega_z.S_{T,p}(\psi, \Lambda)'R_T(\alpha)'$$

and

$$\begin{aligned}
E \left( v_i^{c[1-p,T]} \eta_i' \right) M(\beta)^{[1-p,T]'} &= R_T(\alpha) \cdot S_{T,p}(\psi, \Lambda) E \left( \begin{array}{c} v_i^{c[1-p,0]}(\eta_i^c)' \\ \zeta_i^{[1-q,T]}(\eta_i^c)' \end{array} \right) M(\beta)^{[1-p,T]'} \\
&= R_T(\alpha) \cdot S_{T,p}(\psi, \Lambda) \left( \begin{array}{c} \Gamma_{0\eta} \\ 0_{T+q,3} \end{array} \right) M(\beta)^{[1-p,T]'} \\
&= R_T(\alpha) \cdot \left( \begin{array}{cc} I_p & 0_{p,T+q} \\ 0_{T,p} & \Lambda \cdot M_T(\psi) \end{array} \right) \left( \begin{array}{c} \Gamma_{0\eta} \\ 0_{T+q,3} \end{array} \right) \left( 0_{3,p}, M(\beta)^{[1,T]'} \right) \\
&= R_T(\alpha) \cdot \left( \begin{array}{cc} I_p & 0_{p,T+q} \\ 0_{T,p} & \Lambda \cdot M_T(\psi) \end{array} \right) \left( \begin{array}{cc} 0_{p,p} & \Gamma_{0\eta} M(\beta)^{[1,T]'} \\ 0_{T+q,p} & 0_{T+q,T} \end{array} \right) \\
&= R_T(\alpha) \cdot \left( \begin{array}{cc} 0_{p,p} & \Gamma_{0\eta} M(\beta)^{[1,T]'} \\ 0_{T,p} & 0_{T,T} \end{array} \right)
\end{aligned}$$

Hence,

$$\begin{aligned}
\Omega_u &= R_T(\alpha) \cdot S_{T,p}(\psi, \Lambda) \cdot \Omega_z \cdot S_{T,p}(\psi, \Lambda)' R_T(\alpha)' + M(\beta)^{[1-p,T]} \Sigma_\eta M(\beta)^{[1-p,T]'} \\
&\quad + R_T(\alpha) \cdot \left( \begin{array}{cc} 0_{p,p} & \Gamma_{0\eta} M(\beta)^{[1,T]'} \\ 0_{T,p} & 0_{T,T} \end{array} \right) + \left( \begin{array}{cc} 0_{p,p} & 0_{p,T} \\ M(\beta)^{[1,T]} \Gamma_{0\eta}' & 0_{T,T} \end{array} \right) R_T(\alpha)'
\end{aligned}$$

The two first terms correspond to variances of the dynamic process and the individual heterogeneity factors, the other terms correspond to the correlation between the two processes induced by initial conditions. Note that the parameters of the MA process don't appear in the correlation between the two processes since innovations are assumed to be independent with individual heterogeneity factors. Initial conditions are given by  $\zeta_i^{[1-q,0]}$ ,  $\eta^c$  and  $v_i^{c[1-p,0]}$ .

The Choleski decomposition of matrix  $\Sigma$  can be parametrized expressing the following matrix into a polar coordinate basis.

$$\left( \begin{array}{cccccccc}
1 & 0 & \dots & & & & & 0 \\
0 & \ddots & \ddots & \ddots & & & & \vdots \\
\vdots & 0 & 1 & 0 & & & & \vdots \\
\dots & \dots & 0 & 1 & 0 & 0 & 0 & \ddots \\
& & & \omega_\eta^{12} & 1 & 0 & & \vdots \\
& & & \omega_\eta^{13} & \omega_\eta^{23} & 1 & 0 & \vdots \\
\vdots & & \theta_{1-q,1-p}^{(1)} & \theta_{\eta 1,1-p} & \theta_{\eta 2,1-p} & \theta_{\eta 3,1-p} & 1 & \vdots \\
0 & & & \vdots & \vdots & \vdots & \theta_{2-p,2-p} & \vdots \\
\theta_{0,0}^{(1)} & & & \theta_{\eta 1,0} & \theta_{\eta 21,0} & \theta_{\eta 3,0} & \dots & \theta_{0,0} & 1
\end{array} \right)$$

where  $\theta_{1-q,1-p}^{(1)} = 0$  if  $p > q$  and, more generally,  $\theta_{l,m}^{(1)} = 0$  if  $l > m$ .

## A.III Fixed Effect Estimation, Constrained Effects and Counterfactuals

### A.III.1 Estimates of individual factors given observed earnings

The main equation is:

$$u_i^{[1-p,T]} = M(\beta)^{[1-p,T]} \eta_i^c + v_i^{c[1-p,T]},$$

where  $\eta_i^c$  and  $v_i^{c[1-p,T]}$  are centered by construction and where a row of  $M(\beta)$  is defined as  $M(\beta)^{[t]} = (1, t, 1/\beta^t)$  as in Appendix A.II.

Later on, we shall reintroduce the estimated averages,  $\bar{\eta}_g$ , of the individual effects that we estimate by OLS using the sub-groups defined by age of entry and skill level (21 groups). Define the average in each group as  $\bar{y}_g^{[1-p,T]}$  and define:

$$\widehat{\eta}_g = (M(\beta)^{[1-p,T]'} M(\beta)^{[1-p,T]})^{-1} M(\beta)^{[1-p,T]'} \bar{y}_g^{[1-p,T]}.$$

We now present the fixed effect estimation of  $\eta_i^c$ . We consider first the case with no missing values and extend it to the case with missing values. We finally analyze how to deal in the simulations with constraints on  $\eta_i$ .

### A.III.2 Estimating individual effects

Assume first that there are no missing values. To deal with the correlation between  $\eta_i^c$  and  $v_i$ , we can always write:

$$v_i^{c[1-p,T]} = C\eta_i^c + w_i^{[1-p,T]},$$

where  $E((\eta_i^c)' w_i^{[1-p,T]}) = 0$  so that we get:

$$C = E(v_i^{c[1-p,T]} (\eta_i^c)') (E(\eta_i^c (\eta_i^c)'))^{-1},$$

and:

$$\Omega_w = E(v_i^{c[1-p,T]} v_i^{[1-p,T]'}) - E(v_i^{c[1-p,T]} (\eta_i^c)') (E(\eta_i^c (\eta_i^c)'))^{-1} E(\eta_i^c v_i^{c[1-p,T]'}).$$

This yields the estimating equation for  $\eta_i^c$  :

$$u_i^{[1-p,T]} = D\eta_i^c + w_i^{[1-p,T]} \text{ where } D = M(\beta)^{[1-p,T]} + C,$$

that might be estimated by GLS methods.

It is nevertheless useful to write likelihood functions that will help later to define constrained estimates. Define the conditional (pseudo) likelihood function as:

$$L(u_i^{[1-p,T]} | \eta_i^c) = \frac{1}{(2\pi)^{T/2} \det \Omega_w^{1/2}} \exp \left( -\frac{1}{2} (u_i^{[1-p,T]} - D\eta_i^c)' \Omega_w^{-1} (u_i^{[1-p,T]} - D\eta_i^c) \right),$$

in which  $\Omega_w = V(w_i^{[1-p,T]})$ .

We are seeking the conditional distribution of  $\eta_i^c$  conditional on the observed  $u_i^{[1-p,T]}$  which can be expressed by Bayes law, using a prior for  $\eta_i^c$ ,  $L_0(\eta_i^c)$  as:

$$L(\eta_i^c | u_i^{[1-p,T]}) = \frac{L(u_i^{[1-p,T]} | \eta_i^c) L_0(\eta_i^c)}{\int L(u_i^{[1-p,T]} | \eta_i^c) L_0(\eta_i^c) d\eta_i^c}.$$

Consequently, the distribution function  $L(\eta_i^c | u_i^{[1-p,T]})$  can be written as:

$$H(u_i^{[1-p,T]}) \cdot \exp \left( -\frac{1}{2} (\eta_i^c - B u_i^{[1-p,T]})' \Omega_\eta^{-1} (\eta_i^c - B u_i^{[1-p,T]}) \right) L_0(\eta_i^c)$$

where the constant of integration is derived by setting to one the integral over  $\eta_i^c$ . In the case of a diffuse prior i.e.  $L_0(\eta_i^c) = 1$ , the constant of integration is no longer dependent on  $u_i^{[1-p,T]}$  and is equal to the usual reciprocal of  $(2\pi)^{3/2} \det \Omega_\eta^{1/2}$ . When there are constraints on  $\eta_i^c$ , these constraints can be included in the prior (see below).

As all terms in  $\eta_i^c$  and  $u_i^{[1-p,T]}$  are quadratic, we can derive the unknown matrices  $B$  and  $\Omega_\eta$  by solving:

$$(u_i^{[1-p,T]} - D\eta_i^c)' \Omega_w^{-1} (u_i^{[1-p,T]} - D\eta_i^c) = (\eta_i^c - Bu_i^{[1-p,T]})' \Omega_\eta^{-1} (\eta_i^c - Bu_i^{[1-p,T]}) + u_i^{[1-p,T]'} Au_i^{[1-p,T]}.$$

By identifying quadratic terms in  $(\eta_i^c, \eta_i^c)$ ,  $(u_i^{[1-p,T]}, \eta_i^c)$  and  $(u_i^{[1-p,T]}, u_i^{[1-p,T]})$ , we obtain three equations:

$$\begin{cases} D' \Omega_w^{-1} D &= \Omega_\eta^{-1}, \\ -D' \Omega_w^{-1} &= -\Omega_\eta^{-1} B, \\ \Omega_w^{-1} &= B' \Omega_\eta^{-1} B + A, \end{cases}$$

so that, as  $D' \Omega_w^{-1} D$  is invertible:

$$\begin{cases} \Omega_\eta &= (D' \Omega_w^{-1} D)^{-1}, \\ B &= (D' \Omega_w^{-1} D)^{-1} D' \Omega_w^{-1}, \\ A &= \Omega_w^{-1} - \Omega_w^{-1} D (D' \Omega_w^{-1} D)^{-1} D' \Omega_w^{-1}. \end{cases}$$

If those matrices are known, the (unfeasible) estimator for the individual fixed effects, by reinclusion of the estimated averages, are:

$$\tilde{\eta}_i^c = Bu_i^{[1-p,T]} = B(D\eta_i^c + w_i^{[1-p,T]}) = \eta_i^c + Bw_i^{[1-p,T]}.$$

They are such that:

$$\begin{aligned} V(\tilde{\eta}_i^c) &= EV(\tilde{\eta}_i^c | \eta_i^c) + VE(\tilde{\eta}_i^c | \eta_i^c) \\ \implies V(\tilde{\eta}_i^c) &= B\Omega_w B' + V\eta_i^c = \Omega_\eta + V\eta_i^c. \end{aligned}$$

The term  $\Omega_\eta$  goes to zero at least at the rate  $1/T$  since matrix  $D$  is determined by different factors which are going to zero at least as fast as  $T^{-1}$ .

The feasible estimator is now given by:

$$\hat{\eta}_i^c = \hat{B}u_i^{[1-p,T]},$$

and by reinclusion of the estimated averages for each group,  $\bar{\eta}_{g \ni i} = \bar{\eta}_g$ , we have:

$$\hat{\eta}_i = \bar{\eta}_g + \hat{\eta}_i^c = \bar{\eta}_g + \hat{B}u_i^{[1-p,T]},$$

We now analyse the case with missing values. Suppose that  $u_i^{[1-p,T]}$  is not observable, only  $S_i u_i^{[1-p,T]}$  is where  $S_i$  is the matrix of dimension  $(N_i, T + p + 1)$  selecting non missing values and where  $N_i$  is the number of such non missing values. Consequently, the distribution function  $L(\eta_i^c | S_i u_i^{[1-p,T]})$  becomes:

$$H_i(S_i u_i^{[1-p,T]}) \cdot \exp\left(-\frac{1}{2}(\eta_i^c - B_i S_i u_i^{[1-p,T]})' \Omega_{\eta_i}^{-1} (\eta_i^c - B_i S_i u_i^{[1-p,T]})\right) L_0(\eta_i^c),$$

where by simple analogy to the results of the previous section:

$$\begin{cases} \Omega_{\eta_i} &= (D' S'_i (S_i \Omega_w S'_i)^{-1} S_i D)^{-1}, \\ B_i &= (D' S'_i (S_i \Omega_w S'_i)^{-1} S_i D)^{-1} D' S'_i (S_i \Omega_w S'_i)^{-1}. \end{cases}$$

### A.III.3 Constrained estimator

We reconsider the uncentered version of the individual effects  $\eta_i$  in this section since the constraints apply more naturally to those. Nevertheless, we freely borrow the likelihood expressions derived in the previous section in which we considered the centered version  $\eta_i^c$ .

Using that the likelihood function  $L(\eta_i | y_i^{[1-p, T]})$  is proportional to:

$$\exp\left(-\frac{1}{2}(\eta_i - \hat{\eta}_i)' \Omega_{\eta}^{-1} (\eta_i - \hat{\eta}_i)\right) L_0(\eta_i)$$

where  $\hat{\eta}_i$  is the unconstrained estimator, we solve the following program to compute the constrained estimator of  $\eta_i$

$$\min_{\eta_i} (\eta_i - \hat{\eta}_i)' \Omega_{\eta}^{-1} (\eta_i - \hat{\eta}_i)$$

under the constraints:

$$\eta_{i2} > 0, \eta_{i3} < 0, \eta_{i3} > -\pi_T \eta_{i2}.$$

Denote  $\mu_1, \mu_2$  and  $\mu_3$  the Lagrange multipliers associated to each constraint and write the Lagrangian as:

$$L(\eta_i) = (\eta_i - \hat{\eta}_i)' \Omega_{\eta}^{-1} (\eta_i - \hat{\eta}_i) - \mu_1 \eta_{i2} + \mu_2 \eta_{i3} - \mu_3 (\eta_{i3} + \pi_T \eta_{i2}).$$

Taking derivatives yields:

$$2\Omega_{\eta}^{-1} (\tilde{\eta}_i - \hat{\eta}_i) - \begin{pmatrix} 0 \\ \mu_1 + \pi_T \mu_3 \\ \mu_3 - \mu_2 \end{pmatrix} = 0.$$

We immediately have that:

1. If  $\mu_2 > 0, \mu_1 = 0$  then  $\tilde{\eta}_{i3} = 0$  and  $\tilde{\eta}_{i2} > 0$ , and this implies that  $\pi_T \tilde{\eta}_{i2} + \tilde{\eta}_{i3} > 0$  so that  $\mu_3 = 0$ . Therefore:

$$\begin{pmatrix} \tilde{\eta}_{i1} - \hat{\eta}_{i1} \\ \tilde{\eta}_{i2} - \hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix} + \frac{\Omega_{\eta}}{2} \begin{pmatrix} 0 \\ 0 \\ \mu_2 \end{pmatrix} = 0 \implies \mu_2 e_3' \frac{\Omega_{\eta}}{2} e_3 = \hat{\eta}_{i3},$$

where  $e_3 = (0, 0, 1)'$ . This is compatible if  $\mu_2 = \frac{\hat{\eta}_{i3}}{e_3' \frac{\Omega_{\eta}}{2} e_3} > 0$  and therefore if  $\hat{\eta}_{i3} > 0$  since  $\Omega_{\eta}$  is definite positive. Denoting  $e_2 = (0, 1, 0)'$ , we also have:

$$\tilde{\eta}_{i2} - \hat{\eta}_{i2} = -\mu_2 \cdot e_2' \frac{\Omega_{\eta}}{2} e_3.$$

This satisfies the condition  $\mu_1 = 0$  iff  $\tilde{\eta}_{i2} > 0$ .

2. If  $\mu_3 > 0, \mu_1 = 0$  then  $\tilde{\eta}_{i3} = -\pi_T \tilde{\eta}_{i2}$  and  $\tilde{\eta}_{i2} > 0$ , and this implies that  $\tilde{\eta}_{i3} < 0$  so that  $\mu_2 = 0$ . We have:

$$2\Omega_\eta^{-1}(\tilde{\eta}_i - \hat{\eta}_i) - \begin{pmatrix} 0 \\ \pi_T \\ 1 \end{pmatrix} \mu_3 = 0 \implies (\tilde{\eta}_i - \hat{\eta}_i) = \mu_3 \frac{\Omega_\eta}{2} v_{\pi_T}$$

denoting  $v_\pi = (0, \pi_T, 1)'$ . Given that  $v_\pi' \tilde{\eta}_i = \tilde{\eta}_{i3} + \pi_T \tilde{\eta}_{i2} = 0$ , this implies that :

$$\mu_3 = -\frac{v_\pi' \hat{\eta}_i}{v_\pi' \frac{\Omega_\eta}{2} v_\pi} > 0,$$

if  $v_\pi' \hat{\eta}_i = \hat{\eta}_{i3} + \pi_T \hat{\eta}_{i2} < 0$ . This yields the constrained estimators,  $\tilde{\eta}_{i2}$  and  $\tilde{\eta}_{i3}$ :

$$(\tilde{\eta}_i - \hat{\eta}_i) = \mu_3 \frac{\Omega_\eta}{2} v_\pi'$$

which satisfy the constraint  $\mu_1 = 0$  iff  $\tilde{\eta}_{i2} > 0$ .

3. If  $\mu_1 > 0$  then  $\tilde{\eta}_{i2} = 0$  and thus the constraints  $\pi_T \tilde{\eta}_{i2} + \tilde{\eta}_{i3} \geq 0$  and  $\tilde{\eta}_{i3} \leq 0$  imply that  $\tilde{\eta}_{i3} = 0$ , that  $\mu_2 \mu_3 = 0$  and that one of them is positive.

Summarizing:

- If  $\hat{\eta}_{i3} < 0, \hat{\eta}_{i2} > 0$ , and  $\hat{\eta}_{i3} + \pi_T \hat{\eta}_{i2} > 0$ , constrained estimates,  $\tilde{\eta}_i$ , are equal to unconstrained estimates,  $\hat{\eta}_i$ .
- If  $\hat{\eta}_{i3} > 0, \hat{\eta}_{i3} + \pi_T \hat{\eta}_{i2} > 0$  case 1 applies if  $\tilde{\eta}_{i2} > 0$ .
- If  $\hat{\eta}_{i3} + \pi_T \hat{\eta}_{i2} < 0, \hat{\eta}_{i3} < 0$  case 2 applies if  $\tilde{\eta}_{i2} > 0$ .
- In all other cases,  $\tilde{\eta}_{i2} = \tilde{\eta}_{i3} = 0$ . In this case:

$$\tilde{\eta}_i - \hat{\eta}_i = \begin{pmatrix} \tilde{\eta}_{i1} - \hat{\eta}_{i1} \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix} = \frac{\Omega_\eta}{2} \begin{pmatrix} e_2 & e_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

where  $v_j$  are unknown. They are obtained using:

$$\begin{pmatrix} e_2' \\ e_3' \end{pmatrix} (\tilde{\eta}_i - \hat{\eta}_i) = \begin{pmatrix} e_2' \\ e_3' \end{pmatrix} \begin{pmatrix} 0 \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix} = \begin{pmatrix} e_2' \\ e_3' \end{pmatrix} \frac{\Omega_\eta}{2} \begin{pmatrix} e_2 & e_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Denoting  $I_c^\top = \begin{pmatrix} e_2' \\ e_3' \end{pmatrix}$  so that:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \left[ I_c' \frac{\Omega_\eta}{2} I_c \right]^{-1} I_c' \begin{pmatrix} 0 \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix}$$

so that we get the vector:

$$\tilde{\eta}_i - \hat{\eta}_i = \Omega_\eta I_c [I_c' \Omega_\eta I_c]^{-1} I_c' \begin{pmatrix} 0 \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix}.$$

### A.III.4 Imposing constraints on simulations

Assume that we want to impose the constraints that  $\eta_{i2} > 0$  and that  $\eta_{i3} < 0$  and  $\eta_{i3} > -\pi_T \eta_{i2}$ . Drawing in a multivariate normal distribution with multiple constraints is not as easy as with a single constraint. We use efficient Gibbs sampling as proposed by Rodriguez-Yam, Davis and Scharf (2004).

First, denote  $C_\eta$  the Choleski decomposition of the permutation of matrix  $\Omega_\eta$  (or  $\Omega_{\eta_i}$  in the case of missing values) such that:

$$C_\eta C_\eta' = \Omega_\eta.$$

Without loss of generality, it is convenient to slightly change the order of  $\eta$ s. Assuming that the generic element of the lower diagonal matrix  $C_\eta$  is  $c_{ij}$ , we can write, assuming that the expectation of  $\eta_i$  is  $(\alpha_1, \alpha_2, \alpha_3)$ :

$$\begin{cases} \eta_2 = & \alpha_2 + c_{11}\xi_1, \\ \eta_3 = & \alpha_3 + c_{21}\xi_1 + c_{22}\xi_2, \\ \eta_1 = & \alpha_1 + c_{31}\xi_1 + c_{32}\xi_2 + c_{33}\xi_3. \end{cases}$$

We start from the remark that it is easy to draw in univariate truncated normal distributions conditional to the other variates, for instance,  $f(\eta_1^u \mid \eta_2^u, \eta_3^u, \eta_2^u \leq 0, \eta_3^u \in [-\pi_T \eta_2^u, 0])$ . Second, drawing repetitively in the conditional univariate distributions to construct a Markov chain yields drawings that are distributed according to the joint distribution we are looking for. Furthermore, Rodriguez-Yam, Davis and Scharf (2004) recommends drawing the independent errors  $\xi_1, \xi_2$  and  $\xi_3$  instead of the original variables. For this, we have to rewrite the constraints as (using  $c_{11}, c_{22}$  and  $c_{33}$  are positive, see Section A.III.2):

$$\begin{aligned} \xi_1 &> -\frac{\alpha_2}{c_{11}}, \\ \xi_2 + \frac{c_{21}}{c_{22}}\xi_1 &< -\frac{\alpha_3}{c_{22}}, \\ \xi_2 + \frac{c_{21} + \pi_T c_{11}}{c_{22}}\xi_1 &> -\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}}. \end{aligned} \tag{A.III.2}$$

The algorithm proceeds by considering initial values  $(\eta_2^0, \eta_3^0)$  whose construction we detail below. Then from  $(\eta_2^k, \eta_3^k)$ , we construct  $(\eta_2^{k+1}, \eta_3^{k+1})$  using:

1. Draw  $\xi_2^{k+1}$  in a truncated normal variable, truncated by the bounds  $[-\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}} - \frac{c_{21} + \pi_T c_{11}}{c_{22}}\xi_1^k, -\frac{\alpha_3}{c_{22}} - \frac{c_{21}}{c_{22}}\xi_1^k]$  (a non empty interval because of the constraint  $\xi_1 > -\frac{\alpha_2}{c_{11}}$ ).
2. Draw  $\xi_1^{k+1}$  in a truncated normal variable, truncated by the bounds  $[L_1, L_2]$ . There are five cases:

- If  $c_{21} > 0$ :  $L_1 = \max(-\frac{\alpha_2}{c_{11}}, -\frac{c_{22}}{c_{21} + \pi_T c_{11}}(\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}} + \xi_2^{k+1}))$ ;  $U_1 = -\frac{c_{22}}{c_{21}}(\frac{\alpha_3}{c_{22}} + \xi_2^{k+1})$
- If  $c_{21} = 0$ :  $L_1 = \max(-\frac{\alpha_2}{c_{11}}, -\frac{c_{22}}{c_{21} + \pi_T c_{11}}(\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}} + \xi_2^{k+1}))$ ,  $U_1 = +\infty$
- If  $c_{21} \in (-\pi_T c_{11}, 0)$ :  $L_1 = \max(-\frac{\alpha_2}{c_{11}}, -\frac{c_{22}}{c_{21}}(\frac{\alpha_3}{c_{22}} + \xi_2^{k+1}), -\frac{c_{22}}{c_{21} + \pi_T c_{11}}(\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}} + \xi_2^{k+1}))$ ,  $U_1 = +\infty$

- If  $c_{21} = -\pi_T c_{11}$  :  $L_1 = \max(-\frac{\alpha_2}{c_{11}}, -\frac{c_{22}}{c_{21}}(\frac{\alpha_3}{c_{22}} + \xi_2^{k+1}))$ ,  $U_1 = +\infty$
- If  $c_{21} < -\pi_T c_{11}$  :  $L_1 = \max(-\frac{\alpha_2}{c_{11}}, -\frac{c_{22}}{c_{21}}(\frac{\alpha_3}{c_{22}} + \xi_2^{k+1}))$ ,  $U_1 = -\frac{c_{22}}{c_{21} + \pi_T c_{11}}(\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}} + \xi_2^{k+1})$ .

Then construct .

When the algorithm is said to have converged to  $(\xi_1^\infty, \xi_2^\infty)$  then finish by drawing  $\xi_3$  in a  $N(0,1)$  variate since no constraints are binding on  $\eta_1$ . Construct the final values  $\eta_2^{k+1} = \alpha_2 + c_{11}\xi_1^\infty$ ,  $\eta_3^{k+1} = \alpha_3 + c_{21}\xi_1^\infty + c_{22}\xi_2^\infty$ ,  $\eta_1^{k+1} = \alpha_1 + c_{31}\xi_1^\infty + c_{32}\xi_2^\infty + c_{33}\xi_3$ .

The initial conditions are constructed by neglecting the multivariate aspects of constraints:

- Draw  $\xi_1^0$  in a truncated normal distribution, truncated by the bound  $\xi_1^0 > -\frac{\alpha_2}{c_{11}}$ . Construct  $\eta_2^0 = \alpha_2 + c_{11}\xi_1^0$ .
- Draw  $\xi_2^0$  in a truncated normal distribution, truncated by the bound  $[-\frac{\alpha_3 + \pi_T \alpha_2}{c_{22}} - \frac{c_{21} + \pi_T c_{11}}{c_{22}}\xi_1^0, -\frac{\alpha_3}{c_{22}} - \frac{c_{21}}{c_{22}}\xi_1^0]$ . Construct  $\eta_3^0 = \alpha_3 + c_{21}\xi_1^0 + c_{22}\xi_2^0$ .
- Draw  $\xi_3^0$  in a normal distribution and construct  $\eta_1^0 = \alpha_1 + c_{31}\xi_1^0 + c_{32}\xi_2^0 + c_{33}\xi_3^0$ .

These draws satisfy the constraints but they are not truncated normally distributed.

## A.IV Data and Empirical Appendix

### A.IV.1 Sample selection and the construction of earnings: Comparison with the literature

The data used in this paper differ in several respects from those used in other studies on earnings dynamics. We summarize the main differences with Guvenen (2009) and Hryshko (2012) two recent papers from this literature.

First, datasets have very different characteristics. The two US studies use the PSID, a household panel data, while this paper makes use of DADS, a French administrative panel dataset collected by tax authorities in the private sector only. The PSID household survey includes a large set of demographic characteristics while the French DADS administrative data have larger samples but fewer observable characteristics, namely age at entry on the labor market and a proxy for skills in three groups. The length of the observation period is broadly similar in these three studies. The US studies use waves from 1968 to 1993 in Guvenen (2009) and from 1968 to 1997 in Hryshko (2012), representing a maximum of 25 to 30 years of data. Our study on DADS covers a more recent time period ranging from 1977 to 2007 with 27 yearly effective observations.

As regards the sample selection rules, the two US studies concentrate on male heads of households while DADS data does not report family structure. This is why we focus on males irrespective of their family situation. Guvenen (2009) restricts his sample to individuals between



the ages of 20 and 64 and observed at least 20 times and with positive labor earnings hours. Hryskho (2012) focuses on males in the 25 to 64 age range. An important difference between US studies and the DADS one is the particular definition we use to define a cohort. We define a cohort by the set of individuals entering the labor market at the same point in time but not necessarily at the same age, while we control by age in the earnings equation. Therefore, the cohorts are not defined as usually by the date of birth. It is difficult to assess the consequence of this choice since we are unaware of another study that would compare earnings dynamics using different definitions for the cohorts. However in sections A.IV.2 and A.IV.3 we demonstrate that the 1977 cohort used in the paper does not depart significantly from younger ones extracted from the same dataset in terms of means, variances and autocorrelations of order one.

To filter out extreme values, Guvenen (2009) sets criteria to discard observations with low and high hourly earnings and hours worked. In a similar spirit, this study drops observations in the first and last percentiles of the yearly earnings distribution. Hryskho (2012) use alternative selection rules dropping observations for the years when the percentage change of real labor income in adjacent years is above 500 or below  $-80$ . More importantly, Hryskho (2012) focuses on consecutive spells of positive incomes with at least 9 observations while the DADS study allows any time pattern of missing observations as long as individuals are observed working in the private sector in selected years (ie : in 1977, 1978, 1982 and 1984). Sample size is smaller in PSID data. Guvenen (2009) main sample includes 1270 individuals, while Hryskho (2012) includes 1916 heads. The DADS sample starts with 7446 observations in the initial year and finishes with 4670 in the last year of data.

Finally, the earnings measure is somewhat different between the two US studies and ours. First, Guvenen (2009) defines labor income as including wage income, bonuses, commissions, plus the labor portions of several types of income such as farm income, business income and therefore keeping self-employed or civil servants. Hryskho (2012) introduces a similar definition but excludes self-employment. In this paper we use annualized earnings defined by full earnings divided by the number of days worked and remultiplied by 360.

## A.IV.2 Comparison with younger cohorts

Figure A.i compares the time profile for the mean and the variance of log earnings for our cohort entering the labor market in 1977 with cohorts entering the labor market between 1980 and 2000 and observed up to 2007. We cannot compare the 1977 cohort with older ones (except 1976) since the DADS started in 1976. The left panel displays the change over time in mean log earnings for the various entry cohorts, while the right panel depicts the change in the variances over time by entry cohort. First, mean log earnings display the usual increasing concave profile, with a very homogeneous pattern across entry cohorts. The mean log earnings of the different entry cohorts converge to very comparable level over time. Second, comparing the variance profiles across entry cohorts in the right panel, we see that the decrease in the variance over the

first few years on the labor market is more pronounced for the entry cohort 1977 than for most other entry cohorts. The sharp drop in the variance is present for cohorts entering in 1977, 1980, 1988 and 2000, but it is weaker for other entry cohorts. The variance profile of the first few years on the labor market can be quite heterogeneous across entry cohorts but we control for this with very flexible initial conditions in our earnings equation. After a few years on the labor market the variance for the different entry cohorts converge to very comparable levels except for more recent entry cohorts with a lower mean inequality level for entry cohorts 1997 and 2000 but they are still in the first part of their profile though. Strikingly, in 1994, the different entry cohorts display the same increase in the variance of log earnings making it clear that this survey year is affected by measurement error. Overall, the different profiles are similar and the entry cohort 1977 studied in the paper does not depart significantly from the other ones.

### **A.IV.3 Comparison of first order autocorrelation between cohorts**

In Figure A.ii we draw the first order autocorrelation in earnings residuals relative to potential experience for three different entry cohorts. We compare our cohort entering in 1977 with younger ones entering in 1987 and in 1997. The first group is observed over thirty years from 1977 to 2007, the second one over twenty years from 1987 to 2007 and the third entry cohort is observed over ten years from 1997 to 2007. As the Figure illustrates, the pattern of declining autocorrelation is clearly similar across the three entry cohorts. Starting at one it diminishes up to 0.6 after fifteen years on the labor market for the 1977 and 1987 entry cohorts. For the 1997 entry cohort it declines only to 0.8 but with only ten years of data the trend and level remain very similar to previous entry cohorts.

### **A.IV.4 Attrition**

Table A.i gives a dynamic view of attrition. This Table reports the frequencies of non missing values by pairs of years. For instance, the column 1977, describes the global features of attrition. Attrition is quite severe in the first "normal" (after selection) year, 1985 since 15% of individuals exit between 1984 and 1985. This is true in every adjacent years at the beginning of the sample period (other columns for instance in cell 1987, 1988) but it is decreasing over time to reach 7 or 8% at the end of the panel. Year 1994 confirms its exceptional status as attrition between 1994 and 1995 is very low. More generally though, most individuals reenter the panel quickly since the attrition at two year intervals is only marginally larger than the one observed at one year intervals (for instance the two cells in 1977, 1985 and 1986, indicate attrition of 15% and 16.5%) although this varies somewhat over time. Finally, there is a core of observations which are almost always present in the panel. Looking at row 2007, we can see that out of the 62.7% of the complete sample of individuals present in this year, it is hardly less than 80% of this sample which is not present between 1985 and 2006 – with the exception of 1994 again.

Table A.ii reports the autocorrelation patterns of the first differences in the earnings residuals. Contrary to what is found in some papers in the literature using PSID data (for instance, Meghir and Pistaferri, 2010) we do not find strong evidence that the correlation disappears after taking a two period difference. A few very long difference autocorrelations seem significant and no regular pattern seems to emerge.

#### A.IV.5 Random effect estimation and reduced form parameters

Firstly, we estimate covariance matrices of the permanent and stochastic components of errors as well as their correlation with the initial conditions. The former is composed by three individual unobserved factors  $(\eta_{i1}^c, \eta_{i2}^c, \eta_{i3}^c)$ , while the latter is an ARMA process as explained in Section A.II. Table A.iii provides the values of the Akaike criterion based on the log-likelihood values for specifications in which orders of the autoregressive and moving average components vary from (1,1) to (3,3). Unsurprisingly, enlarging the number of AR or MA components strongly increases the value of the sample likelihood function. Nonetheless, increasing it beyond 3 lags is difficult to implement since it involves a year, 1981, in which observations are missing altogether. This is why we did not pursue further the exploration of higher orders for the ARMA processes. According to the Akaike criterion we should choose the ARMA(3,3) specification, a much more persistent specification than in most studies in the literature. Nevertheless, the estimates of the ARMA(3,3) exhibit some estimates which are very imprecise, specifically the ones describing the correlations between initial conditions and the MA components (Table A.iv). That is why in the rest of the analysis we will use as our preferred estimates, results from the ARMA(3,1) model.

Table A.iv presents parameter estimates. Each column reports results for different ARMA(p,q) specifications for  $(p, q) \in \{1, 2, 3\}^2$ . In every model, autoregressive coefficients remain largely lower than one. Their sum reflects the high persistence of shocks though it is far enough from one to reject a unit root. A formal statistical test rejects at levels less than 1% that the process is non-stationary (see Magnac and Roux, 2009). This result parallels the result of Alvarez and Arellano (2004) on US and Spanish data or of Guvenen (2009) but differs from Hryshko (2012). Autoregressive coefficients are ranging from 0.2 to 0.02 in the ARMA(3,1) specification and describe the persistence of shocks due to unemployment spells or mobility for instance while the MA coefficient is negative and might stand for measurement errors.

The correlations between initial conditions and these individual factor loadings are also informative. They are significant and have an economically significant magnitude of around .2 or .3 in absolute value. The estimated correlations between the linear trend and geometric factors  $\eta_2$  and  $\eta_3$ , and the initial conditions are similar to the estimated correlations between both of them and the level factor. They are respectively significantly positive and negative. More surprisingly, the correlation between  $\eta_1$  and the initial conditions is also negative. That would indicate that individuals endowed with higher starting human capital stock have more difficulties to acquire

immediately the level of potential earnings that correspond to their skill levels.

Finally, the estimated variance of the idiosyncratic terms is reported in Table A.v. Note first that these parameters are identified even in years 1981, 1983 and 1990 for which information is missing. Nonetheless, estimates for those years are imprecise and have a magnitude that can differ widely from the others and across ARMA specifications because they are identified only out of the structural restrictions that we placed by assuming an ARMA process. Regarding the "normal" years, period-specific variances start from a rather high level in the first three years between .20 and .30. They generally decrease over the sample periods albeit very slowly. Between 1984 and 2000 they are quite precisely estimated at a level around .18, except the exceptional year 1994 in which we know that the measurement error is large, and levels off at around .14 after 2000 (except the exceptional year 2003). These estimates certainly pick up the patterns of autocorrelations increasing over time that we spotted in the raw data (see Table 2). Part of it is certainly attributable to measurement errors although another part of it could be attributed to a decreasing impact of shocks along the life cycle.

#### **A.IV.6 Constrained and structural estimates**

To confirm the diagnostic of fat tails of fixed effect estimates, we also evaluate the distribution of the QLR statistic in a different way. We computed the distance between the unconstrained and the constrained estimates and compare this distance with the distance between the same constrained estimates and simulated unconstrained estimates using normal random draws for the simulations. In all these experiments, we use the covariance matrix of the  $\eta$ s as a weighting matrix to compute the distance and as the basis for simulating the normal errors. Table A.vii reports the quantiles of the distributions of the actual and simulated distances. The two distributions coincide rather well for all quantiles until 0.6 but the divergence becomes severe over .6 and specifically at the upper end. This can either be due to the rejection of the constraints or to the non normality of the factors which is a standard finding in studies that assess the normality of individual effects in earnings functions (Hirano, 2002 for instance).

#### **A.IV.7 Non-linear dynamics of earnings: transition matrices across quintiles**

Table A.ix reports the observed and predicted (using both unconstrained and simulated constrained estimates) transition matrices between quintiles, between the entry in the labor market (1977) and 15 years later (1992). Between these two dates, the model predicts more mobility than the actual data, since the diagonal of the observed transition matrix is for every quintile higher than the predicted ones (either unconstrained or simulated). The discrepancy is maximal for the highest quintile since the probability of staying in the highest quintile is observed to be equal to 0.5 although its prediction is equal to 0.3.

The fit of the model is better when considering the mobility between quintiles between 1993 and 2007, as reported in table A.x, in particular between the transitions based on the simulated parameters and the observed ones. Yet, the model predicts more mobility than what is truly observed (see the diagonal terms corresponding to the 2nd, 3rd and 4th quintiles). The fit is very good for the transitions related to the 1st and 5th quintiles since the prediction error is less than two percentage points.

#### **A.IV.8 Estimation of human capital prices by a flat spot condition and robustness checks**

We follow Bowlus and Robinson (2012). From the DADS, average log daily real earnings by age and year can be computed on full-time males employees in the Private Sector from 1976 to 2010. To identify the “flat spot” region where human capital remains stable, we run regressions of the average log daily real earnings on potential experience (difference between current age and 16), an exponential term reflecting the curvature of the earnings profile with respect to potential experience, and year dummies. We have run different regression changing the contributing individuals with respect to their potential experience and selected the population with the broader range of potential experience for which the coefficients on potential experience and the exponential term were statistically non significant. This leads us to select individuals who are aged between 43 and 58 whose average log-earnings profile did not exhibit any slope or curvature. The results of the regressions and the human capital prices values are available upon request.

We then repeat the procedures that lead to the construction of Tables 4, 7, 8 and 9. They are reproduced in Tables A.xi; A.xii, A.xiii and A.xiv.

Table A.i: Non Missing Values

	1977	1979	1980	1985	1986	1987	1988	1989	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	
1977	1																								
1978	1																								
1979	.882	.882																							
1980	.868	.786	.868																						
1982	1	.882	.868																						
1984	1	.882	.868																						
1985	.849	.751	.743	.849																					
1986	.834	.739	.731	.75	.834																				
1987	.804	.714	.704	.718	.737	.804																			
1988	.765	.675	.668	.694	.690	.691	.765																		
1989	.777	.689	.677	.701	.694	.691	.689	.777																	
1991	.743	.658	.65	.67	.663	.655	.649	.678	.743																
1992	.736	.653	.647	.663	.655	.649	.642	.662	.679	.736															
1993	.749	.665	.653	.657	.666	.654	.631	.652	.659	.673	.749														
1994	.581	.515	.506	.508	.518	.511	.492	.506	.513	.517	.544	.581													
1995	.725	.643	.634	.636	.644	.632	.609	.628	.63	.635	.661	.535	.725												
1996	.721	.641	.631	.631	.638	.627	.603	.622	.622	.627	.652	.521	.671	.721											
1997	.71	.629	.621	.622	.63	.619	.596	.613	.612	.618	.642	.511	.649	.661	.71										
1998	.708	.628	.619	.618	.625	.615	.591	.61	.609	.614	.636	.506	.642	.649	.667	.708									
1999	.708	.628	.617	.617	.623	.614	.59	.61	.605	.609	.63	.502	.635	.639	.652	.665	.708								
2000	.701	.622	.611	.612	.62	.61	.583	.6	.595	.601	.623	.497	.625	.629	.637	.649	.662	.701							
2001	.687	.61	.598	.599	.605	.595	.573	.589	.584	.587	.605	.479	.608	.612	.62	.629	.639	.65	.687						
2002	.67	.595	.586	.588	.591	.581	.559	.575	.568	.573	.592	.471	.59	.594	.597	.606	.613	.617	.621	.67					
2003	.616	.547	.539	.544	.542	.532	.516	.533	.526	.53	.539	.425	.538	.541	.546	.553	.561	.564	.563	.577	.616				
2004	.63	.559	.551	.552	.556	.545	.523	.541	.534	.539	.555	.441	.555	.557	.559	.567	.573	.574	.574	.584	.565	.63			
2005	.634	.560	.552	.554	.558	.548	.526	.544	.536	.541	.558	.446	.557	.558	.559	.566	.570	.574	.571	.574	.543	.574	.634		
2006	.634	.561	.553	.556	.557	.549	.525	.544	.535	.541	.556	.444	.553	.556	.557	.563	.568	.570	.567	.574	.538	.566	.586	.634	
2007	.627	.557	.547	.55	.552	.542	.521	.538	.531	.535	.548	.436	.547	.549	.551	.556	.560	.562	.557	.561	.525	.552	.570	.591	

Notes: Frequencies of observations present in the sample at years described by row and column, relative to the full sample

Table A.ii: Autocorrelation matrix of earnings residuals in differences

	1978	1979	1980	1985	1986	1987	1988	1989	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	
1979	-.400																							
1980	-.009	-.277																						
1985	-.018	-.016	-.084																					
1986	.003	-.031	.090	-.434																				
1987	.043	.093	-.013	-.058	-.345																			
1988	.004	.035	.011	-.055	-.046	-.299																		
1989	.041	-.036	-.008	.028	-.054	-.020	-.323																	
1992	-.053	.060	-.055	-.014	-.006	-.074	-.003	-.039																
1993	-.021	.015	-.019	.007	.013	.048	-.072	.000	-.351															
1994	.018	-.013	.024	.000	-.021	-.013	.003	-.037	-.108	-.385														
1995	.021	-.001	.017	-.027	.029	.038	.001	.032	.043	-.070	-.519													
1996	.012	-.013	-.034	.008	-.020	.000	.036	.046	.029	-.021	.026	-.440												
1997	-.052	.032	-.047	.026	-.046	.006	-.022	-.058	-.007	-.005	-.004	-.019	-.520											
1998	.010	-.010	.052	-.047	.049	-.040	.004	.000	.009	.015	-.031	.036	-.015	-.391										
1999	.056	-.017	-.017	.013	.006	-.013	.040	-.004	.014	-.067	.004	-.020	.003	-.010	-.244									
2000	-.087	.085	-.059	.008	.016	-.014	-.006	-.023	.041	.023	.005	-.042	.022	-.003	-.047	-.420								
2001	.024	-.051	.051	.009	-.082	.044	-.028	.052	-.046	-.018	.032	-.009	-.062	.051	.044	-.013	-.539							
2002	.008	.001	-.037	.027	.010	-.090	.046	-.025	-.019	-.002	-.043	.013	.031	.024	-.028	.005	-.010	-.298						
2003	.005	-.050	.001	.041	-.040	-.108	.001	-.015	.061	-.028	.062	-.025	-.049	.052	-.006	.025	.027	-.010	-.247					
2004	-.036	.068	.008	-.061	.057	.144	-.005	.004	-.047	.013	-.031	.012	.025	-.043	.005	-.024	-.025	.014	-.157	-.705				
2005	.073	-.011	.001	-.021	-.017	.026	-.011	-.010	-.019	.020	.005	.004	-.001	-.009	-.013	.056	.014	-.043	.002	.012	-.227			
2006	-.031	.063	-.042	.009	.035	-.025	.021	-.031	.055	-.014	.034	-.023	-.002	-.031	-.013	-.015	.013	-.028	-.002	.039	-.069	-.375		
2007	-.002	-.022	-.010	-.042	-.003	-.026	.026	-.036	-.016	.079	-.070	.022	.015	-.015	-.035	.035	-.015	.020	.030	-.028	-.006	.053	-.254	

Table A.iii: AIC criterion

ARMA(p,q)	q=1	q=2	q=3
p=1	-344885 (43)	-344899 (45)	-344906 (47)
p=2	-345301 (47)	-345447 (50)	-345733 (53)
p=3	-345839 (51)	-346133 (54)	-346293 (58)

AIC criterion computed as  $-2\log(L) + 2K$ , with  $L$  the likelihood and  $K$  the number of parameters. Number of parameters in brackets.



Table A.iv: Estimated parameters of the Random Effects Model

	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
$\alpha_1$	.702 (.005)	.729 (.006)	.711 (.007)	.263 (.011)	.186 (.011)	.220 (.011)	.200 (.012)	.203 (.011)	.194 (.011)
$\alpha_2$				.145 (.004)	.324 (.008)	.143 (.009)	.191 (.005)	.143 (.009)	.161 (.009)
$\alpha_3$							.022 (.003)	.087 (.004)	.187 (.008)
$\psi_1$	.369 (.005)	.391 (.005)	.373 (.007)	-.091 (.011)	-.172 (.011)	-.135 (.012)	-.164 (.012)	-.166 (.011)	-.189 (.011)
$\psi_2$		.020 (.003)	.017 (.003)		.170 (.006)	-.028 (.008)		-.046 (.008)	-.046 (.008)
$\psi_3$			-.012 (.004)			-.080 (.004)			.114 (.007)
$\sigma_{\eta_1}$	.302 (.001)	.302 (.003)	.301 (.003)	.310 (.003)	.306 (.003)	.304 (.003)	.306 (.003)	.300 (.003)	.298 (.004)
$\sigma_{\eta_2}$	.038 (.005)	.039 (.001)	.039 (.001)	.038 (.001)	.039 (.001)	.036 (.001)	.038 (.001)	.037 (.001)	.037 (.001)
$\sigma_{\eta_3}$	.255 (.005)	.259 (.006)	.256 (.006)	.263 (.004)	.260 (.005)	.248 (.005)	.258 (.005)	.247 (.006)	.242 (.007)
$\rho_{\eta_1, \eta_2}$	.473 (.016)	.413 (.021)	.454 (.021)	.571 (.013)	.486 (.017)	.610 (.013)	.505 (.017)	.485 (.020)	.365 (.030)
$\rho_{\eta_1, \eta_3}$	-.604 (.003)	-.548 (.020)	-.586 (.019)	-.694 (.011)	-.618 (.015)	-.729 (.012)	-.636 (.016)	-.620 (.019)	-.509 (.029)
$\rho_{\eta_2, \eta_3}$	-.946 (.023)	-.948 (.003)	-.947 (.003)	-.945 (.002)	-.946 (.002)	-.941 (.003)	-.946 (.002)	-.943 (.003)	-.944 (.004)
$\sigma_{y_0}$	.491 (.000)	.506 (.007)	.496 (.007)	.448 (.004)	.479 (.005)	.429 (.004)	.442 (.004)	.455 (.005)	.494 (.008)
$\sigma_{y_{-1}}$				.381 (.004)	.424 (.005)	.359 (.004)	.387 (.004)	.386 (.005)	.428 (.008)
$\sigma_{y_{-2}}$							.264 (.004)	.270 (.006)	.299 (.008)
$cov(\eta_1, y_0)$	-.227 (.019)	-.257 (.017)	-.237 (.017)	-.156 (.015)	-.214 (.016)	-.149 (.016)	-.186 (.016)	-.201 (.017)	-.282 (.019)
$cov(\eta_1, y_{-1})$				-.127 (.016)	-.183 (.017)	-.113 (.017)	-.153 (.017)	-.168 (.018)	-.253 (.020)
$cov(\eta_1, y_{-2})$							-.169 (.018)	-.185 (.019)	-.267 (.022)
$cov(\eta_2, y_0)$	.358 (.022)	.402 (.020)	.374 (.021)	.232 (.017)	.335 (.019)	.155 (.021)	.219 (.020)	.253 (.022)	.361 (.026)
$cov(\eta_2, y_{-1})$				.218 (.019)	.331 (.021)	.119 (.024)	.242 (.022)	.235 (.025)	.352 (.029)
$cov(\eta_2, y_{-2})$							.239 (.024)	.253 (.027)	.351 (.032)
$cov(\eta_3, y_0)$	-.290 (.018)	-.333 (.023)	-.305 (.023)	-.179 (.020)	-.270 (.022)	-.107 (.023)	-.163 (.023)	-.195 (.024)	-.291 (.029)
$cov(\eta_3, y_{-1})$				-.169 (.021)	-.272 (.023)	-.077 (.025)	-.190 (.023)	-.181 (.027)	-.287 (.032)
$cov(\eta_3, y_{-2})$							-.181 (.026)	-.194 (.029)	-.282 (.035)
$cov(y_0, \zeta_0)$	.809 (.023)	.036 (8.525)	-.024 (26.529)	-.823 (.269)	.826 (.059)	-.931 (.207)	.841 (.061)	-.795 (.416)	.812 (.096)
$cov(y_0, \zeta_{-1})$		.779 (.438)	-.012 (1.245)		.408 (.102)	-.352 (17.542)		-.208 (152.666)	.361 (31.114)
$cov(y_{-1}, \zeta_{-1})$			.798 (.813)		.722 (.062)	-.066 (.148)		.830 (41.955)	.234 (17.858)
$cov(y_0, \zeta_{-2})$						-.805 (3.931)			-.719 (76.705)
$cov(y_{-1}, \zeta_{-2})$						-.382 (11.249)			-.202 (44.061)
$cov(y_{-2}, \zeta_{-2})$									.752 (.094)

Table A.v: Yearly standard deviation of earnings

	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
1978	.311 (.001)	.312 (.002)	.312 (.002)						
1979	.254 (.001)	.257 (.001)	.255 (.001)	.222 (.001)	.232 (.001)	.219 (.001)			
1980	.223 (.005)	.223 (.001)	.223 (.001)	.222 (.001)	.227 (.001)	.221 (.001)	.224 (.002)	.224 (.002)	.230 (.002)
1981	.264 (.005)	.260 (.005)	.263 (.005)	.000 (.096)	.103 (.040)	.002 (.066)	.004 (.082)	.006 (.076)	.001 (.060)
1982	.152 (.005)	.150 (.005)	.150 (.005)	.194 (.002)	.193 (.002)	.197 (.002)	.193 (.002)	.195 (.002)	.198 (.002)
1983	.244 (.004)	.243 (.005)	.247 (.005)	.040 (.063)	.175 (.017)	.096 (.037)	.023 (.048)	.039 (.049)	.193 (.021)
1984	.154 (.001)	.149 (.004)	.149 (.004)	.189 (.002)	.184 (.001)	.187 (.002)	.188 (.001)	.188 (.001)	.182 (.002)
1985	.182 (.001)	.182 (.001)	.182 (.001)	.181 (.001)	.183 (.001)	.183 (.001)	.181 (.001)	.183 (.001)	.183 (.001)
1986	.187 (.001)	.187 (.001)	.187 (.001)	.189 (.001)	.189 (.001)	.190 (.001)	.190 (.001)	.190 (.001)	.192 (.001)
1987	.181 (.001)	.182 (.001)	.181 (.001)	.176 (.001)	.176 (.001)	.177 (.001)	.176 (.001)	.177 (.001)	.177 (.001)
1988	.180 (.001)	.180 (.001)	.181 (.001)	.181 (.001)	.181 (.001)	.181 (.001)	.181 (.001)	.182 (.001)	.183 (.001)
1989	.171 (.008)	.172 (.001)	.172 (.001)	.168 (.001)	.170 (.001)	.169 (.001)	.169 (.001)	.170 (.001)	.171 (.001)
1990	.012 (.002)	.021 (.007)	.005 (.008)	.358 (.012)	.303 (.008)	.375 (.015)	.349 (.012)	.395 (.016)	.363 (.013)
1991	.182 (.001)	.184 (.002)	.180 (.002)	.153 (.002)	.167 (.001)	.156 (.002)	.161 (.001)	.157 (.002)	.163 (.001)
1992	.162 (.001)	.162 (.001)	.162 (.001)	.159 (.001)	.155 (.001)	.159 (.001)	.157 (.001)	.160 (.001)	.161 (.001)
1993	.207 (.001)	.207 (.001)	.207 (.001)	.209 (.001)	.209 (.001)	.209 (.001)	.210 (.001)	.209 (.001)	.211 (.001)
1994	.237 (.001)	.236 (.001)	.237 (.001)	.250 (.001)	.250 (.001)	.251 (.001)	.252 (.001)	.253 (.001)	.254 (.001)
1995	.193 (.001)	.195 (.001)	.194 (.001)	.177 (.001)	.179 (.001)	.177 (.001)	.177 (.001)	.178 (.001)	.180 (.001)
1996	.177 (.001)	.177 (.001)	.177 (.001)	.176 (.001)	.178 (.001)	.177 (.001)	.177 (.001)	.177 (.001)	.178 (.001)
1997	.167 (.001)	.167 (.001)	.167 (.001)	.162 (.001)	.162 (.001)	.162 (.001)	.162 (.001)	.162 (.001)	.164 (.001)
1998	.137 (.001)	.138 (.001)	.138 (.001)	.134 (.001)	.137 (.001)	.135 (.001)	.135 (.001)	.136 (.001)	.138 (.001)
1999	.152 (.001)	.152 (.001)	.152 (.001)	.155 (.000)	.157 (.000)	.157 (.000)	.156 (.000)	.157 (.000)	.158 (.001)
2000	.159 (.001)	.159 (.001)	.159 (.001)	.159 (.001)	.159 (.001)	.159 (.001)	.159 (.001)	.159 (.001)	.160 (.001)
2001	.158 (.001)	.158 (.001)	.158 (.001)	.159 (.001)	.159 (.001)	.160 (.001)	.159 (.001)	.160 (.001)	.161 (.001)
2002	.153 (.001)	.153 (.001)	.153 (.001)	.146 (.001)	.146 (.001)	.146 (.001)	.146 (.001)	.147 (.001)	.149 (.001)
2003	.168 (.001)	.167 (.001)	.168 (.001)	.178 (.001)	.178 (.001)	.179 (.001)	.179 (.001)	.180 (.001)	.181 (.001)
2004	.147 (.001)	.148 (.001)	.148 (.001)	.133 (.001)	.133 (.001)	.134 (.001)	.133 (.001)	.134 (.001)	.135 (.001)
2005	.128 (.001)	.128 (.001)	.128 (.001)	.130 (.001)	.132 (.001)	.130 (.001)	.131 (.001)	.131 (.001)	.133 (.001)
2006	.123 (.001)	.124 (.001)	.123 (.001)	.124 (.000)	.124 (.000)	.124 (.000)	.125 (.000)	.125 (.000)	.127 (.000)
2007	.117 (.003)	.117 (.001)	.117 (.001)	.115 (.001)	.116 (.001)	.116 (.001)	.115 (.001)	.117 (.001)	.118 (.001)

Table A.vi: Estimates of the covariance of individual effects: Bias-corrected

Sample periods	$Var(\eta_1)$	$Cov(\eta_1, \eta_2)$	$Cov(\eta_1, \eta_3)$	$Var(\eta_2)$	$Cov(\eta_2, \eta_3)$	$Var(\eta_3)$
(3,15]	2.5 (11)	0.2 (0.89)	-2.7 (12)	0.024 (0.077)	-0.25 (1)	3.1 (13)
(15,22]	0.31 (0.13)	0.031 (0.016)	-0.31 (0.16)	0.005 (0.0023)	-0.043 (0.022)	0.41 (0.22)
(22,26]	0.1 (0.014)	0.0076 (0.0021)	-0.065 (0.017)	0.0018 (0.00051)	-0.013 (0.0038)	0.096 (0.029)
(26,28]	0.047 (0.0072)	0.0021 (0.001)	-0.016 (0.0073)	0.00029 (0.00025)	-0.001 (0.0017)	0.0043 (0.012)
Complete sample	0.65 (2.3)	0.053 (0.19)	-0.67 (2.6)	0.0069 (0.017)	-0.068 (0.22)	0.78 (2.8)
Random effects	0.093 (0.0035)	0.0059 (0.00053)	-0.05 (0.0041)	0.0015 (1e-04)	-0.0093 (0.00075)	0.066 (0.0058)

Notes: See Table above

Table A.vii: Distances between unconstrained and constrained estimates for observations and simulations

Quantiles	Observed distance	Simulated distance
0.175	0	0
0.225	0.0021	0.00180
0.275	0.0141	0.0132
0.325	0.0370	0.0391
0.375	0.0763	0.0761
0.425	0.126	0.125
0.475	0.194	0.194
0.525	0.276	0.282
0.575	0.401	0.395
0.625	0.568	0.531
0.675	0.763	0.714
0.725	1.04	0.945
0.775	1.48	1.21
0.825	2.14	1.57
0.875	3.17	2.10
0.925	5.32	2.93
0.975	12.7	4.74

Notes: Distances use as a metric the inverse covariance matrix of  $\eta$ s. Simulations are performed by adding to the constrained estimates a normal noise and by reprojecting on the constrained set.

Table A.viii: Quantiles of individual effects  $\eta_i$ : simulated constrained estimates

Individual effects	Sample periods	Quantiles						
		0.05	0.2	0.35	0.5	0.65	0.8	0.95
$\eta_1$	(3,15]	1.58	1.96	2.1	2.23	2.35	2.52	2.98
		(0.811)	(0.0365)	(0.0263)	(0.0234)	(0.0204)	(0.0197)	(0.0481)
	(15,22]	1.94	2.16	2.3	2.4	2.53	2.72	3.18
		(0.0288)	(0.0204)	(0.0177)	(0.0118)	(0.012)	(0.016)	(0.0384)
	(22,26]	1.99	2.22	2.35	2.46	2.57	2.75	3.21
		(0.0193)	(0.0151)	(0.0126)	(0.0125)	(0.0106)	(0.0116)	(0.0218)
	(26,28]	2.11	2.34	2.45	2.54	2.65	2.81	3.18
		(0.0189)	(0.0155)	(0.0123)	(0.0119)	(0.0099)	(0.0113)	(0.0178)
$\eta_2$	(3,15]	0.0161	0.0393	0.0598	0.0786	0.103	0.136	0.241
		(0.0442)	(0.059)	(0.0659)	(0.0762)	(0.0897)	(0.116)	(0.465)
	(15,22]	0.00891	0.0207	0.0294	0.038	0.0475	0.0615	0.103
		(0.0223)	(0.0277)	(0.0295)	(0.0305)	(0.0337)	(0.0358)	(0.0513)
	(22,26]	0.00832	0.0195	0.0276	0.0353	0.0435	0.056	0.0969
		(0.0208)	(0.0271)	(0.0294)	(0.03)	(0.0312)	(0.0316)	(0.0396)
	(26,28]	0.00868	0.0213	0.0282	0.0345	0.0425	0.0537	0.0852
		(0.0251)	(0.0291)	(0.0269)	(0.0287)	(0.0302)	(0.0305)	(0.0386)
$\eta_3$	(3,15]	-0.683	-0.386	-0.282	-0.21	-0.157	-0.101	-0.0411
		(0.682)	(0.198)	(0.149)	(0.122)	(0.103)	(0.0961)	(0.0694)
	(15,22]	-0.348	-0.182	-0.137	-0.109	-0.0837	-0.0589	-0.0244
		(0.114)	(0.0663)	(0.0614)	(0.0579)	(0.051)	(0.0463)	(0.0402)
	(22,26]	-0.355	-0.184	-0.135	-0.103	-0.0787	-0.0554	-0.0253
		(0.0918)	(0.0737)	(0.0611)	(0.0558)	(0.0494)	(0.0443)	(0.0366)
	(26,28]	-0.328	-0.192	-0.137	-0.102	-0.0788	-0.0557	-0.025
		(0.0801)	(0.0714)	(0.0635)	(0.0581)	(0.0579)	(0.0541)	(0.0418)

Notes: Sample period: Number of observed periods. Standard errors (sampling and parameter uncertainty, 1000 MC simulations) in brackets.

Table A.ix: Transitions  $t = 1$  to  $t = 15$ 

Quintiles in t=15	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	35.3	25.1	17.9	13.7	7.96
in t=1	(24.1)	(22.5)	(19.7)	(18.2)	(15.6)
	[26]	[24.1]	[17.7]	[18.3]	[13.9]
Quintile 2	29.2	27.6	19.5	13.9	9.7
	(22.4)	(21.6)	(20.3)	(19.2)	(16.5)
	[20.8]	[22.5]	[21.4]	[19.7]	[15.7]
Quintile 3	22.3	24.8	22.4	19	11.4
	(21.4)	(19.4)	(20.5)	(20.8)	(17.8)
	[21.8]	[18.4]	[21.3]	[20]	[18.4]
Quintile 4	9.95	16.5	26.4	27.1	20
	(18.5)	(20.2)	(20.4)	(20.4)	(20.5)
	[18]	[20.4]	[19]	[20.9]	[21.6]
Quintile 5	3.23	5.97	13.7	26.3	50.9
	(13.6)	(16.3)	(19)	(21.5)	(29.6)
	[13.4]	[14.6]	[20.5]	[21.1]	[30.4]

Notes: Observed, unconstrained between brackets, simulated between square brackets. Quintiles of observed data are computed using the non-missing observations. Predicted data are computed adding the permanent component and simulated draws from the distribution of the transitory component. Quintiles are computed on all observations.

Table A.x: Transitions  $t = 16$  to  $t = 31$ 

Quintiles in t=31	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	51.1	31.7	12.2	4.04	0.974
in t=16	(38.1)	(32.1)	(20.3)	(7.24)	(2.23)
	[46.5]	[29.5]	[16.8]	[4.31]	[2.92]
Quintile 2	30.9	35.1	24.1	6.96	2.92
	(30.3)	(26.2)	(23.9)	(15.7)	(3.76)
	[29.4]	[27.8]	[25.7]	[13.2]	[3.76]
Quintile 3	11.4	24.8	36.9	23.8	3.2
	(20.3)	(22.4)	(27.4)	(21.8)	(8.07)
	[18.5]	[24.8]	[28.1]	[22.7]	[5.98]
Quintile 4	4.45	6.82	21.7	46.8	20.2
	(8.35)	(16.1)	(22.4)	(33.5)	(19.5)
	[5.57]	[15.6]	[23.4]	[36.9]	[18.5]
Quintile 5	2.23	1.67	5.01	18.4	72.8
	(2.92)	(3.06)	(5.98)	(21.6)	(66.5)
	[0.139]	[2.23]	[5.98]	[22.8]	[68.9]

Notes: see table A.ix

Table A.xi: Group effects  $\eta_g$ 

Skill group	Age group	Nb Obs	$\bar{\eta}_{g1}$	$\bar{\eta}_{g2}$	$\bar{\eta}_{g3}$
2	17	1268	2.5 (0.037)	0.024 (0.0069)	-0.049 (0.054)
3	17	1224	2.5 (0.04)	0.023 (0.006)	-0.051 (0.044)
1	19	41	2.8 (0.043)	0.053 (0.007)	-0.23 (0.055)
2	19	934	2.6 (0.037)	0.027 (0.0053)	-0.073 (0.039)
3	19	994	2.6 (0.041)	0.025 (0.0073)	-0.073 (0.053)
1	21	117	2.9 (0.087)	0.035 (0.0087)	-0.12 (0.071)
2	21	710	2.8 (0.017)	0.031 (0.0031)	-0.1 (0.024)
3	21	512	2.7 (0.017)	0.025 (0.0032)	-0.09 (0.024)
1	23	171	3.1 (0.019)	0.038 (0.0035)	-0.14 (0.026)
2	23	348	2.8 (0.027)	0.033 (0.0045)	-0.11 (0.034)
3	23	254	2.8 (0.043)	0.035 (0.0057)	-0.15 (0.042)
1	25	191	3.4 (0.054)	0.045 (0.0063)	-0.19 (0.047)
2	25	146	2.9 (0.061)	0.021 (0.0061)	-0.044 (0.042)
3	25	93	2.7 (0.024)	0.016 (0.0037)	0.0083 (0.031)
1	27	114	3.5 (0.027)	0.03 (0.006)	-0.12 (0.047)
2	27	87	3.1 (0.021)	0.044 (0.0037)	-0.22 (0.028)
3	27	63	2.7 (0.039)	0.013 (0.0059)	0.02 (0.046)
1	29	58	3.3 (0.058)	0.024 (0.0074)	-0.042 (0.056)
2	29	67	2.9 (0.083)	0.021 (0.013)	-0.097 (0.11)
3	29	55	2.7 (0.053)	-6.2e-05 (0.0082)	0.1 (0.064)

Note: Standard error in parentheses. A flat spot deflator is used

Table A.xii: Frequencies of violations of single-dimensional restrictions: Robustness

Restrictions $\rightarrow$	$\eta_2 \geq 0$	$\eta_3 \leq 0$	$\eta_3 + \pi_T \eta_2 \geq 0$
Sample periods			
(3,15]	0.13 (0.041)	0.16 (0.046)	0.15 (0.05)
(15,22]	0.12 (0.019)	0.18 (0.021)	0.19 (0.022)
(22,26]	0.068 (0.012)	0.13 (0.018)	0.21 (0.024)
(26,28]	0.039 (0.01)	0.1 (0.02)	0.18 (0.024)

Notes: Sample periods = number of observed periods. 5 per cent level rejection frequency of single-dimensional tests of restrictions. Standard errors (sampling and parameter uncertainty, 1000 MC simulations) between brackets. A flat spot deflator is used.

Sample periods	P-values < 0.10	0.05	0.01
(3,15]	0.17 (0.012)	0.13 (0.011)	0.085 (0.0087)
(15,22]	0.23 (0.011)	0.18 (0.01)	0.13 (0.0086)
(22,26]	0.24 (0.0092)	0.18 (0.0084)	0.13 (0.0072)
(26,28]	0.22 (0.0092)	0.17 (0.0082)	0.11 (0.0069)
Complete sample	0.22 (0.0051)	0.17 (0.0047)	0.11 (0.0039)

Notes: Sample periods = number of observed periods. Frequency of p-values of the test of restrictions satisfying the conditions. Standard errors (sampling and parameter uncertainty, 20 Monte Carlo simulations) between brackets. Statistic distribution obtained by 150 replications. A flat spot deflator is used.

Table A.xiii: Frequencies of violations of the global restriction: Robustness



Table A.xiv: Quantiles of structural estimates: Robustness

	Sample periods	0.05	0.2	0.35	0.5	0.65	0.8	0.95
$\varkappa$	(3,15]	6.4 (0.297)	7.5 (0.235)	8.1 (0.25)	8.62 (0.282)	9.13 (0.303)	9.76 (0.224)	11 (0.276)
	(15,22]	4.67 (0.26)	6.32 (0.182)	7.09 (0.241)	7.67 (0.253)	8.29 (0.274)	9.18 (0.74)	11.2 (0.324)
	(22,26]	2.85 (0.166)	5.11 (0.159)	6.36 (0.169)	7.31 (0.296)	8.17 (0.249)	9.41 (0.339)	12.8 (0.335)
	(26,28]	1.91 (0.145)	4.43 (0.189)	5.92 (0.215)	7.29 (0.259)	8.66 (0.331)	10.8 (0.46)	14.7 (0.386)
	(3,15]	0.161 (0.0768)	0.225 (0.0788)	0.286 (0.0849)	0.362 (0.104)	0.469 (0.137)	0.713 (0.195)	2.03 (0.526)
	(15,22]	0.158 (0.0881)	0.255 (0.0799)	0.336 (0.0913)	0.434 (0.115)	0.602 (0.125)	0.896 (0.222)	2.59 (0.67)
	(22,26]	0.102 (0.0819)	0.23 (0.104)	0.347 (0.116)	0.48 (0.153)	0.697 (0.207)	1.23 (0.249)	4.3 (0.664)
	(26,28]	0.0797 (0.0576)	0.158 (0.123)	0.287 (0.159)	0.459 (0.177)	0.752 (0.226)	1.4 (0.305)	5.1 (0.793)
$c_L$	(3,15]	7.88 (1.05)	29.2 (1.75)	58.9 (2.56)	130 (3.97)	324 (6.69)	1048 (14.3)	13612 (221)
	(15,22]	15.1 (1.65)	78.3 (3.09)	200 (3.66)	434 (5.62)	983 (11.5)	3196 (33.1)	51121 (969)
	(22,26]	3.57 (0.865)	59.4 (3.04)	208 (5.19)	528 (9.89)	1603 (14.3)	6547 (39)	101232 (1065)
	(26,28]	1.48 (0.521)	17.9 (2.3)	113 (5.02)	460 (8.77)	1689 (17.7)	7271 (46.6)	158715 (1327)

Notes: A flat spot deflator is used.

Figure A.i: Change over time in mean and variance of log earnings for cohorts 1977-2000

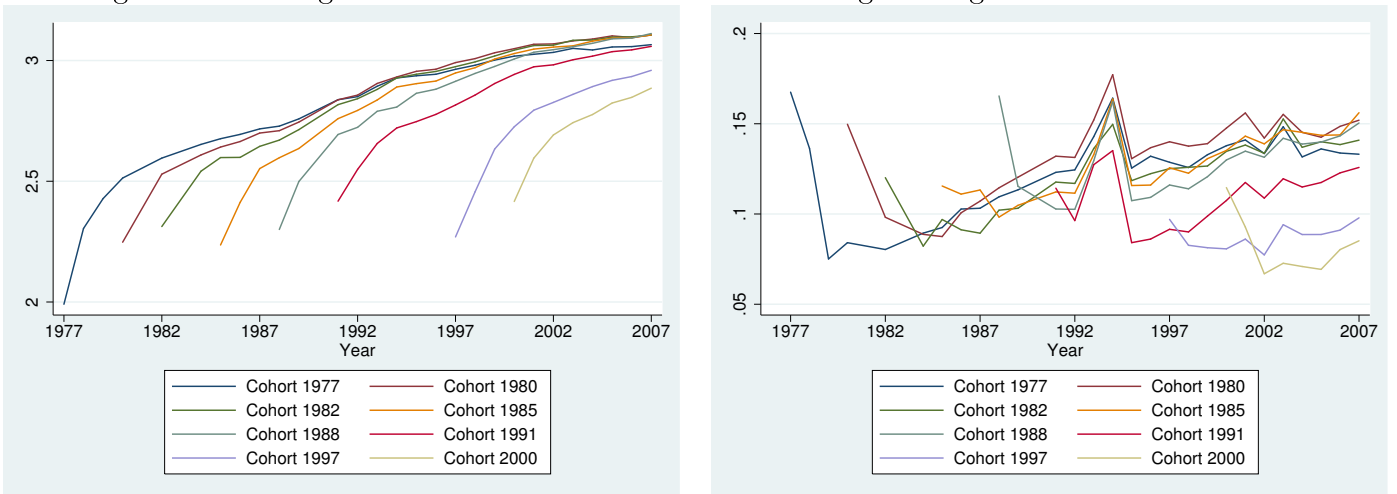


Figure A.ii: First order autocorrelation relative to potential experience for 1977, 1987 and 1997 entry cohorts

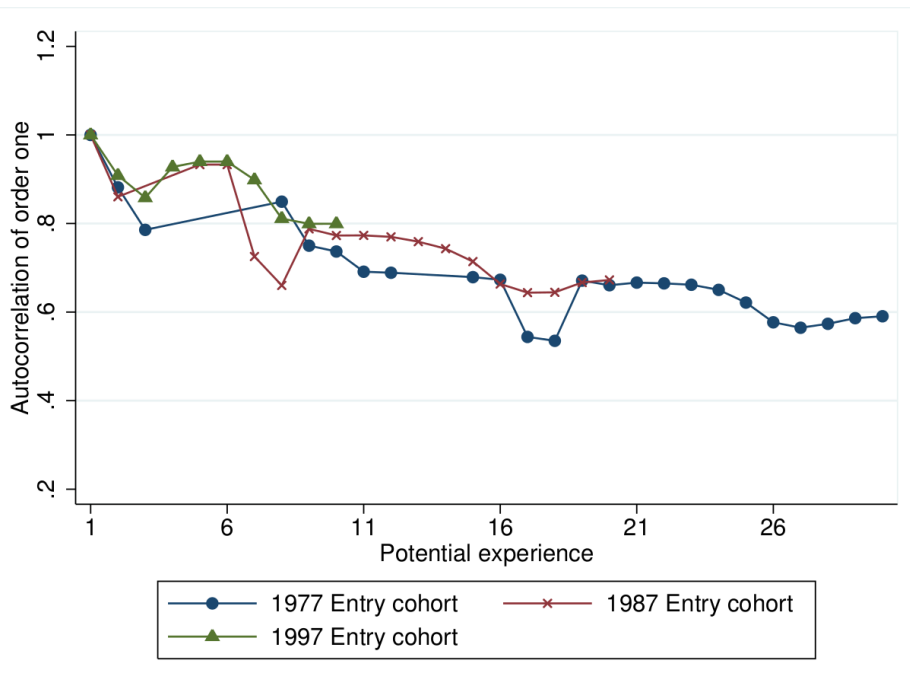


Figure A.iii: Unconstrained estimates: variance of residuals  $v_{it}$  by age and skill group

