WORKING PAPERS

N°TSE- 380

January 2013

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

Thierry Magnac, Nicolas Pistolesi and Sébastien Roux

INSTITUT D'ECONOMIE INDUSTRIELLE

Toulouse School of Economics

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

Thierry Magnac*, Nicolas Pistolesi†, Sébastien Roux‡

First version: April 2011

This version: January 2013

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 rate of return, their costs of human capital investments and their terminal values of human capital at a fixed 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, a random effect model is estimated first by pseudo maximum likelihood methods. This step is followed by a simple second step fixed effect method by which individual-specific structural parameters are estimated. This allows us to test restrictions, compute counterfactual profiles and evaluate how earnings inequality over the life-cycle is affected by changes in structural parameters. Under some conditions, even small changes in life expectancy seem to imply large changes 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)

‡CREST INSEE, Paris

1 Introduction[1]

Since the seminal work by Lillard and Willis (1978) on the estimation of reduced form earnings dynamics an extensive literature has emerged. Its main motivation comes from the assessment of differences between short-run and long-run earnings inequalities (see for instance Bönke, Corneo and Lüthen, 2011 for a survey) and from the joint modeling of consumption and income variance (for instance, Meghir and Pistaferri, 2010). While a very large set of empirical studies estimating ARMA models on earnings residuals has been conducted over recent years, the literature has not reached any consensus on a unique specification of the earnings process. Most authors admit that a mixed process with individual-specific effects along with autoregressive and moving average components seems necessary to fit the longitudinal change in earnings dispersion that is commonly observed although they do not agree on the description of earnings growth. Several papers have considered a beauty contest between a specification in which earnings growth is random and a specification in which earnings growth is governed by a linear trend multiplied by a fixed individual effect (see Baker, 1997, Guvenen, 2009, and Hryshko, 2012, for instance). Yet, the theoretical structural background justifying the reduced forms used in these papers are unclear although additional structure would help discriminating between them.

This is why the first contribution of this paper is to develop an empirically tractable theoretical model of human capital investments accommodating substantial unobserved heterogeneity and from which we derive a convenient reduced-form for the dynamics of earnings – in logarithms as this is the most popular specification. We follow Mincer (1974) and more specifically his research program on post schooling wage growth *Accounting identity model* as presented by Heckman, Lochner and Todd (2006) and as formalized in the theoretical model of Ben Porath (1967). We explain differences in individual earnings life-cycle profiles by heterogenous choices of human capital investments driven by heterogeneous individual characteristics. In a sense, we are extending to post-school investments what has been developed times ago by Heckman (see Heckman, Lochner and Todd, 2006, for a survey) and Card (for instance in the Econometrica lecture in 2001) for schooling investments in human capital.

The model delivers the well known predictions of a human capital setting (Rubinstein and Weiss, 2006). Earnings profile are increasing and concave and this reflects the shortening of the

[1]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 and Bernard Salanié for helpful 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 and Hong Kong University for their helpful comments. All errors remain ours.

investment horizon. Second, the variance of earnings has a U-shape along the life-cycle because large-returns investors have a steeper earnings profile than low-returns individuals experiencing a flatter profile and these profiles cross after a few years. 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.

Adopting a highly stylized human capital model comes at the price of symplifying other elements that might drive earnings dynamics. 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). Some reduced-form specifications such as Alvarez, Browning and Erjnaes (2010) who try to model the whole distributions of earnings are richer in terms of heterogeneity but ours is enough to model life-cycle profiles of mean and variances of earnings that condition the main diagnostics about life-cycle earnings inequality. 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). Finally, we do not model general equilibrium effects as in Heckman, Lochner and Taber (1998).

In a nutshell, the model developed in this paper summarizes life-cycle profiles of individual earnings by a limited number of individual-specific components which are economically interpretable. Individuals 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 as in Mincer's original model. We also assume that the marginal cost of producing skills 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 of this set-up, 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 identified while others are only partially identified. Ironically, our set-up is able to generate the two most popular specifications – random growth and random walk – used in the reduced form literature albeit in a sequential way along the life cycle.

Our second contribution is to 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 data is an administrative dataset collecting earnings in the private sector and having advantages and drawbacks for our purpose. The first key advantage is that it includes enough observations so that we can study a large single cohort of individuals (more than 7,000). They enter the labor market simultaneously 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). Secondly, 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. We will see that we find much longer dependence for transitory earnings than what is usually found in the literature. These data have shortcomings as well since first, few other individual characteristics than age and broad skill groupings are available. The panel is also affected by attrition since some individuals leave the private sector, temporarily or definitely, because of unemployment, self-employment, non-participation or because they start working in the public sector. This explains why we choose to use male earnings data only in order to mitigate the non-participation selection issue.

Our third contribution is an original empirical strategy that uses a sequence of random and fixed effect methods in order to be able to compute interesting counterfactuals i.e. the non-linear impacts of changes in the environment. We first estimate the model by random effect pseudo maximum likelihood (Alvarez and Arellano, 2004) and then derive fixed effect estimates of the individual factors in a second step. As fixed effect estimates are biased, we evaluate their bias and show that it becomes second-order when the number of period observations is roughly above 20. We also correct for bias and find that it tends to overcorrect. Using those fixed effect estimates, we evaluate structural restrictions and compute estimates of the structural unobserved heterogeneity terms. This enables us to construct counterfactuals measuring the impact of changes in those structural estimates. The alternative strategy of estimating distributions of individual-specific effects as in Cunha, Heckman and Schennach (2010) turns out to be difficult because of structural constraints on individual effects while direct fixed effect estimation is performed at a reasonable cost.

Our main results can be summarized as follows. In the first-step random effect estimation, we find that ARMA orders for individual-and-period specific shocks are much larger than in the

literature. Our preferred specification is an ARMA(3,1) in which period-heteroskedastic variance decreases over time. This indicates much longer persistence than usually thought although the presence of a unit root is strongly rejected as in for instance Alvarez et al (2010) or Alvarez and Arellano (2004). Levels and growth of earnings are positively correlated in the long run and long-run and initial levels are negatively correlated which corroborates one of the predictions of the human capital setting as seen above. Finally, the larger the level and the slope of earnings profiles, the more concave they are and this stems from the horizon effect.

The second step fixed effect estimates show that structural restrictions are satisfied in most of the sample although there seems to exist a small fraction of earnings profiles which do not agree with the set-up. It is our maintained assumption that human capital investments are positive until the end of the observation period that seems mostly rejected.

Finally, a counterfactual analysis 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 and those increases can be attributed to investment heterogeneity between individuals. Cross-section inequality increases by around 20% at the end of the period although this figure has quite a large standard error.

In the next Section we briefly review the literature on the estimation and the empirics of earnings equations. Next in Section 3, we describe the model of human capital accumulation and derive the structural equation for log earnings. In Section 4 we present our empirical strategy and detail the econometric estimation methods that we use. Data are described in Section 5 and results are presented in Section 6. After a discussion of a possible extension, a final Section concludes.

2 A Brief Review of the Literature

The literature connected to what we are doing is huge and this brief presentation cannot summarize all these connections in a comprehensive way.

First, the empirical literature on earnings dynamics, as reviewed in Meghir and Pistaferri (2010), began with the seminal works by Lillard and Willis (1978), Lillard and Weiss (1979) and MaCurdy (1982). These papers are the starting points of the *random growth* and *random walk* specifications that are designed to fit the evolution of variance of earnings over the life cycle – or more exactly of their residuals after a first-stage regression on covariates like education and age – as well as their autocorrelation.

The *random growth* or *heterogeneous income profile* model consists in having unobserved heterogeneity in levels as well as in first differences or growth in earnings. This model was estimated on various datasets, though mostly the PSID, and their results are reported in papers by Lillard and coauthors cited above, and for instance by Hause (1980), Lillard and Reville (1999) on US data, Dickens (2000) on U.K. data, Cappellari (2004) on Italian data, Sologon and O'Donoghue (2009) on European data and many others. Importantly, this specification allows to test Mincer's (1974) theoretical prediction that the variance of earnings should decrease at the beginning of the life cycle until those highly investing in human capital catch up weaker investors. Empirically this would translate into a negative covariance between the individual heterogeneity terms in level and growth and this result has been confirmed by these studies.

The *random walk* model of MaCurdy (1982), alternatively called restricted income profile posits that earnings residuals are the sum of a random walk and a transitory earnings process which is of an ARMA type. The same specification has also been estimated by Abowd and Card (1989), Moffit and Gottschalk (1995, 2002, 2008), Kalwij and Alessie (2007) although there are variations in the orders of ARMA processes which are used in those papers.

Baker (1997) was the first to compare the performance of random growth and random walk models. He primarily concluded that tests of one against the other had low power even if the randow walk seems to slightly dominate the other. Guvenen (2007) followed up and studied the implications of the form of the income process on consumption inequality. He compares predictions of *random walk* and *random growth* models using life cycle consumption and simulated data. Guvenen concludes that a model with heterogenous earnings growth is better able to replicate the observed change in consumption inequality than a model with a unit root. In Guvenen (2009) the sources of identification between the two income processes are more deeply investigated. A major difference between the model in which agents have heterogenous earnings profiles and the model in which they are subject to persistent shocks is that in the former case, the autocorrelation of first differences of earnings residuals remains significant because of the presence of unobserved heterogeneity in earnings growth. Guvenen's analysis favours the heterogenous growth specification. In contrast, Hryshko (2012) arrives at the opposite conclusion that random walk specification offers the best fit when trying to test random walk against random growth specifications using PSID data on earnings and a fixed number of ARMA lags.

Other contributions have generalized the model in the direction of non linear and non normal models that would allow a less parametrically driven view of what happens in the tails of the earnings distribution. Geweke and Keane (2000) implements Bayesian inference methods

and show that the share of variance explained by permanent individual heterogeneity terms is larger than under a Gaussian model. Hirano (2002) uses a Bayesian framework to propose a semi-parametric estimator for autoregressive panel data models. Bonhomme and Robin (2009) focus on the same issue and model the change over time in earnings using copula. Marginal distributions of earnings are fully non parametric and the joint distribution is flexibly modelled over a three-year span of panel data.

Alvarez, Browning and Ejrnaes (2010) allows for a lot of heterogeneity and non-linearities in earnings distributions in order to get a better fit of the tails of the earnings distribution and estimate the model using indirect inference. In contrast with Hryshko (2012), they do not find any evidence of a unit root in the dynamics. In a different vein, Meghir and Pistaferri (2004) postulate a non-linear ARCH(1) data generating process for the permanent and for the transitory shocks. Estimating the model by educational group, Meghir and Pistaferri (2004) conclude that the variance of shocks is persistent in some education groups. In a similar framework, Hospido (2010) models the heterogenous variance of earnings but instead of implementing a GMM approach, she uses bias-corrected likelihood methods. Finally, an alternative strand of research simultaneously consider earnings dynamics and mobility on the labor market (see for example Altonji, Smith and Vidangos, 2009).

Methodological issues also arise in this context. The model of earnings residuals that we specify in this paper can be viewed as resulting into a specific covariance structure over time that can be fitted to the empirical covariance of earnings. Minimum distance as in Abowd and Card (1989) is severely small-sample biased (see Arellano, 2004 for a review) and although the emphasis in the dynamic panel data literature is slightly different, the lessons from this literature are useful to remember here. As is well known in GMM estimation, the range of moments involved when the time dimension becomes larger makes first order asymptotics a poor guide in empirical research. 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 quasi-maximum likelihood estimators (Hsiao, Pesaran and Tahmiscioglu, 2002, Dhaene and Jochmans, 2009). Another direction was recently proposed by Han, Philips and Sul (2010) in the case of AR(p) models under mean stationarity whose properties are robust and simple to derive under both stationary and non stationary cases.

As T is neither large nor small in our application and as we stick to a framework in which the initial conditions are supposed to have been generated by another stochastic process so that asymptotic stationarity properties are not satisfied, the GMM framework remains our reference.

Alvarez and Arellano (2003) analyses the asymptotic properties of GMM estimators using double asymptotics in N and T. Okui (2009) derives the small sample biases not only in the mean but also in the variance of GMM estimates because of the presence of too many moments even in the case in which T is small. Okui suggests some moment selection mechanism in order to limit the importance of these biases by, to put it briefly, selecting out moments between variables which are too far apart in time. Those moments are far more likely to contribute to larger bias and not to smaller variance. There is another route through quasi-maximum likelihood methods that reduces the number of moments available for estimation as suggested by Alvarez and Arellano (2004). In a comparison with other fixed T consistent estimators, this estimator seems to dominate in most Monte Carlo exercises the maximum likelihood estimator using differenced data (Hsiao et al., 2002) and the corrected within group estimator. In their application to PSID, they do not find any evidence of a unit root.

Closely related to our model is an empirical factor model using panel data on earnings whose use was pioneered by Jim Heckman through a series of papers with diverse coauthors. The first objective of this research was to restrict the set of joint distributions of two or more potential outcomes. If these outcomes are selectively observed such as in the case of a binary treatment, their joint distribution albeit not identifiable in the generic case becomes identified in the linear factor case. Aakvik, Heckman and Vytlacil (2005) makes this point in a general Roy model using one factor and Carneiro, Hansen and Heckman (2003) extends it to the multiple factor case. Furthermore, this setting allows to address the issue of discriminating between heterogeneity from uncertainty in educational decisions (Cunha, Heckman and Navarro, 2007) and to investigate the empirics of skill formation (Cunha and Heckman, 2008). Finally, Cunha, Heckman and Schennach (2010) extends these results and results from Schennach (2004) to a non linear factor set up and show how non parametric estimates of moments of latent variables can be constructed from various measurements of these variables using empirical characteristic functions and inverse Fourier transforms.

In our factor model, factors are known. Arellano and Bonhomme (2010) look in detail to the identification of the distribution of individual effects or factor loadings when the time dimension is fixed and show that its variance is identified under restrictions of the dynamics. They also propose the construction of non parametric estimates for the distribution of the individual factor loadings. That factors are known is in contrast with Bai (2009) who derives MLE estimates in factor models in which the time factors are unknown and in the presence of covariates. In contrast to the linear factor approach we adopt, deconvolution methods might also be an interesting

but more arduous route to follow. Horowitz and Markatou (1996) estimate semi-parametrically the distributions of idiosyncratic terms and individual effects. However, in their approach the dynamic dimension has to be restricted to be AR(1). Geweke and Keane (1998) and Hirano (2002) generalized the model in the same direction by implementing a Bayesian approach to estimate posterior distributions of the parameters. Bonhomme and Robin (2010) construct an estimator of the distribution of factors using empirical characteristic functions and apply this estimator to analyze the distributions of permanent and transitory components of earnings using the PSID i.e. in a random growth setting.

Finally, general equilibrium effects with microeconomic foundations is another direction taken in the literature. Heckman, Lochner and Taber (1998) were among the first to model human capital investments at school and later in life in a dynamic and stochastic general equilibrium set-up. This allows them to estimate the effects on inequality of counterfactual productivity shocks. In the recent literature, 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, the latter resulting to be the larger component of variance. The previous literature on this theme is reviewed in these two last papers.

3 The Model

We present a model of human capital investment in discrete time and in which agents face individual specific costs, individual specific rates of return and individual specific terminal values of human capital stocks. As in Ben Porath (1967) and Mincer (1974), we characterize the optimal sequence of human capital investments over the life cycle. The key new point is that the reduced form of the life cycle earnings equation is a log-linear factor model with three factors whose factor loadings are in relation with the individual specific structural determinants. We analyze the transformation between parameters of the reduced and structural forms and the ensuing structural restrictions on factor loadings.

3.1 The set up

Individuals enter the labor market at a period which is normalized to $t = 1$. The entry decision in the labour market is endogenous and depends on previous human capital accumulation. We take, however, these initial conditions as given and depending on unobserved variables, among which the human capital stock at entry. These initial conditions are potentially correlated with all shocks affecting the life-cycle dynamics of earnings.

From period 1 onwards, agents can acquire human capital through part-time training. Human capital is supposed to be single-dimensional so that skills are general and costs are borne by the workers. Labor supply is inelastic and potential individuals earnings, $y_i^P(t)$ are given by an individual-specific stock of human capital, $H_i(t)$, times an individual specific rental rate, $\exp(\delta_i(t))$ that is $y_i^P(t) = \exp(\delta_i(t))H_i(t)$. Individuals face uncertainty through the variability of the rental rate of human capital $\delta_i(t)$ which is mainly affected by aggregate shocks but also by individual ones if there are some frictions in the labor market. Firms might temporarily value individual specific human capital differently than the market in order to attract, retain or discourage specific individuals. The rental rate is supposed to follow a stochastic process and $\delta_i(t)$ 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).

By substracting human capital investments, current individual earnings are assumed to be given by:

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

where $1 - \exp(-\tau_i(t))$ can be interpreted as the fraction of working time devoted to investing in human capital as in the original Ben Porath formulation. It might also be interpreted as the level of effort put in the acquisition of human capital at the cost of losing a fraction of potential earnings. We call, $\tau_i(t)$, somewhat abusively the level of investment in human capital at time t and actual earnings are equal to potential earnings when $\tau_i(t) = 0$. There is no upper bound on $\tau_i(t)$ although an infinite investment value would mean that the individual has not yet entered the labor market.

Because of these investments, individuals accumulate human capital in a way that is described by the following equation

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

where $H_i(t)$ is the stock of human capital, ρ_i an individual specific rate of return of human capital investments and $\lambda_i(t)$ is the depreciation of human capital in period t. This latter component

embeds individual-specific shocks or innovations at the economy level as these innovations 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 or because of voluntary moves across firms or sectors. As $\delta_i(t)$, the variable $\lambda_i(t)$ is supposed to be revealed at period t to the agent and is uncertain before. We also take the stochastic process $\lambda_i(t)$ as a given.[2]

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, we assume that period t utility is equal to current log earnings net of investment costs so that there is no consumption smoothing over time.[3] Period-t utility is written as :

$$\delta_i(t) + log\ H_i(t) - \left(\tau_i(t) + c_i \frac{\tau_i(t)^2}{2}\right)$$

in which c_i represent between-individual differences in the cost of human capital accumulation in utility terms and the cost is quadratic. Note that the coefficient of $\tau_i(t)$ is set to 1 because it corresponds to the standard formulation of Ben Porath (1967) in which the objective function would be a function of current earnings or their logarithm only :[4]

$$\delta_i(t) + log\ H_i(t) - \tau_i(t). \tag{2}$$

Quadratic costs adds richness to the setting and it fits well with the interpretation of $\tau_i(t)$ in terms of effort exerted for human capital investments and not only time as in the simple specification. Quadratic costs makes the solution in $\tau_i(t)$ unique (see below).

Nonetheless, the costs of investments do not depend on the level of human capital $H_i(t)$ as in a Ben Porath setting in which equation (1) would include a non linear term $H_i(t)^\alpha$ instead of setting $\alpha = 1$ as we do. This is in fact another way of making the solution $\tau_i(t)$ uniquely determined. Furthermore, our specification avoids the "regression to the mean" effect emphasized by Huggett, Ventura and Yaron (2006) that makes individuals closer and closer at the end of

[2]We shall also assume additional technical assumptions such as $E_{t-h}(|\delta_i(t)|) < \infty$ and $E_{t-h}(|\lambda_i(t)|) < \infty$ so that the dynamic program is well defined. For the sake of readability these standard assumptions are not stated here (see Stokey and Lucas, 1989).

[3]Our conjecture is that there does not exist a dynamic model with financial and human capital accumulation that would generate a log-earnings equation if the financial asset accumulation equation is written linearly in income. In contrast, there does exist a dynamic model which generates a factor-like earnings equation in levels, allows both financial and human capital and has a factor format. This case is developed in a companion paper.

[4]We investigated the case in which the linear cost parameter is left free and this parameter is difficult to separately identify from ρ_i and c_i (see also below).

their working life. Moreover, Section 7 proposes a convenient generalization of our setting to the case of increasing costs of investment with the level of human capital. It comes at the price of having additional factors in the econometric model.

Returning to the main argument, the decision program of individuals maximizing their discounted expected utility stream over the present and future is given by the following Bellman equation:

$$V_t(H_i(t), \tau_i(t)) = \delta_i(t) + \log H_i(t) - \left(\tau_i(t) + c_i \frac{\tau_i(t)^2}{2}\right) + \beta E_t \left[W_{t+1}(H_i(t+1))\right] \tag{3}$$

in which β is the discount factor and:

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

This dynamic program is completed by a terminal condition that at a future date $T+1$ the value function or the discounted value of utility stream from $T+1$ onwards is given by:

$$W_{T+1}(H_i(T+1)) = \delta^* + \kappa_i \log H_i(T+1). \tag{4}$$

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

$$\kappa_i = 1 + \beta_{T+2} + \beta_{T+3} + ...$$

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

$$\kappa_i \leq \frac{1}{1-\beta}. \tag{5}$$

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.[5] As human capital investments are embodied, a smaller discount rate is a source of decreasing returns to investment as in the original argument used by Mincer and this explains the concavity of earnings profiles.

3.2 The life-cycle profile of investments

When human capital investments are always positive, the profile of investments is summarized in:

[5] As we will see in the empirical section, we fix the value of β at .95 because of weak identification issues. As usual in empirical dynamic models, experiments show that the likelihood function is flat wrt to this parameter. This also explains why we do not make this parameter individual specific and assume that it is homogenous.

Proposition 1 *Suppose that :*

$$\beta\rho_i\kappa_i > 1, \tag{6}$$

then:

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

Proof. See Appendix A.1 ■

Equation (7) expresses the well known result that human capital investments decrease with time. The term in β^{-t} indeed means that it is always better to invest earlier than later because the horizon over which investments are valuable is becoming smaller and smaller. This is the negative value of $\kappa_i - \frac{1}{1-\beta}$ (condition 5) that commands the intensity of the decrease. In addition, levels of investments increase with returns, ρ_i, and decrease with costs, c_i. Finally, condition (6) ensures that investments as given by equation (7) are positive until period T.

It is now easy to analyze cases in which investments in human capital 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* $S_i \in \{1, ., T+1\}$ *so that :*

$$\forall t \geq S_i, t < T+1, \tau_i(t) = 0, \text{ and } \tau_i(S_i - 1) > 0$$

if and only if:

$$\frac{1}{\kappa_{i,S_i}} < \beta\rho_i \leq \frac{1}{\kappa_{i,S_i+1}}, \tag{8}$$

where $\kappa_{i,T} = \kappa_i$ *and* $\kappa_{it} = 1 + \beta\kappa_{i,t+1}$ *for all* $t < T+1$ *(and by convention* $\frac{1}{\kappa_{i,T+1}} = +\infty$, $\frac{1}{\kappa_{i,1}} = 0$*). Additionally, equation (7) describing human capital investments remains valid for all* $t < S_i$*.*

Proof. See Appendix A.2 ■

Because, the sequence κ_{it} of the previous proposition is given by:

Lemma 3 *For all* $t \in (1, T+1)$ *:*

$$\kappa_{it} = \frac{1}{1-\beta} + \beta^{T-t}(\kappa_i - \frac{1}{1-\beta}) \tag{9}$$

Proof. Using definitions in Proposition 2 and by backward induction from T. ■

we can summarize the two propositions into the following:

Corollary 4 *There exists* $S_i \in \{1, ., T+1\}$ *such that:*

$$\frac{1}{\kappa_{i,S_i-1}} < \beta\rho_i \leq \frac{1}{\kappa_{i,S_i}},$$

and:

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

Proof. From Proposition 2, human capital investments stop at period S_i. We can then use equations (7) and (9). ■

This corollary proves that the profile of life-cycle investments is truncated at zero but there are no dynamic effects of the truncation. The profile remains similar even if investments stop. This corollary also shows that if we had information about the duration of the sequence of human capital investments, we would be able to relate this information to parameters ρ_i and κ_i only. In particular, note that the cost parameter, c_i, does not affect this duration and only the level of investments. This is a strong prediction of our set-up and this is due to the separability between investment costs and human capital stocks.

3.3 The Lifecycle Profile of Earnings

We start by deriving the earnings profile when human capital investments remain positive until period T. First, the stock of human capital in period t depends on previous investment choices and past depreciation that is

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

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

$$log\ y_i(t) = \delta_i(t) + log\ H_i(1) + \sum_{l=1}^{t-1}\rho_i\tau_i(l) - \sum_{l=1}^{t-1}\lambda_i(l) - \tau_i(t). \tag{10}$$

It shows that returns to human capital $\delta_i(t)$ cannot be distinguished from depreciation effects $\sum_{l=1}^{t-1}\lambda_i(l)$ and we will therefore write that transitory earnings are equal to:

$$\delta_i^y(t) = \delta_i(t) - \sum_{l=1}^{t-1}\lambda_i(l).$$

Furthermore, inserting into the first sum the structural expression for $\tau_i(\cdot)$ given by equation (7) we get :

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

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.

Finally, using equation (7):

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

and rearranging expression (10) we have the following reduced form expression for log earnings

$$\log y_i(t)=\eta_{i1}+\eta_{i2}t+\eta_{i3}\beta^{-t}+\delta_i^y(t), \tag{11}$$

in which:

$$\eta_{i1} = \log H_i(1)-\frac{\rho_i^2}{c_i}\left(\kappa_i-\frac{1}{1-\beta}\right)\frac{\beta^{T+2}}{1-\beta}-\frac{\rho_i+1}{c_i}\left(\rho_i\frac{\beta}{1-\beta}-1\right), \tag{12}$$

$$\eta_{i2} = \frac{\rho_i^2}{c_i}\frac{\beta}{1-\beta}-\frac{\rho_i}{c_i}, \tag{13}$$

$$\eta_{i3} = \frac{\rho_i^2}{c_i}\left(\kappa_i-\frac{1}{1-\beta}\right)\frac{\beta^{T+2}}{1-\beta}-\frac{\rho_i}{c_i}\beta^{T+1}(\kappa_i-\frac{1}{1-\beta}). \tag{14}$$

From these reduced form equations, it is clear that different permanent and transitory factors contribute to individual earnings trajectories. On the one hand, three types of permanent heterogeneities drive earnings dynamics. Firstly, differences in initial capital investment at school, $H_i(1)$, lead to permanent differences in log earnings. Secondly, between-individual differences in marginal return to investment, ρ_i, and in the cost parameter c_i make earnings growth rates individual specific. Thirdly, the interaction between the marginal return and the cost of investment, ρ_i/c_i, makes earnings profiles differ in amplitude. We shall look below at the form of transitory earnings.

In the case in which human capital investments stop before period T, the previous results can be adapted by replacing period $T+1$ by period S_i as developed in Corollary 4. This affects the definitions of the factors $(\eta_{i1}, \eta_{i2}, \eta_{i3})$ as derived in equations (12) to (14) although it does not affect the form of the earnings equation (11) before and including period $S_i - 1$. Nonetheless after period S_i, human capital investments are equal to zero and the earnings equation (11) is derived by using potential earnings and the accumulation equation:

$$log\ y_i(t) = \delta_i(t) + log\ H_i(t), log\ H_i(t+1) = log\ H_i(t) - \lambda_i(t), \forall t \geq S_i$$

so that we have:

$$log\ y_i(t+1) = log\ y_i(t) + \delta_i(t+1) - \delta_i(t) - \lambda_i(t). \tag{15}$$

Earnings growth becomes stochastic and is no longer determined by the terms η_{i2} and η_{i3}.

An interesting conclusion of these theoretical developments is therefore that the two most popular specifications, the *heterogeneous growth* and *random walk* ones, are both predicted by the theoretical model although not concurrently but as a sequence in the working life of each worker. As long as human capital investments are positive the heterogeneous growth specification applies. It is only when investments stop that the random walk hypothesis becomes the rule. We shall assume in the empirical section, for want of better identification, that the econometric model is given by the heterogeneous growth model and equation (11) so that investments in human capital are positive until the last period of observation T. Next section shows that this condition is testable. If this condition were not true, the earnings equation would be a mixture between a generalized random growth model (11) and a random walk (15) and identification would rely on specific distributional assumptions (see below).

3.4 From the Reduced to the Structural Forms

The structural model not only imposes a three-factor structure on the reduced form but it also imposes restrictions on reduced-form parameters, $(\eta_{i1}, \eta_{i2}, \eta_{i3})$. In addition, the transformation formulas between reduced and structural forms help recovering the distribution of unobserved heterogeneity structural components. We answer in this section two questions. Do restrictions on structural parameters and this system of equations imply any restrictions on the reduced form parameters? Second, are structural parameters identified?

First, equation (12) which describes unobserved heterogeneity in levels in earnings equations allows us to identify the level of initial human capital if the other individual specific terms are fixed. It thus imposes no constraint.

The other equations (13) and (14) are more interesting and can be rewritten as:

$$\eta_{i2} = \frac{\rho_i}{c_i}\left(\rho_i \frac{\beta}{1-\beta} - 1\right), \tag{16}$$

$$\eta_{i3} = \frac{\rho_i}{c_i}\beta^{T+1}(\kappa_i - \frac{1}{1-\beta})\left(\rho_i \frac{\beta}{1-\beta} - 1\right). \tag{17}$$

This is a non linear system of two equations with three unknowns: ρ_i, c_i and κ_i so that parameters are underidentified. Some structural restrictions can nevertheless be binding.

Namely, structural restrictions consist in statements about the lower discount factor after period $T+1$ and about costs and returns parameters i.e.:

$$\kappa_i \in [0, \frac{1}{1-\beta}], c_i > 0, \rho_i > 0. \tag{18}$$

As developed at the end of the previous section, we shall also impose that human capital investments remain positive so that:

$$\tau_i(t) > 0 \text{ for all } t \leq T. \tag{19}$$

We can now summarize reduced-from restrictions and the identification of structural parameters as:

Proposition 5 *Structural restrictions (18) and (19) imply the following restrictions on reduced form parameters :*

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

Parameter κ_i is identified and:

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

Furthermore, parameters (ρ_i, c_i) are partially identified in the sense that there exists values (ρ_i^L, c_i^L) such that

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

and there exits a one-to-one relationship:

$$c_i = c(\rho_i, \eta_{i2}, \eta_{i3}).$$

Proof. See Appendix A.3 ■

Its interpretation is intuitive. The random growth parameter η_{i2} is positive because human capital investments are productive and the curvature term η_{i3} is negative because the horizon is finite and profiles are concave. It is also this curvature relative to the random 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 $T+1$.

3.5 Transitory earnings

In equation (11), transitory earnings $\delta_i^y(t)$ are due to individual specific and aggregate shocks, $\delta_i(t)$ net of human capital depreciations, $\lambda_i(t)$. To this we add measurement errors $\zeta_i(t)$ to obtain that random shocks are described by:

$$\delta_i^y(t) = \delta_i(t) - \sum_{l=s}^{t-1} \lambda_i(l) + \zeta_i(t).$$

Even if measurement errors are independent over time, the effects of the first two transitory components may persist across periods and generates autocorrelation in the earnings residuals. Indeed, the deviation of the rate of return $\delta_i(t)$ from the market rental rate is due to individual specific factors and the match each worker forms with a specific firm. This is likely to persist over time. Depreciation factors included in $\sum_{l=0}^{t-1} \lambda_i(l)$ are highly persistent if $\lambda_i(t)$ is independent over time. It indeed generates a random walk if $\lambda_i(t)$ is iid over time. Nevertheless it needs not be so if $\sum_{l=0}^{t-1} \lambda_i(l)$ is stationary, that is that depreciation shocks are partly compensated in the future. Layoff shocks that force agents to change sectors might be an example of a long persistence in these factors. In order to identify the individual specific parameters ηs, we shall impose in the econometric model, a period-heteroskedastic ARMA structure on these shocks though alternatives such as factor structures might be an interesting route to explore.

Another interpretation stems from 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 controled 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 the earnings is equal to the productivity process because adverse shocks on that process make employee and employer renegociate the wage contract; or alternatively, the wage is equal to the outside offer in the case the employee can either renegociate with his employer or take the outside offer if the productivity is lower that the outside option. We do not impose these structural constraints in this paper and we treat them as an element of idiosyncratic shocks.

Next section describes how we deal with estimation and inference in this model of the earnings formation process.

4 Econometric Modelling of earnings Dynamics

In this section we state our empirical and estimation strategy and details our arguments for the steps we use.

The first key high-level assumption is our choice of estimating the generalized random growth model as given by earnings equation (11) only. It corresponds to estimating parameters of earnings profiles under the null hypothesis that investments are positive until the end of the period of observation. This assumption is one of our structural restrictions and is testable. One reason to proceed this way is that it seems difficult to identify the model under the alternative hypothesis unless one is ready to adopt more parametric assumptions for individual heterogeneity terms. The alternative would indeed imply that the data generating process is a random mixture between a random growth model and a random walk model where the random mixture depends on the value of individual heterogeneity terms. We leave these developments for future research.

Our second key assumption consists in fixing the discount rate at a value equal to 0.95. This solves the identification issue that we face in decomposing empirical variances and covariances of log earnings over time into the effects of the individual specific factors and the effects of the idiosyncratic error terms. Arellano and Bonhomme (2010) shows that along with a finite lag specification assumption about the ARMA process, this assumption is sufficient to get identification. Experiments that we performed in simpler identically and independent settings indeed seemed to indicate that the discount rate parameter is not well identified.

Furthermore, we adopt a strategy in two steps. We first specify a model that is estimable by random effect methods and specifically, we use the pseudo-likelihood estimator suggested by Alvarez and Arellano (2004). Under a normality assumption, the implicit moment selection underlying this estimation method is optimal and though the method loses optimality in the general case, it is still useful for moment selection and for small-sample bias reduction. Though we recover consistent estimates of covariance matrices of individual effects and transitory idiosyncratic terms, using those estimates to impose restrictions, derive structural estimates and compute counterfactuals is computationally difficult. One route would be to use deconvolution techniques although it would require the development of estimation under structural constraints on distributions.

As a simpler next step, we chose to turn to fixed effects estimation which is simple to implement when covariance matrix estimates are known or estimated in the previous random effect step. These fixed effect estimates are admittedly biased if the time span is not long enough

since the order of the bias is $1/T$. Nonetheless we show that for individuals observed over a sufficiently long time period, the bias is empirically of a second order magnitude by comparing variance estimates across fixed and random effect specifications. Using fixed effect estimates, structural restrictions become easily testable and estimates of the reduced-form parameters under structural constraints are easily computed. Yet, these constrained parameters are likely to be at the frontier of these structural restrictions and because the frontier structural parameters are implausible in economic terms, we adopt a simulated approach to draw more plausible estimates of the structural parameters. We directly draw those estimates into the normal approximation of the asymptotic distribution of the fixed effect estimates. The last leg of our empirical strategy is to compute counterfactuals by changing the values of those structural parameters.

In Section 4.1, we specify the covariance structure implied by the reduced form earnings equation (11) and the time-heteroskedastic ARMA assumption that we adopt for transitory earnings. We estimate covariance parameters by random effect methods using the pseudo-likelihood approach as explained in Section 4.2 and then turn to fixed effect estimation and the imposition of structural constraints. We end this section with the computation of counterfactuals.

4.1 Model Specification

Equation (11) can be written as a three-factor model with factor loadings, $\eta_i = (\eta_{i1}, \eta_{i2}, \eta_{i3})$:

$$\log(y_{it}) = \eta_{i1} + \eta_{i2}t + \eta_{i3}\frac{1}{\beta^t} + \delta_i^y(t) \text{ for any } t = 1, ., T. \tag{20}$$

We follow the literature and take deviations from the mean of $\log(y_{it})$ using the finest groupings that are observed in the data as a function of age of entry, skill level and time, say:

$$\log(y_{it}) = E(\log(y_{it}) \mid i \in g) + u_{it}.$$

Denote $\overline{\eta}_{gk} = E(\eta_{ik} \mid i \in g)$ for $k = 1, 2$ or 3 and the centered individual effects as:

$$\eta_{ik}^c = \eta_{ik} - \overline{\eta}_{gk} \text{ if } i \in g,$$

so that equation (20) becomes:

$$u_{it} = \eta_{i1}^c + \eta_{i2}^c t + \eta_{i3}^c \frac{1}{\beta^t} + v_{it} \text{ for any } t = 1, ., T, \tag{21}$$

which is the equation of interest in the random effect estimation below. It will also be useful to remember the between group equation:

$$E(\log(y_{it}) \mid i \in g) = \overline{\eta}_{g1} + \overline{\eta}_{g2}t + \overline{\eta}_{g3}\frac{1}{\beta^t} + E(\delta_i^y(t) \mid i \in g)$$

which allows the identification of the group means of $(\eta_{i1}, \eta_{i2}, \eta_{i3})$ under a restriction on group and period specific effects $E(\delta_i^y(t) \mid i \in g)$. Estimates of these quantities will be used in the fixed effect estimation of the η_is.

On the one hand, our main parameters of interest are factor loadings in equation (21). On the other hand, we specify the stochastic process followed by the shock v_{it} as a "reduced-form" process. Its variances and autocovariances are given by an ARMA structure whose order is known in advance with the additional feature of period heteroskedasticity. Similar specifications of the dependence structure are developed in Alvarez and Arellano (2004), Guvenen (2009) and Arellano and Bonhomme (2010). We define v_{it} as

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

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

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

Whereas alternatives could be the composition of permanent and transitory shocks (Bonhomme and Robin, 2009) or general factor models (Bai, 2009), we chose ARMA models in order to easily test for the presence of any non stationary elements in those stochastics.

4.2 Random Effect Estimation

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

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

The structure of Ω_z structure is detailed in Appendix B although it can be summarized easily. The correlations between initial conditions and individual effects are not constrained, while innovations ζ_{it} are supposed orthogonal to any previous terms including initial conditions. However, the initial conditions $(v_{i(1-p)}, ., v_{i0})$ can be correlated with previous shocks as $\zeta_{i0}, ., \zeta_{i(1-q)}$.

As for the individual effects $(\eta_{i1}^c, \eta_{i2}^c, \eta_{i3}^c)$ we assume that they are independent of the idiosyncratic shocks $\zeta_{i(1-q)}, ., \zeta_{iT}$ while they can be correlated with the initial conditions of the

autoregressive process $(v_{i(1-p)}, ., v_{i0})$ in an unrestricted way. From these restrictions it is possible to build the covariance matrix of the observed variables

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

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

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.

The pseudo likelihood setting is particularly well adapted to the case in which there are mssing data in earnings dynamics. In the case of GMM estimation procedures, we would have to rewrite each moment condition in which there are missing data by replacing the missing variables by their expressions as a function of observed variables. This is untractable in such a dataset in which the number of different missing structures is very large while this is handled with parsimony in a pseudo likelihood setting. 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.

4.3 Fixed Effect Estimates

It is not possible to impose structural constraints on parameters at the estimation stage in the random effect model. It is nonetheless useful to use random effect estimates in order to construct fixed effect estimates of individual factor loadings in a second step. In a log likelihood framework, we obtain fixed effect estimates as linear combinations of residuals, the linear combinations being given by the covariance matrix estimated in the random effect model. Appendix C.1 develops the corresponding analytic computations that lead to define the individual effects estimates as:

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

in which matrix B is a function of random effect parameters and $\hat{B}$ its plug-in estimate.

The bias for the estimated variance of earnings can be computed as in Arellano and Bonhomme (2010). To abstract first from sampling errors, an unfeasible estimate is naturally defined as:

$$\tilde{\eta}_i^c = Bu_i^{[1-p,T]} = \eta_i^c + Bw_i^{[1-p,T]},$$

in which random vector $w_i^{[1-p,T]}$ has mean zero and covariance matrix, Ω_w whose expression is computed in Appendix C.1. We have :

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

The bias term is given by $B\Omega_w B'$ and it is easy to show that the dominating term is of order $1/T$.[6]

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

$$\hat{\eta}_i^c = \hat{B}u_i^{[1-p,T]} = \tilde{\eta}_i^c + (\hat{B} - B)w_i^{[1-p,T]},$$

although this term is in $1/\sqrt{N}$ and thus dominated in large N and moderate T samples by the bias in $1/T$. Note that these biases can be estimated and bias-corrected estimates of those variance terms can be constructed. We shall evaluate them in the empirical section below.

4.4 Constraints and Structural Parameters

From those individual-specific estimates, we now show how to impose the structural constraints derived in Proposition 5. If those constraints are satisfied, structural individual specific estimates can be derived from fixed effect estimates.

Indeed, estimates $\hat{\eta}_i$ do not necessarily satisfy the constraints:

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

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

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

As we know the asymptotic distribution of each factor loadings, we can test each single restriction at the individual level.

[6] Because our factors are a constant, a linear trend and a geometric one, there are also bias terms in $1/T^2$ and $exp(-\beta T)$ that are dominated by the leading one, $1/T$.

Furthermore, we can 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 C.3. We have to reintroduce first the estimates of individual effect averages i.e.:

$$\hat{\eta}_i = \bar{\eta}_g + \hat{\eta}_i^c \text{ if } i \in g$$

We can then construct the distribution of the distance in the data between the unconstrained and the constrained estimates, $\hat{\eta}_i^R$:

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

which 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 is a mixture of chi-square distributions and we evaluate this distribution by simulation in the empirical section below.

Nevertheless, constrained estimates are on the frontier of structural restrictions by construction when the unconstrained estimates are outside the set of structural constraints. This happens quite often even when the null hypothesis is true and when the number of observed periods T_i is small. 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 κ_i is equal to 0. Because $\rho_i > 1/\beta\kappa_i$, this would generate an implausible large estimate for ρ_i.

This is why we use simulation to sample into the asymptotic distribution of constrained estimates. We use that the likelihood function of an individual earnings profile is given by:

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

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 η_i using the developments in Appendix C.4. Some additional trimming to avoid frontier points is used.

4.5 Counterfactuals

We want to analyze the impact of a change in the levels of the terminal capitalization rate κ_i or in the rate of return ρ_i for instance. Other parameters will remain fixed since it would require a proper modelling or specific assumptions. This applies in particular to the initial level of human capital as well as to the rental price of human capital over time which will assumed constant across these experiments. It is fair to note that counterfactuals we construct are independent of these variables only under restrictions that we state below.

4.5.1 Survival probabilities

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 κ_i into κ_i^*:

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

as if we were prolonging, all of a sudden, life expectancy by K years. Other parameters ρ_i and 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 remain the same. We assume that there is infinite demand for human capital at the rental prices that were effectively observed and we assume that decumulation shocks remain the same so that the transitory earnings process also remains the same.

Evaluating equations (12) to (14) at the new values $(\kappa_i^*, \rho_i, c_i, H_i(1))$ demonstrate 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:

$$\eta_{i1}^* - \eta_{i1} = -\frac{\rho_i^2}{c_i}\left(\kappa_i - \frac{1}{1-\beta}\right)\frac{\beta^{T+2}}{1-\beta}(\beta^K - 1). \tag{22}$$

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

Nonetheless, parameters ρ_i and c_i in equation (22) are not identified and only a lower bound (ρ_i^L, c_i^L) on their values can be computed. We shall then proceed by making different assumptions like $\rho_i = \rho_i^L$, $\rho_i^L = 1.20\rho_i$ etc to assess the robustness of this construction.

4.5.2 Human capital technology

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 led 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 these rates.

5 Descriptive Analysis of the Data

5.1 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 wage 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 it is impossible to 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 4% since 1976. Regarding the sampling device, all individuals who were born in October of even years should be included in this sample. Nevertheless, there are two main reasons why observations can be missing. 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 only. 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.

In addition, we aim at keeping only employees with a permanent full-time attachment to the private sector. Firstly, we considered persons employed full time only and censor information about part-time jobs. We also restricted the sample to men entering the labor market in 1977 and working in the private sector in 1978, 1982 and 1984 to avoid non-participation and also because the bulk of entries as a public servant occurs at the beginning of the working life. The definition of entry here is the same as in Le Minez and Roux (2002). We consider that an individual has entered the labor market as soon as this individual has occupied the same job for more than 6 months and is still employed the following year, possibly in a different firm. The date of entry defines the cohort to which the individual belongs and we focus on a single cohort of entrants in 1977, the first year of our panel which lasts until 2007.

We impose these restrictions in order to concentrate on a relatively homogeneous sample of workers with a long term attachment to the labor market to which private firms have access.

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 full earnings divided by the number of days worked and remultiplied by the average annual workload. 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 using administrative data is that few observable characteristics are available apart from a rough measure of age at labor market entry and a rough measure of education grouping the first job into three categories. As a measure of skill, we can also use a grouping given by the age of entry. The first group includes individuals entering the labor market when they are less than 20 years old, the second group of individuals enters between age 20 and 23 and the last group after age 24.

We analyze deviations of log earnings with respect to the mean log earnings of workers within the same age of entry and education group at each point in time. That is, we compute earnings residuals, u_{it} as in equation (21):

$$u_{it} = \log(y_{it}) - \overline{\log(y_{it})_{gt}}, \text{ for } i \in g = 1, ., G$$

in which g is the index of groups formed by age of entry and education.

5.2 Earnings Inequality in France Over the Period

As regard earnings inequalities in France, a few recent studies investigate their evolution over roughly the same period as our data. First and foremost, wage dispersion in France has not increased over the last thirty years in contrast with the US or the UK (Moffit and Gottschalk, 2008, Dickens, 2000). It represents an important distinction from previous studies mostly using US data that one should keep in mind.

Charnoz, Coudin and Gaini (2011) use quantile regressions using the same DADS data on full time private sector earnings. They show that earnings inequality in France has been rather stable from 1976 to 1992 and has been slightly decreasing from 1995 to 2004. They point out the important role of the natural replacement of older cohorts by younger ones and the large part played by several increases in the minimum wage over the period. Cornilleau (2012) uses the same data set and looks at changes in means and dispersion of earnings from the nineteen sixties

to 2004. This author concludes to a slight decrease in inequality that balances two counteracting effects. A reduction in inequality at the bottom of the distribution (the ratio of the minimum wage over the median increasing from 51 to 66%) was slightly mitigated by a small increase in dispersion in the upper part of the distribution (the ratio of median wage to the limit of the last decile decreasing from 52 to 50%). In a different vein, Verdugo (2012) relates changes in the wage structure with changes in educational attainment using various datasets. He concludes to a reduction in wage inequality in the lower part of the distribution and a stable dispersion in the upper part.

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

5.3 Data description

Table 1 reports descriptive statistics of the sample. The sample size is 7446 observations in 1977 and 4670 in 2007. Age of entry groups defined above are of unequal size, the low skill group being the largest. Attrition follows a somewhat irregular pattern which is partly due in the first years to our sampling design since we required that wage earners be present in 1977, 1978, 1982 and 1984. 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 and missing data patterns in 1994 are very similar across age of entry groups.

To complete this information, Table 2 gives a dynamic view of attrition. This Table reports the frequencies of reported 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.

We report in Figure 1, the increase of average log-earnings over the period for the three groups defined by age of entry. These are log-earnings at current prices although the shape of real log-earnings is hardly different. Inflation, as measured by consumer prices, leads to a substracting 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. We do not report the evolution of average log-earnings by groups defined by education and age of entry, the only individual characteristics that are available in the dataset although these evolutions are parallel to the ones graphed in Figure 1. Nonetheless, as already said, the variance of log-earnings that we consider from now on are computed by taking deviations of log-earnings with respect to their means in groups defined by covariates and periods.

The left panel of Figure 2 represents the change in the cross-sectional variance of (log) earnings for the full sample, while the right panel represents the variance by groups defined by age of entry.[7] The first few years witness a strong variability of earnings. Until the sixth year 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 earnings variance levels and a larger rate of variance growth over their life-cycle. The full covariance matrix is reported in Table 3 and gives information about correlations although this is easier to use graphs to describe the main features of this matrix. 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 asymetric pattern over time which is quite robust to the choice of these specific years (1986 and 2007). The correlation between earnings at years $t-k$ and posterior t is quickly disappearing between t and $t-k$ in early years of the panel while it is roughly linear in lags in

[7] 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.

late years. Figure 4 takes a different view that confirm 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 we move to the end of the period, the larger the autocorrelation coefficients are.

Finally, Table 4 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.

6 Results

We first present the estimated parameters of the reduced form earnings equation by random effect ML estimation and we discuss the selection of the ARMA specification. In section 6.2 we detail the procedure we implement to estimate unconstrained individual factor loadings or fixed effects. Next, we test and impose structural constraints on estimates. This leads us in Subsection 6.4 to the estimation of structural parameters which are identified (the terminal value coefficient) or partially identified (rates of return). Then we assess the counterfactual impact of changes in life expectancy on the variance of earnings.

6.1 Random effect estimation and reduced form parameters

Firstly, we estimate covariance matrices of the permanent and transitory components of the errors as well as their correlation with the initial conditions. The former is composed by three individual unobserved factors $(\eta^c_{i1}, \eta^c_{i2}, \eta^c_{i3})$, while the latter is represented by an ARMA process as explained in the previous section. Table 5 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 6). That is why in the rest of the analysis we will use as a pivot, results from the ARMA(3,1) model.

Table 6 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 concludes with no doubt that the process is 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). Autoregressive coefficients are ranging from .2 to .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 estimate of the covariance matrix of individual factor loadings is quite stable across the different specifications. Their variances are very precisely estimated at around .30 for the fixed level factor, η_1, and .25 for the geometric factor, η_3, and at around .04 for the linear trend factor, η_2. The correlation between the linear trend and geometric factors is very strongly negative and equal to -.95 consistently across specifications. This is to be expected if the structural constraint derived above between η_2 and η_3 ($\eta_3 \in [-\pi_T \eta_2, 0]$) is verified and we will analyze this issue more in detail below. The correlation coefficient between the geometric and the level factors, η_3 and η_1 is also significantly negative – around -0.6 – and the one between the level and linear trend individual specific terms 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.

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 η_2 and η_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 η_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 7. 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 3). 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.

Long-run vs Short-run decompositions and Goodness-of-Fit

Table 8 provides a decomposition of the cross-sectional inequality into permanent and transitory components. The first column indicates the variance of logs earnings at three points in time 1977, 1981 and 2007 and the mean value over the sample period 1977-2007 for the full sample and three different sub-samples by age of entry on the labor market (for which we estimated the same random effect model). The second column measures the share of the variance due to permanent factors and the third measures the share due to the dynamic component. Firstly, on average, 64% of the variance is due to permanent factors. This share displays a sharp increase over the life cycle from 3% at entry on the labor market to 88% thirty years later. Secondly, there are strong differences by sub-groups. On the one hand, individuals entering the labor market earlier witness less inequality on average than more skilled ones (0.104 vs 0.233), on the other hand they experience a larger share of transitory inequality, while permanent individual heterogeneity is more important for individuals entering the labor market at an older age (72.1% vs 57.4%).

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 evolution of variances is reproduced but 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.

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 variances depend on the number of periods of observation we have for each person. Variances are higher for persons being present in the panel shorter periods of time and biases do not seem to be able to fully explain this feature. Nevertheless, correcting for non random attrition seems out of the scope of this paper and we leave it for further research.

6.2 Fixed effect estimation

Using previous estimates, it is easy to construct fixed effect estimates of the three individual factor loadings. Appendix C.1 shows how we use, for this purpose, random effect estimates of the covariance matrices of log earnings residuals as well as the way we impute back the earnings averages to individual factors. 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.

It is worth recalling that fixed effect estimates are not consistent if the number of periods T is fixed (for instance, Arellano and Bonhomme, 2010). Table 9 presents the estimates of quantiles of their distributions distinguishing subsamples of observations according to the number of periods they are observed (between 4 and 28). The bias in $1/T$ is noticeable as the larger the number of observed periods is, the lower the inter-quartile ratio for all three factors. Overall the median of the coefficient attached to the level factor is of the order of magnitude of the mean earnings at around 2.5 and the range between the 20th and 80th percentile is .5 if the number of observed periods is maximal ($T = 28$). The median of the coefficient of the linear trend factor which can be interpreted as the return to experience at the initial stage 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 time is this coefficient multiplied by $(\log \beta)^2 = 2.5.10^{-3}$. 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.

Table 10 presents estimates of the covariance matrix of individual effects or factor loadings

obtained by fixed and random effect methods. We find again that for a small number of periods of observation (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 grouping 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 $1/T$ bias but also to different underlying stochastics which characterize these two sub-populations. Random effect estimates would be the mixture of these two groups. Table 11 reports the statistics using bias corrected estimates. First, bias-correction at the first order does not seem sufficiently precise to correct the bias for observations 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 observations observed more than 26 periods.

This interpretation finds some confirmation in the representation of 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 graphed 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 β^{-t} (see Appendix B). Transitory earnings are fixed and their passive rôle obscures these comparisons so that we found it better not to include them. Figure 5 graphs the prediction of the variance profiles that can be computed using random effects first, fixed effects second and bias-corrected fixed effects third. We use the subsample in which the number of observed periods is larger than 23 because Table 10 and Table 11 show that the bias is much less severe for such observations. Earnings profiles using fixed effect estimates reproduce the random effect profile at a higher level in Figure 5 although correcting the bias tend to overcorrect. Discrepancies with random effect estimates seem nevertheless second order and this validates the use of this selected sample in the following.

6.3 Structural restrictions and Constrained estimates

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 positive ($\eta_2 > 0$), the coefficient of the geometric factor should be negative ($\eta_3 < 0$) and a weighted sum of these two coefficients should be positive ($\eta_3 + \pi_T \eta_2 > 0$). Parameter $\pi_T > 0$ is fixed in the population and a function of β (see Section 4.4).

An informal way of representing those restrictions is brought about by Figure 6. The clouds

of points for η_2 and η_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. This is no doubt attributable to the very different asymptotic behaviour of the two factors, one being a linear trend and the other being geometric. Second, points in orange (or light) refer to observations for which the sample periods are few (less than 20) and they 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 the satisfaction of the constraints are very sensitive to two key elements. The position of the origin point (0,0) whose estimation depends on the model we have for average earnings that is described in Appendix C.1 and that leads to the imputation of averages for ηs. Second, the π_T parameter which determines the slope of the bottom-left side of the triangle.

More formally, Tables 12 and 13 report frequencies of restriction violations using previous estimates and the same presentation regarding the number of observed periods. In Table 12, we report the sample frequency of *individual* rejections at level 5% of each of the three single restrictions using a standard asymptotic approximation. This frequency tends to decrease with the number of observed periods 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 first two restrictions that the random growth parameter, η_{i2}, is positive and the curvature parameter, η_{i3}, is negative is plausible.

The last restriction involving both parameters is less acceptable at least in the group (22,26] and this restriction is related to the assumption that investments are positive until period T. This hypothesis seems to be more acceptable in the last group in the almost complete subsample, (26,28]. 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 13 reports testing experiments of the three restrictions, globally and it is convenient to compute first constrained estimates of the individual factors. Unconstrained factors are projected on the set of constraints using the quadratic metric given by the covariance matrix of the ηs estimated in the first-step random effect estimation. This procedure is explained in Appendix C.3. The value of the quadratic metric measuring the distance between constrained and unconstrained estimates is the quasi likelihood ratio (QLR) statistic associated to the global restriction (Silvapulle and Sen, 2005). P-values are obtained by simulation in the distribution of the stat-

istic under the null. Table 13 reports the sampling frequency of *individual* rejections using three different levels (0.01, 0.05 and 0.10) using the standard asymptotic approximation.[8] 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 ($T_i \in (22, 26]$) and when the level is small (0.01, 0.05). There seems to exist a "fat tail" of observations for which we reject these restrictions.

To confirm this diagnostic by evaluating the distribution of the QLR statistic, 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 ηs as a weighting matrix to compute the distance and as the basis for simulating the normal errors. Table 14 reports the quantiles of the distributions of the actual and simulated distances. The two distributions coincide rather well for all quantiles until .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).

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

6.4 Structural parameter estimates and Counterfactuals

These constrained reduced form estimates can be used to construct structural parameter estimates and counterfactuals. There are two empirical strategies: keep the whole sample or keep only observations for which the period of observation is long enough and the bias is second-order. We chose the second strategy and kept for these estimations and counterfactuals only observations for which the number of observations is larger than 23.

Parameter κ_i that governs the magnitude of post-retirement returns in human capital can be easily estimated using the distribution of η_2 and η_3 and its distribution is bounded between 0 and $1/(1-\beta) = 1/(1-.95) = 20$ (Proposition 5). As said above, reduced-form parameters at the

[8] 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).

frontier of the constrained set generate implausible economic estimates of structural parameters κ_i or ρ_i. This is why instead of using constrained estimates, we draw simulated estimates into the posterior distribution of the constrained estimates as explained in Appendix C.4. Figure 8 presents a standard kernel estimate of the distribution of the structural parameter κ_i (with no corrections for the left-hand side bound at 0) that shows that the distribution of this parameter is skewed towards value 0. Only a lower bound of parameter ρ_i is identified and its distribution is reported in Table 15. The lower bound lies between .05 and a large maximal value that is due to the constraint that $\rho_{iL} > 1/\beta\kappa_i$.

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 the previous section to compute those counterfactuals. Nonetheless, as developed in Section 4.5, 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 ρ.[9] In Figure 9, the top panel reports the effect on mean earnings for those individuals who are observed more than 23 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 9, bottom panel. Variances are increasing in particular at the end of the period. Because rates of return are heterogenous, a 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 quite imprecise because standard errors are quite large. Nonetheless, we chose an experiment of $K = 2$ years because it is the limit case for which observed and counterfactual 95% confidence intervals are contiguous. When $K = 3$, they are well separated.

7 Extensions

In order to show the versatility of our theoretical approach, we finally develop an extended model in which investment costs depend on human capital levels. Human capital depreciates at a common and exogenous exponential rate $\alpha \in (0, 1)$ in the human capital accumulation equation

[9] In theory, the true identified set is obtained by making the parameters controling partial identification individual specific. Analyzing results in this case is left for future research.

$$H_i(t+1) = H_i(t)^{\alpha} \exp[\rho_i \tau_i(t) - \lambda_i(t)].$$

This is equivalent to making human capital investments more and more costly when human capital levels increase.

Individuals maximize the present discounted value of their earnings streams, and their objective function is given by

$$V_t(H_i(t), \tau_i(t)) = \delta(t) + \log H_i(t) - \left(\tau_i(t) + c_i \frac{\tau_i(t)^2}{2}\right) + \beta E_t \left[W_{t+1}(H_i(t+1))\right]$$

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

$$-\left[1 + c\tau_i(t)\right] + \beta \rho_i H_i(t+1) E_t \left[\frac{\partial W_{t+1}}{\partial H_i(t+1)}\right] = 0. \tag{23}$$

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

$$\frac{\partial W_t}{\partial H_i(t)} = \frac{1}{H_i(t)} + \alpha\beta E_t \left[\frac{\partial W_{t+1}}{\partial H_i(t+1)}\right] \frac{H_i(t+1)}{H_i(t)} \tag{24}$$

Introducing condition (23) into condition (24) we obtain

$$\frac{\partial W_t}{\partial H_i(t)} = \frac{1}{H_i(t)} + \frac{\alpha}{\rho_i} \frac{\left[1 + c_i \tau_i(t)\right]}{H_i(t)}.$$

Inserting this condition at lead $t+1$ in condition (23), we obtain the Euler equation for $\tau_i(\cdot)$

$$\left(1 + c_i \tau_i(t)\right) = \beta \left[\rho_i + \alpha E_t \left(1 + c\tau_i(t+1)\right)\right],$$

which can written, denoting $m_i(t) = \left(1 + c_i \tau_i(t)\right)$, as:

$$m_i(t) = \beta \left[\rho_i + \alpha E_t m_i(t+1)\right]. \tag{25}$$

For $t = T$, condition (24) writes more simply as:

$$\frac{\partial W_T}{\partial H_i(T)} = \frac{\kappa}{H_i(T)},$$

so that condition (25) at time $T-1$ becomes:

$$m_i(T-1) = \beta \rho_i \kappa.$$

We can solve forward equation (25):

$$m_i(t) = \beta \rho_i \sum_{j=1}^{T-t-1} (\alpha\beta)^j + \frac{(\alpha\beta)^{T-t}}{\alpha} \rho_i \kappa$$

so that:

$$(1 + c_i\tau_i(t)) = \rho_i a_t \tag{26}$$

with

$$a_t = \frac{(\alpha\beta)^{T-t}}{\alpha}\kappa + \beta\frac{1-(\alpha\beta)^{T-t-1}}{1-\alpha\beta}$$

and therefore:

$$\tau_i(t) = \frac{1}{c_i}\{\rho_i a_t - 1\} \quad \forall t < T \tag{27}$$

Moreover, the stock of human capital in period t depends on previous investment choices. Using lower case letters to denote log variables (ie: $h_i(t) = \log H_i(t)$):

$$\begin{aligned}
h_i(t+1) &= \alpha^{t-s}h_i(s) + \rho_i\sum_{l=s}^{t}\alpha^{t-l}\tau_i(l) - \sum_{l=s}^{t}\alpha^{t-l}\lambda_i(l) \quad \text{for } t > s. \\
&= \alpha^{t-s}h_i(s) + \rho_i\sum_{l=s}^{t}\alpha^{t-l}\left[\frac{1}{c_i}\left(\rho_i\{\frac{(\alpha\beta)^{T-l}}{\alpha}\kappa + \beta\frac{1-(\alpha\beta)^{T-l-1}}{1-\alpha\beta}\} - 1\right)\lambda_i(l)\right] \\
&= \alpha^{t-s}h_i(s) - \sum_{l=s}^{t}\alpha^{t-l}\lambda_i(l) + \frac{\rho_i^2}{c_i}\sum_{l=s}^{t}\alpha^{t-l}\left[\frac{(\alpha\beta)^{T-l}}{\alpha}\kappa + \beta\frac{1-(\alpha\beta)^{T-l-1}}{1-\alpha\beta})\right] \\
&- \frac{\rho_i}{c_i}\left(\frac{\alpha^{s-t-1}-1}{\alpha-1}\right)\alpha^{t+1-s}
\end{aligned}$$

Since $\log y_i(t) = \delta(t) + h_i(t)$ we have

$$\begin{aligned}
\log y_i(t) = \delta_t - \sum_{l=s}^{t}\alpha^{t-l}\lambda_i(l) + \alpha^{t-s}h_i(s) + \frac{\rho_i}{c_i}\left(\frac{\alpha^{s-t-1}-1}{\alpha-1}\right)\alpha^{t+1-s} \\
+ \frac{\rho_i^2}{c_i}\sum_{l=s}^{t}\alpha^{t-l}\left[\frac{(\alpha\beta)^{T-l}}{\alpha}\kappa + \beta\frac{1-(\alpha\beta)^{T-l-1}}{1-\alpha\beta})\right]
\end{aligned}$$

which is a model with several geometric factors.

8 Conclusion

In this paper, we proposed a structural model of human capital investments that leads to a three linear factor model describing unobserved heterogeneity components of the earnings equation. Using a long panel on a single cohort of wage earners in France from 1977 to 2007, we estimated the reduced form parameters by random effect maximum likelihood methods that deliver the covariance matrix of the random effects. We constructed fixed effect estimates of factor loadings and assess their bias and degree of accuracy. This procedure enables us to evaluate the relevance

of structural restrictions and construct constrained estimators. We then derive estimates of some structural components in the original model in terms of returns and terminal capitalized returns to investments. This allows us to compute richer counterfactuals than the ones that are directly available using random effect procedures.

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 random walk models are hard to discriminate. Furthermore, a simple counterfactual analysis shows 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 this conclusion if 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, 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 modelling is the analysis of a mixture of the heterogenous growth and random walk specifications of the earnings equation and specifically identification issues. Finally, regarding theoretical issues, we are developing in a companion paper the case of linear investment technologies in human capital which allows to analyze financial and human capital investments simultaneously.

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.

Altonji, J., Smith, A. and Vidangos, I., 2009, "Modeling earnings dynamics", Cowles Foundation, 2009-08.

Alvarez, J. and M., Arellano, 2003, "The Time Series and Cross Section Asymptotics of Dynamic Panel Data Estimators", *Econometrica*, 71:1121-1159.

Alvarez, J. and Arellano, M., 2004, "Robust likelihood estimation of dynamic panel data models" Working Paper 0421.

Alvarez, J., Browning, M. and Ejrnaes, M., 2010, "Modelling income processes with lots of heterogeneity", forthcoming *Review of Economic Studies*

Arellano, M., 2004, *Panel data,* Cambridge UP: Cambridge.

Arellano M., and S., Bonhomme, 2010, "Identifying Distributional Characteristics in Random Coefficient Panel Data Models", unpublished manuscript.

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.

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.

Bönke, T., G., Corneo and H., Lüthen, 2011, "Lifetime Earnings Inequality in Germany", working paper, DIW.

Cappellari, L., 2004, "The dynamics and inequality of italian men's earnings: Long-term changes or transitory fluctuations?", *The Journal of Human Resources,* 39(2):475-499.

Card., D., 2001, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica,* 69(5):1127-60.

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", unpublished manuscript.

Cornilleau, G., 2012, "Inégalité de salaires et de revenus, la stabilité dans l'hétérogénéité", OFCE working paper

Cunha F., and J.J., Heckman, 2008, "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation", *Journal of Human Resources*, 43(4):738-782.

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.

Dhaene, G. and K., Jochmans, 2009, "An Adjusted Profile Likelihood for Nonstationary Panel Data Models with Fixed Effects", unpublished manuscript.

Dickens, R., 2000,. "The evolution of individual male earnings in Great-Britain: 1975-95", *The Economic Journal* 110(January): 27-49.

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.

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.

Han, C., P. C.B. Phillips and D. Sul, 2010, "X-Differencing and Dynamic Panel Model Estimation", Cowles Foundation WP 1747.

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, "General-Equilibrium Treatment Effects: A Study of Tuition Policy", *The American Economic Review*, 88(2), 381 - 86.

Heckman, J. J., Lochner, L. L. and Todd, P. E., 2006, Earnings functions, rates of return and treatment e ects: 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.

Horowitz, J. and M., Markatou, 1996, "Semiparametric estimation of regression models for panel data", *Review of Economic Studies,* 63:145-68.

Hospido, L., 2010, "Modelling heterogeneity and dynamics in the volatility in individual wages", forthcoming *Journal of Applied Econometrics.*

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.

Kalwij, A. s. and Alessie, R., 2007, "Permanent and transitory wages of British men 1975-2001: Year, age and cohort effects", *Journal of Applied Econometrics,* 22: 1063-1093.

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 Weiss, Y., 1979, "Components of variation in panel earnings data : American scientists 1960-70", *Econometrica* 47(2): 437-454.

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

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 Roux, S., 2009, "Dynamique des salaires dans une cohorte", *Economie et Prévision,* 187:1-24.

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,* NBER.

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., 1995, "Trends in the autocovariance structure of earnings in the U.S.: 1969-1987". WP Brown University.

Moffitt, R. A. and Gottschalk, P., 2002, "Trends in the transitory variance of earnings in the united states", *The Economic Journal* 112 (March): C68-C73.

Moffitt, R. A. and Gottschalk, P., 2008, "Trends in the transitory variance of male earnings in the US", 1970-2004", unpublished manuscript.

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

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.

Sologon, D. M. and O'Donoghue, C., 2009, "Earnings dynamics and inequality among men across 14 EU countries 1994-2001: Evidence from ECHP", unpublished manuscript.

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

Verdugo, G., 2012, "The great compression of the wage structure in France 1969-2008", WP Banque de France.

APPENDICES

A Proofs of Propositions

A.1 Proof of Proposition 1

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

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

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

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

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

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

so that, by backward induction, we obtain:

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

and so on. This yields:

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

Replacing in equation (A.1) yields:

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

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

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

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

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

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

A.2 Proof of Proposition 2

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

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

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

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

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

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

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), $\beta\rho_i \leq 1/\kappa_{i,t+1}$. Conversely, if $\beta\rho_i \leq 1/\kappa_{i,t+1}$ 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 $\kappa_{iT+1-1} = \kappa_i$ and the linearity of the previous two equations lead to the condition derived by induction again:

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

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

Second, assume that $\tau_i(t) > 0$. Proposition 1 can be recast in a set-up where the last period becomes $S_i = t+1$ 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:

$$\tau_i(t) = \frac{1}{c_i}\left\{\rho_i\left[\frac{\beta}{1-\beta} + \beta(\kappa_{it} - \frac{1}{1-\beta})\right] - 1\right\} > 0,$$

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

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 $t \in \{0, ., T\}$. By convention we set $\frac{1}{\kappa_{i0}} = 0$ in order to cover all cases since $\rho_i > 0$.■

A.3 Proof of Proposition 5

The two equations (16) and (17) simplify to:

$$\begin{cases} \eta_{i2} = \frac{\rho_i}{c_i}\left(\rho_i \frac{\beta}{1-\beta} - 1\right), \\ \eta_{i3} = \frac{\rho_i}{c_i}\beta^{T+1}(\kappa_i - \frac{1}{1-\beta})\left(\rho_i \frac{\beta}{1-\beta} - 1\right). \end{cases} \tag{A.5}$$

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

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

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

$$\frac{\eta_{i3}}{\eta_{i2}} \in [-\frac{\beta^{T+1}}{1-\beta}, 0]. \tag{A.6}$$

Conversely, if this restriction is valid, then κ_i is given by:

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

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.5):

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

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 ρ_i. Therefore $c_i \geq c_L = c(\rho_L, \eta_{2i})$.■

B 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 (21) writes

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

where $u_i^{[1,T]} = (u_{i1}, ..., u_{iT})'$, $v_i^{[1,T]} = (v_{i1}, ..., v_{iT})'$, $\eta_i^c = (\eta_{i1}^c, \eta_{i2}^c, \eta_{i3}^c)$ are the centered versions of the ηs 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 some initial conditions, the number of which depends on the order of the autoregressive process. Denote p this order and write the initial conditions as:

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

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

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

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} 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} is a (time heteroskedastic) ARMA process whose innovations are independent of the individual heterogeneity terms, η_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 supposed to be independent between them although they are both correlated with initial conditions, $u_i^{[1-p,0]}$.

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

B.1 Time heteroskedastic ARMA specification

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

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

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

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

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

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

As $v_i^{[1-p,T]} = \left(v_{i(1-p)}, ..., v_{iT}\right)$, we have:

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

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

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

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

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

$$\begin{pmatrix} I_p & 0 \\ M_T(\alpha) \end{pmatrix}.v_i^{[1-p,T]} = \begin{pmatrix} I_p & 0_{p,T+q} \\ 0_{T,p} & \Lambda.M_T(\psi) \end{pmatrix} \begin{pmatrix} v_i^{[1-p,0]} \\ \zeta_i^{[1-q,T]} \end{pmatrix}. \tag{B.7}$$

To compute the covariance of $v_i^{[1-p,T]}$, we derive the covariance matrix of $\begin{pmatrix} v_i^{[1-p,0]} & \zeta_i^{[1-q,T]} \end{pmatrix}$. 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 $Vy_i^{[1-p,0]} = \Gamma_{00}$. Assuming as usual that $E(y_{i\tau}\zeta_{it}) = 0$ for any $\tau < t$, we have that $E(v_i^{[1-p,0]}.(\zeta_i^{[1,T]})') = 0$. Only $E(y_i^{[1-p,0]}.(\zeta_i^{[1-q,0]})')$ remains to be defined:

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

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

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

because the innovation ζ_i^s is drawn after r and is supposed to be not correlated with y_i^r.

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

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

B.2 Individual heterogeneity

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

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

Consider the covariance matrix of initial conditions Σ :

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

and define,

$$\begin{aligned} R_T(\alpha) &= \begin{pmatrix} \begin{pmatrix} I_p & 0 \end{pmatrix} \\ M_T(\alpha) \end{pmatrix}^{-1} \\ S_{T,p}(\psi,\Lambda) &= \begin{pmatrix} I_p & 0_{p,T+q} \\ 0_{T,p} & \Lambda.M_T(\psi) \end{pmatrix} \end{aligned}$$

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

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

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

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

and

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

Hence,

$$\begin{aligned}\Omega_y &= R_T(\alpha).S_{T,p}(\psi,\Lambda).\Omega_z.S_{T,p}(\psi,\Lambda)'R_T(\alpha)' + M(\beta)^{[1-p,T]}\Sigma_\eta M(\beta)^{[1-p,T]\prime} \\ &+R_T(\alpha).\begin{pmatrix} 0_{p,p} & \Gamma_{0\eta}M(\beta)^{[1,T]\prime} \\ 0_{T,p} & 0_{T,T}\end{pmatrix} + \begin{pmatrix} 0_{p,p} & 0_{p,T} \\ M(\beta)^{[1,T]}\Gamma'_{0\eta} & 0_{T,T}\end{pmatrix} R_T(\alpha)'\end{aligned}$$

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 does not appear in the correlation between the two processes since innovations are supposed to be independent with individual heterogeneity factors. Initial conditions are given by $\zeta_i^{[1-q,0]}$, η^c and $v_i^{[1-p,0]}$.

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

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

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

C Fixed Effects, Constrained Effects and Counterfactuals

C.1 Estimates of individual factors given observed wages

The main equation is:

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

where η_i^c and $v_i^{[1-p,T]}$ 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 B.

Later on, we shall reintroduce the estimated averages, $\overline{\eta}_g$, 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:

$$\bar{\eta}_g = (M(\beta)^{[1-p,T]\prime}M(\beta)^{[1-p,T]})^{-1}M(\beta)^{[1-p,T]\prime}\bar{y}_g^{[1-p,T]}.$$

We now present the fixed effect estimation of η_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 η_i.

C.2 Estimating individual effects

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

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

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

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

and:

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

This yields the estimating equation for η_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:

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

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

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

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

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

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

where the constant of integration is derived by setting to one the integral over η_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 η_i^c, these constraints can be included in the prior (see below).

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

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

By identifying quadratic terms in (η_i^c, η_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:

$$\left\{ \begin{array}{rl} D'\Omega_w^{-1}D & = \Omega_\eta^{-1}, \\ -D'\Omega_w^{-1} & = -\Omega_\eta^{-1}B, \\ \Omega_w^{-1} & = B'\Omega_\eta^{-1}B + A, \end{array} \right.$$

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

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

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:

$$\begin{array}{rcl} V(\tilde{\eta}_i^c) & = & EV(\tilde{\eta}_i^c \mid \eta_i^c) + VE(\tilde{\eta}_i^c \mid \eta_i^c) \\ & \Longrightarrow & V(\tilde{\eta}_i^c) = B\Omega_w B' + V\eta_i^c = \Omega_\eta + V\eta_i^c. \end{array}$$

The term Ω_η goes to zero at least at the rate $1/T$ since matrix D is determined by different factors. Some are going to zero faster than T but they are dominated by the simple factors.

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 $(T_i, T+p+1)$ selecting non missing values and where T_i is the number of such non missing values. Consequently, the distribution function $L(\eta_i^c \mid S_i u_i^{[1-p,T]})$ becomes:

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

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

$$\left\{ \begin{array}{rl} \Omega_{\eta i} & = (D'S_i'(S_i\Omega_w S_i')^{-1}S_i D)^{-1}, \\ B_i & = (D'S_i'(S_i\Omega_w S_i')^{-1}S_i D)^{-1}D'S_i'(S_i\Omega_w S_i')^{-1}. \end{array} \right.$$

C.3 Constrained estimator

We reconsider the uncentered version of the individual effects η_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 η_i^c.

Using that the likelihood function $L(\eta_i \mid y_i^{[1-p,T]})$ 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 η_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 μ_1, μ_2 and μ_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:

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

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

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

   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:

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

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

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

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

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

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:

$$\tilde{\eta}_i - \hat{\eta}_i = \begin{pmatrix} \tilde{\eta}_{i1} - \hat{\eta}_{i1} \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix} = \frac{\Omega_\eta}{2} \begin{pmatrix} e_2 & e_3 \end{pmatrix} \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix}$$

where υ_j are unknown. They are obtained using:

$$\begin{pmatrix} e_2' \\ e_3' \end{pmatrix} (\tilde{\eta}_i - \hat{\eta}_i) = \begin{pmatrix} e_2' \\ e_3' \end{pmatrix} \begin{pmatrix} 0 \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix} = \begin{pmatrix} e_2' \\ e_3' \end{pmatrix} \frac{\Omega_\eta}{2} \begin{pmatrix} e_2 & e_3 \end{pmatrix} \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix}$$

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

$$\begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix} = \left[I_c' \frac{\Omega_\eta}{2} I_c \right]^{-1} I_c' \begin{pmatrix} 0 \\ -\hat{\eta}_{i2} \\ -\hat{\eta}_{i3} \end{pmatrix}$$

so that we get the vector:

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

C.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_η the Choleski decomposition of the permutation of matrix Ω_η (or $\Omega_{\eta i}$ in the case of missing values) such that:

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

Furthermore, it is convenient to slightly change the order of ηs without loss of generality. Assuming that the generic element of the lower diagonal matrix C_η is c_{ij}, we can write, assuming that the expectation of η_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 ξ_1, ξ_2 and ξ_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 C.2):

$$\begin{array}{c} \xi_1 > -\frac{\alpha_2}{c_{11}}, \\ \xi_2 + \frac{c_{21}}{c_{22}}\xi_1 < -\frac{\alpha_3}{c_{22}}, \\ \xi_2 + \frac{c_{21}+\pi_T c_{11}}{c_{22}}\xi_1 > -\frac{\alpha_3+\pi_T\alpha_2}{c_{22}}. \end{array} \tag{C.8}$$

The algorithm proceeds by considering initial values (η_2^0, η_3^0) whose construction we detail below. Then from (η_2^k, η_3^k), we construct $(\eta_2^{k+1}, \eta_3^{k+1})$ using:

1. Draw ξ_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 ξ_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 ξ_3 in a N(0,1) variate since no constraints are binding on η_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 ξ_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 ξ_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 ξ_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.

Table 1: Sample size

	Age of Entry			
	Below 20	Between 20 and 23	Above 23	All
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: 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 3: Autocorrelation matrix of earnings residuals

	1978	1979	1980	1982	1984	1985	1986	1987	1988	1989	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
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

Table 4: 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 5: AIC criterion

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

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

Table 6: Estimated parameters

	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
α_1	.702	.729	.711	.263	.186	.220	.200	.203	.194
	(.005)	(.006)	(.007)	(.011)	(.011)	(.011)	(.012)	(.011)	(.011)
α_2				.145	.324	.143	.191	.143	.161
				(.004)	(.008)	(.009)	(.005)	(.009)	(.009)
α_3							.022	.087	.187
							(.003)	(.004)	(.008)
ψ_1	.369	.391	.373	- .091	- .172	- .135	- .164	- .166	- .189
	(.005)	(.005)	(.007)	(.011)	(.011)	(.012)	(.012)	(.011)	(.011)
ψ_2		.020	.017		.170	- .028		- .046	- .046
		(.003)	(.003)		(.006)	(.008)		(.008)	(.008)
ψ_3			- .012			- .080			.114
			(.004)			(.004)			(.007)
σ_{η_1}	.302	.302	.301	.310	.306	.304	.306	.300	.298
	(.001)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.004)
σ_{η_2}	.038	.039	.039	.038	.039	.036	.038	.037	.037
	(.005)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
σ_{η_3}	.255	.259	.256	.263	.260	.248	.258	.247	.242
	(.005)	(.006)	(.006)	(.004)	(.005)	(.005)	(.005)	(.006)	(.007)
ρ_{η_1,η_2}	.473	.413	.454	.571	.486	.610	.505	.485	.365
	(.016)	(.021)	.021	(.013)	(.017)	(.013)	(.017)	(.020)	(.030)
ρ_{η_1,η_3}	- .604	- .548	- .586	- .694	- .618	- .729	- .636	- .620	- .509
	(.003)	(.020)	.019	(.011)	(.015)	(.012)	(.016)	(.019)	(.029)
ρ_{η_2,η_3}	- .946	- .948	- .947	- .945	- .946	- .941	- .946	- .943	- .944
	(.023)	(.003)	.003	(.002)	(.002)	(.003)	(.002)	(.003)	(.004)
σ_{y_0}	.491	.506	.496	.448	.479	.429	.442	.455	.494
	(.000)	(.007)	(.007)	(.004)	(.005)	(.004)	(.004)	(.005)	(.008)
$\sigma_{y_{-1}}$				.381	.424	.359	.387	.386	.428
				(.004)	(.005)	(.004)	(.004)	(.005)	(.008)
$\sigma_{y_{-2}}$							.264	.270	.299
							(.004)	(.006)	(.008)
$cov(\eta_1, y_0)$	- .227	- .257	- .237	- .156	- .214	- .149	-.186	- .201	- .282
	(.019)	(.017)	.017	(.015)	(.016)	(.016)	(.016)	(.017)	(.019)
$cov(\eta_1, y_{-1})$				- .127	- .183	- .113	- .153	- .168	- .253
				(.016)	(.017)	(.017)	(.017)	(.018)	(.020)
$cov(\eta_1, y_{-2})$							- .169	- .185	- .267
							(.018)	(.019)	(.022)
$cov(\eta_2, y_0)$	.358	.402	.374	.232	.335	.155	.219	.253	.361
	(.022)	(.020)	.021	(.017)	(.019)	(.021)	(.020)	(.022)	(.026)
$cov(\eta_2, y_{-1})$				.218	.331	.119	.242	.235	.352
				(.019)	(.021)	(.024)	(.022)	(.025)	(.029)
$cov(\eta_2, y_{-2})$							.239	.253	.351
							(.024)	(.027)	(.032)
$cov(\eta_3, y_0)$	- .290	- .333	- .305	- .179	- .270	- .107	- .163	- .195	- .291
	(.018)	(.023)	.023	(.020)	(.022)	(.023)	(.023)	(.024)	(.029)
$cov(\eta_3, y_{-1})$				- .169	- .272	- .077	- .190	- .181	- .287
				(.021)	(.023)	(.025)	(.023)	(.027)	(.032)
$cov(\eta_3, y_{-2})$							- .181	- .194	- .282
							(.026)	(.029)	(.035)
$cov(y_0, \zeta_0)$	.809	.036	- .024	- .823	.826	- .931	.841	- .795	.812
	(.023)	(8.525)	26.529	(.269)	(.059)	(.207)	(.061)	(.416)	(.096)
$cov(y_0, \zeta_{-1})$		.779	- .012		.408	- .352		- .208	.361
		(.438)	1.245		(.102)	(17.542)		(152.666)	(31.114)
$cov(y_{-1}, \zeta_{-1})$			.798		.722	- .066		.830	.234
			(.813)		(.062)	(.148)		(41.955)	(17.858)
$cov(y_0, \zeta_{-2})$						- .805			- .719
						(3.931)			(76.705)
$cov(y_{-1}, \zeta_{-2})$						- .382			- .202
						(11.249)			(44.061)
$cov(y_{-2}, \zeta_{-2})$									.752
									(.094)

Table 7: 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	.312	.312						
	(.001)	(.002)	(.002)						
1979	.254	.257	.255	.222	.232	.219			
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)			
1980	.223	.223	.223	.222	.227	.221	.224	.224	.230
	(.005)	(.001)	(.001)	(.001)	(.001)	(.001)	(.002)	(.002)	(.002)
1981	.264	.260	.263	.000	.103	.002	.004	.006	.001
	(.005)	(.005)	(.005)	(.096)	(.040)	(.066)	(.082)	(.076)	(.060)
1982	.152	.150	.150	.194	.193	.197	.193	.195	.198
	(.005)	(.005)	(.005)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
1983	.244	.243	.247	.040	.175	.096	.023	.039	.193
	(.004)	(.005)	(.005)	(.063)	(.017)	(.037)	(.048)	(.049)	(.021)
1984	.154	.149	.149	.189	.184	.187	.188	.188	.182
	(.001)	(.004)	(.004)	(.002)	(.001)	(.002)	(.001)	(.001)	(.002)
1985	.182	.182	.182	.181	.183	.183	.181	.183	.183
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1986	.187	.187	.187	.189	.189	.190	.190	.190	.192
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1987	.181	.182	.181	.176	.176	.177	.176	.177	.177
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1988	.180	.180	.181	.181	.181	.181	.181	.182	.183
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1989	.171	.172	.172	.168	.170	.169	.169	.170	.171
	(.008)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1990	.012	.021	.005	.358	.303	.375	.349	.395	.363
	(.002)	(.007)	(.008)	(.012)	(.008)	(.015)	(.012)	(.016)	(.013)
1991	.182	.184	.180	.153	.167	.156	.161	.157	.163
	(.001)	(.002)	(.002)	(.002)	(.001)	(.002)	(.001)	(.002)	(.001)
1992	.162	.162	.162	.159	.155	.159	.157	.160	.161
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1993	.207	.207	.207	.209	.209	.209	.210	.209	.211
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1994	.237	.236	.237	.250	.250	.251	.252	.253	.254
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1995	.193	.195	.194	.177	.179	.177	.177	.178	.180
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1996	.177	.177	.177	.176	.178	.177	.177	.177	.178
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1997	.167	.167	.167	.162	.162	.162	.162	.162	.164
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1998	.137	.138	.138	.134	.137	.135	.135	.136	.138
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
1999	.152	.152	.152	.155	.157	.157	.156	.157	.158
	(.001)	(.001)	(.001)	(.000)	(.000)	(.000)	(.000)	(.000)	(.001)
2000	.159	.159	.159	.159	.159	.159	.159	.159	.160
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
2001	.158	.158	.158	.159	.159	.160	.159	.160	.161
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
2002	.153	.153	.153	.146	.146	.146	.146	.147	.149
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
2003	.168	.167	.168	.178	.178	.179	.179	.180	.181
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
2004	.147	.148	.148	.133	.133	.134	.133	.134	.135
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
2005	.128	.128	.128	.130	.132	.130	.131	.131	.133
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
2006	.123	.124	.123	.124	.124	.124	.125	.125	.127
	(.001)	(.001)	(.001)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
2007	.117	.117	.117	.115	.116	.116	.115	.117	.118
	(.003)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)

Table 8: Short term inequalities and their decomposition

	Short term	Decomposition	
		Perm. (%)	Trans. (%)
Full sample			
1977	.167	.033	.966
1981	.095	.336	.663
2007	.151	.886	.113
Mean	.129	.648	.351
Age of entry < 20			
1977	.195	.008	.991
1981	.089	.225	.774
2007	.113	.872	.127
Mean	.104	.574	.425
Age of entry ≥ 20 and < 24			
1977	.121	.084	.915
1981	.091	.432	.567
2007	.187	.900	.099
Mean	.154	.682	.317
Age of entry ≥ 24			
1977	.134	.197	.802
1981	.125	.606	.393
2007	.276	.889	.110
Mean	.233	.721	.278

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 due to the transitory component.

Individual effects	Sample periods	Quantiles 0.05	0.2	0.35	0.5	0.65	0.8	0.95
	(3,15]	-2.97	0.483	1.74	2.43	3.04	4.31	8.14
		(2.31)	(0.603)	(0.194)	(0.0557)	(0.132)	(0.488)	(1.97)
	(15,22]	1.52	2.11	2.32	2.51	2.7	3	3.8
η_1		(0.0708)	(0.023)	(0.0191)	(0.0221)	(0.0208)	(0.031)	(0.0522)
	(22,26]	2	2.25	2.39	2.52	2.67	2.87	3.46
		(0.0226)	(0.0182)	(0.0161)	(0.0159)	(0.0175)	(0.021)	(0.0345)
	(26,28]	2.13	2.35	2.47	2.58	2.7	2.86	3.26
		(0.024)	(0.0159)	(0.0147)	(0.0136)	(0.0138)	(0.0151)	(0.0318)
	(3,15]	-0.435	-0.162	-0.0538	0.0203	0.0937	0.205	0.539
		(0.17)	(0.0478)	(0.0182)	(0.00945)	(0.013)	(0.0412)	(0.148)
	(15,22]	-0.124	-0.032	0.00555	0.033	0.0601	0.0973	0.194
η_2		(0.00983)	(0.00439)	(0.00321)	(0.00313)	(0.00322)	(0.00415)	(0.00951)
	(22,26]	-0.0471	-0.00103	0.0203	0.0388	0.0567	0.0834	0.141
		(0.00396)	(0.00264)	(0.00246)	(0.00216)	(0.00245)	(0.00267)	(0.00554)
	(26,28]	-0.0218	0.00914	0.0254	0.0383	0.0526	0.073	0.114
		(0.00333)	(0.00221)	(0.00214)	(0.00207)	(0.00216)	(0.00252)	(0.00415)
	(3,15]	-6.27	-2.14	-0.706	-0.00853	0.751	2.24	5.71
		(2.11)	(0.576)	(0.147)	(0.0858)	(0.212)	(0.712)	(2.47)
	(15,22]	-1.44	-0.622	-0.324	-0.125	0.0874	0.395	1.29
η_3		(0.0847)	(0.033)	(0.0254)	(0.0221)	(0.0287)	(0.0315)	(0.121)
	(22,26]	-0.907	-0.443	-0.275	-0.142	-0.024	0.131	0.445
		(0.0469)	(0.019)	(0.0167)	(0.0174)	(0.0171)	(0.0171)	(0.0344)
	(26,28]	-0.632	-0.36	-0.242	-0.147	-0.0542	0.0559	0.269
		(0.0232)	(0.0165)	(0.0138)	(0.0136)	(0.0147)	(0.0163)	(0.0267)

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

Table 9: Quantiles of the distribution of individual effects: unconstrained estimates

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]	11	0.93	-12	0.093	-1.1	14
	(15)	(1.2)	(16)	(0.095)	(1.3)	(17)
(15,22]	0.5	0.057	-0.57	0.01	-0.09	0.83
	(0.081)	(0.011)	(0.11)	(0.0016)	(0.015)	(0.15)
(22,26]	0.14	0.011	-0.099	0.0038	-0.027	0.2
	(0.0073)	(0.0011)	(0.0091)	(0.00032)	(0.0024)	(0.018)
(26,28]	0.076	0.0043	-0.038	0.002	-0.013	0.09
	(0.0039)	(0.00058)	(0.0041)	(0.00015)	(0.00097)	(0.0066)
Complete sample	2.6	0.22	-2.8	0.024	-0.27	3.3
	(3.2)	(0.25)	(3.4)	(0.021)	(0.28)	(3.8)
Random effects	0.093	0.0059	-0.05	0.0015	-0.0093	0.066
	(0.0034)	(0.00049)	(0.0038)	(0.00011)	(0.00077)	(0.0058)

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 10: Estimates of the covariance of individual effects

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

Notes: See Table above

Table 11: Estimates of the covariance of individual effects: Bias-corrected

Restrictions $\longrightarrow$ Sample periods	$\eta_2 < 0$	$\eta_3 < 0$	$\eta_3 + \pi_T \eta_2 < 0$
(3,15]	0.065	0.073	0.092
	(0.014)	(0.015)	(0.013)
(15,22]	0.058	0.087	0.11
	(0.0094)	(0.013)	(0.014)
(22,26]	0.023	0.037	0.079
	(0.0049)	(0.0066)	(0.011)
(26,28]	0.0066	0.013	0.033
	(0.0026)	(0.0049)	(0.0076)

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.

Table 12: Frequencies of violations: single-dimensional restriction

Sample periods	P-values <0.10	0.05	0.01
(3,15]	0.18	0.14	0.09
	(0.01)	(0.0093)	(0.0093)
(15,22]	0.19	0.14	0.076
	(0.01)	(0.009)	(0.0073)
(22,26]	0.13	0.093	0.06
	(0.0077)	(0.0068)	(0.0054)
(26,28]	0.062	0.038	0.018
	(0.006)	(0.0047)	(0.0029)
Complete sample	0.13	0.096	0.058
	(0.0045)	(0.0038)	(0.003)

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.

Table 13: Frequencies of violations: global restriction

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

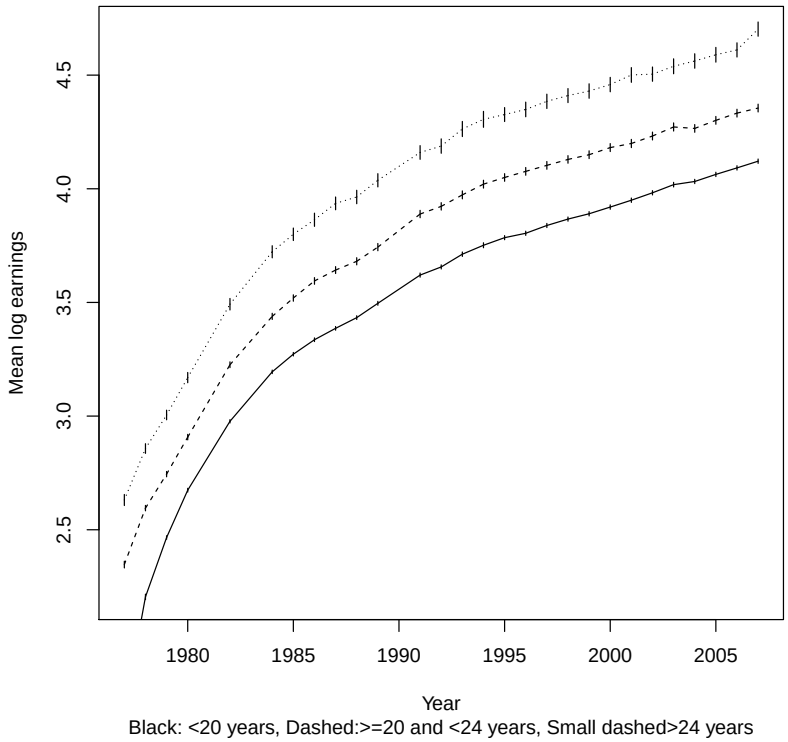
Notes: Distances use as a metric the inverse covariance matrix of ηs. Simulations are performed by adding to the constrained estimates a normal noise and by reprojecting on the constrained set.

Table 14: Distances between unconstrained and constrained estimates for observations and simulations

Summaries	Mean	Std error
Min.	0.0567	3.18e-05
1st Qu.	0.123	7.07e-05
Median	0.129	7.92e-05
Mean	0.14	0.000138
3rd Qu.	0.141	0.000154
Max	1.25	0.00662

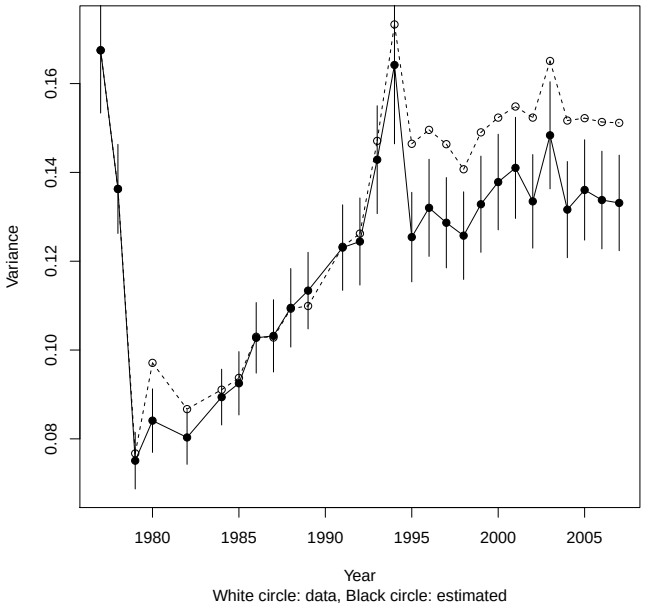
Notes: 4292 observations for which the number of periods is over 22.

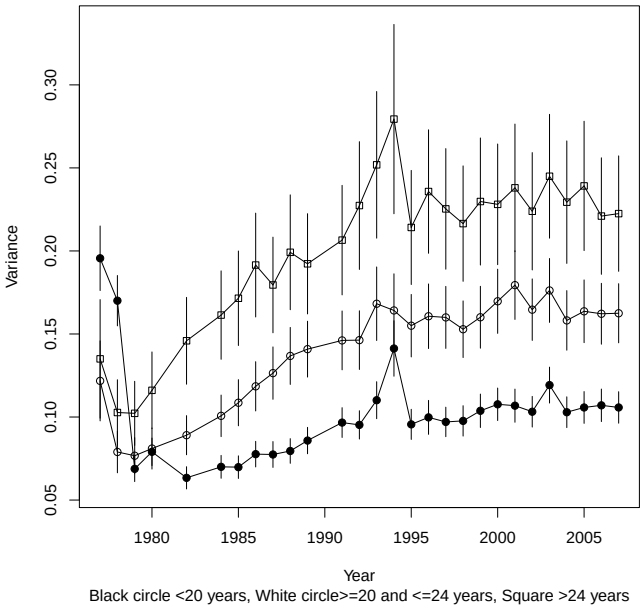
Table 15: Distribution of the returns to investment (lower bound)



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

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





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

(A) full sample (B) by age of entry

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

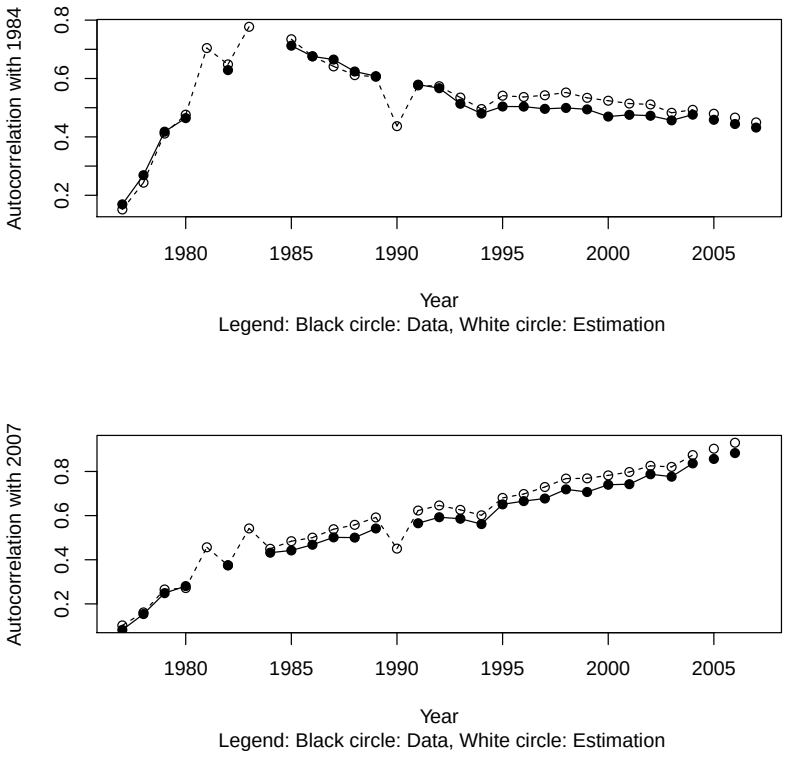


Figure 3: Autocorrelations with 1986 and 2007

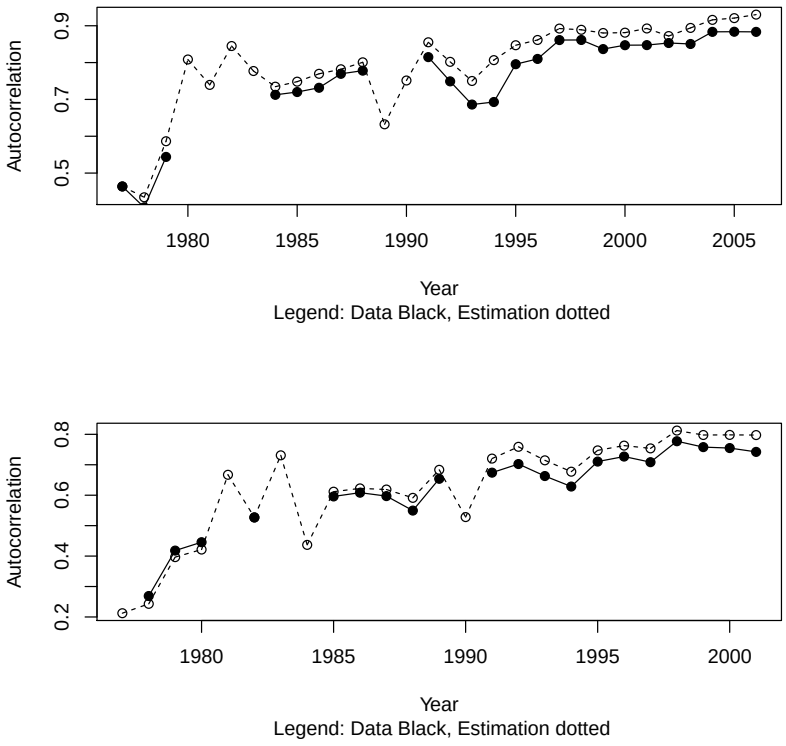


Figure 4: Autocorrelations of order 1 and of order 6

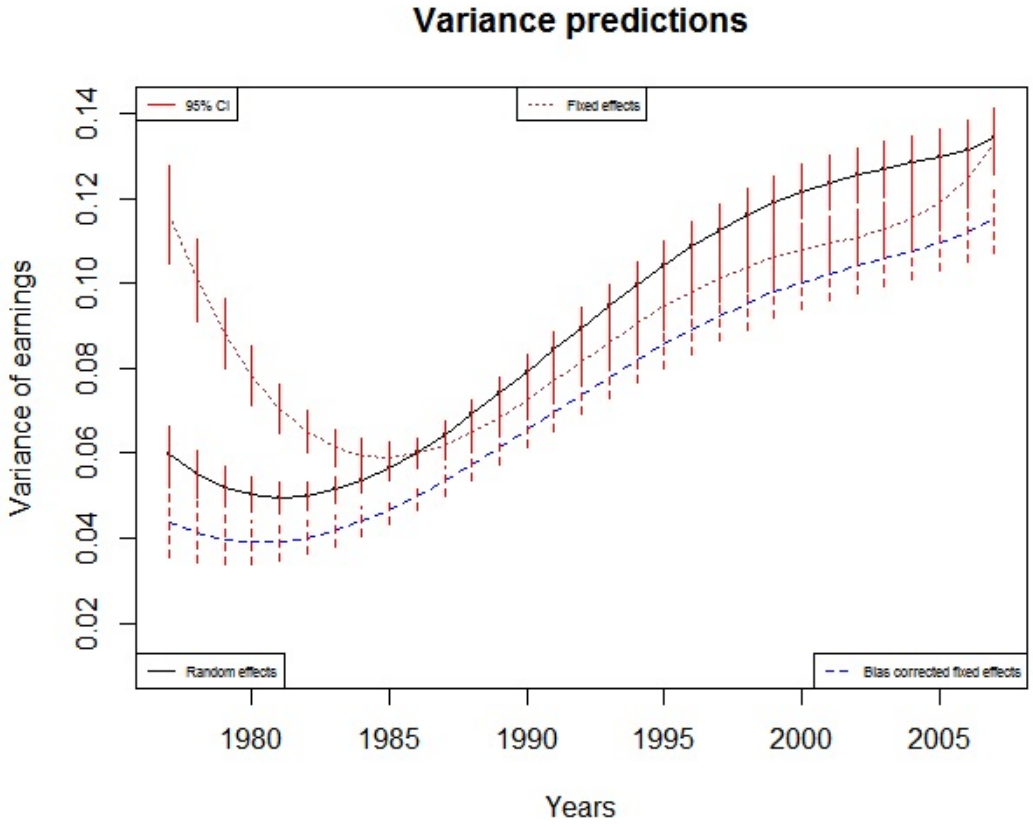


Figure 5: Random, fixed effect and biased-corrected fixed effect predictions of earnings variances using permanent components

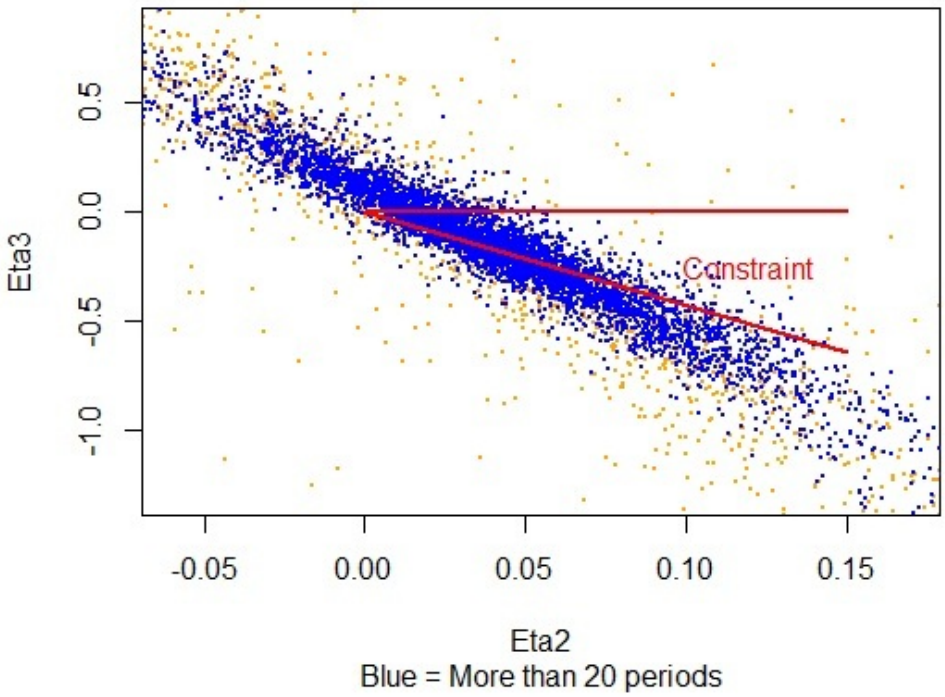


Figure 6: Scatter plot of η_2 and η_3 and the area describing the structural constraint

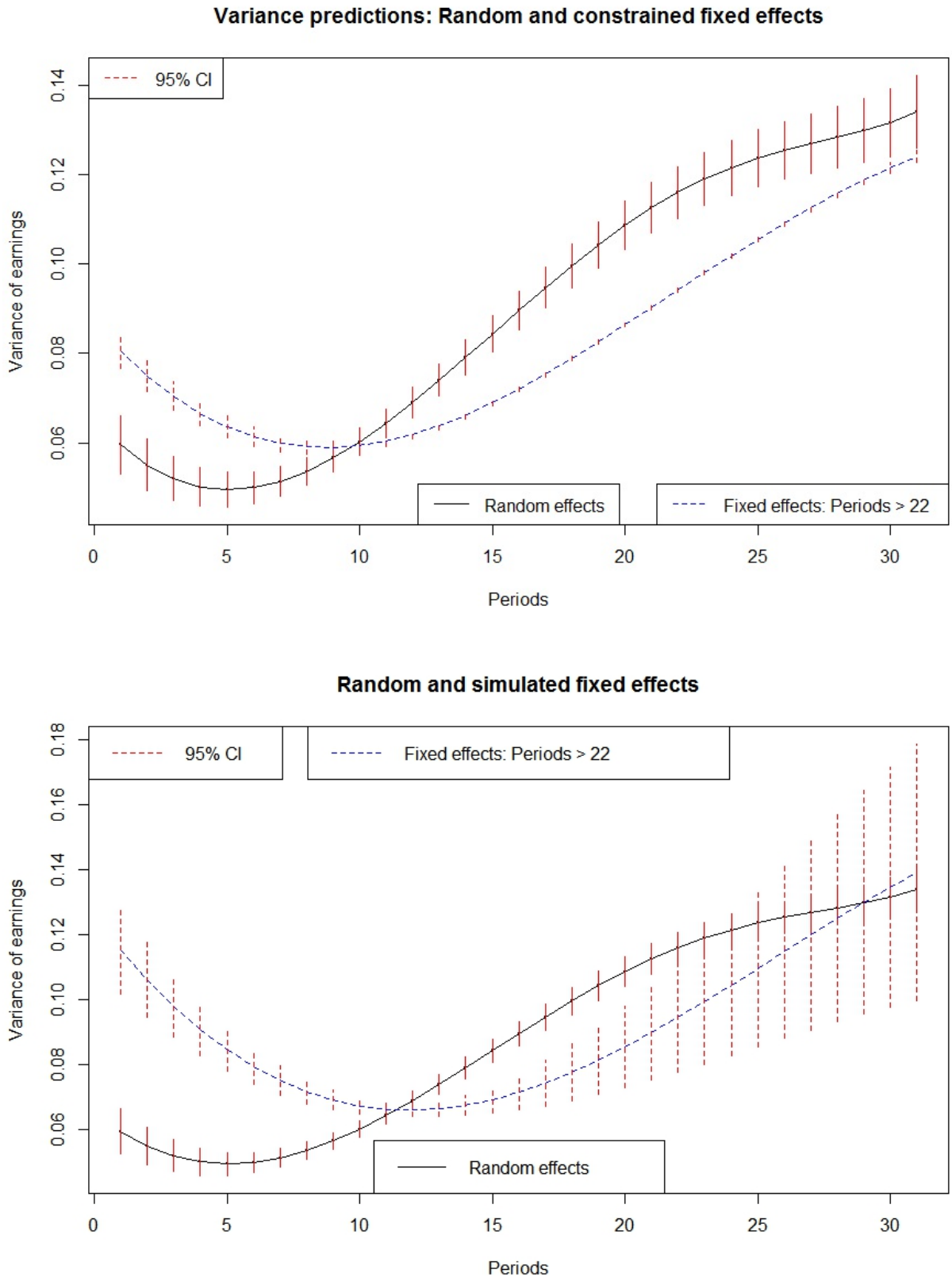


Figure 7: Earnings variances (permanent components): Constrained estimates and simulated constrained estimates

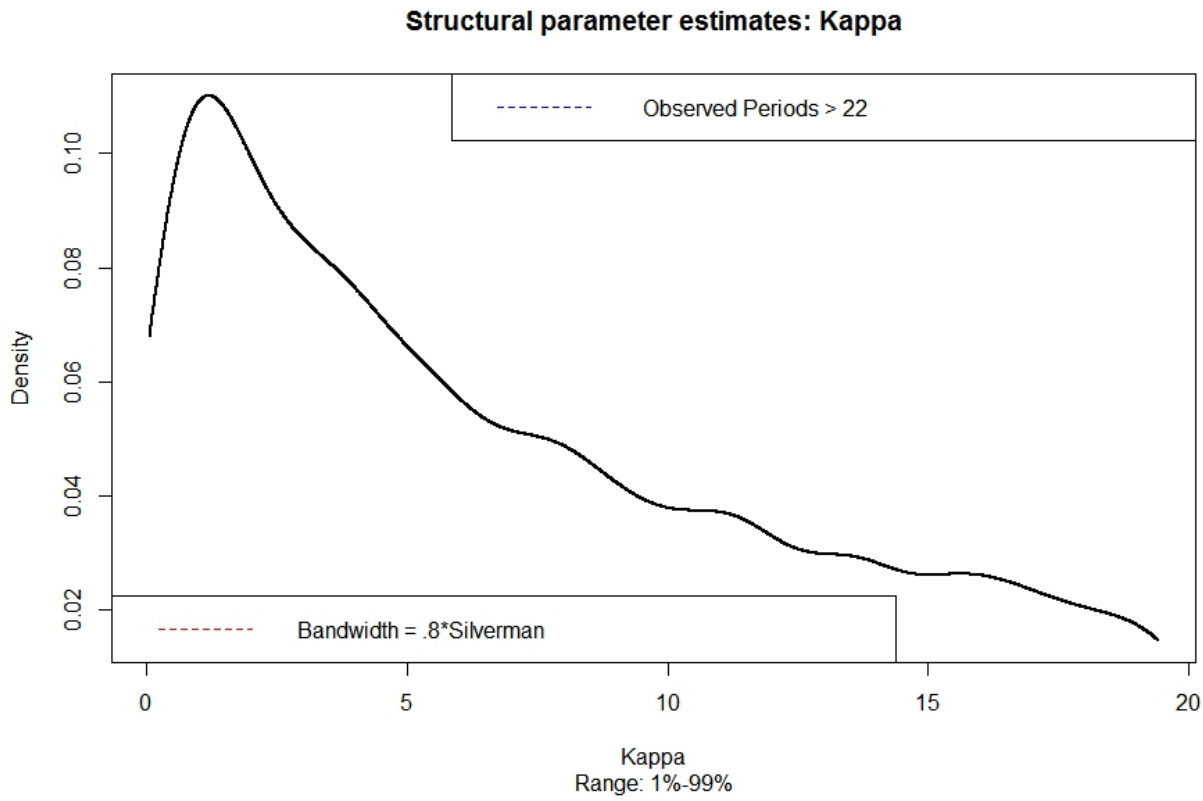
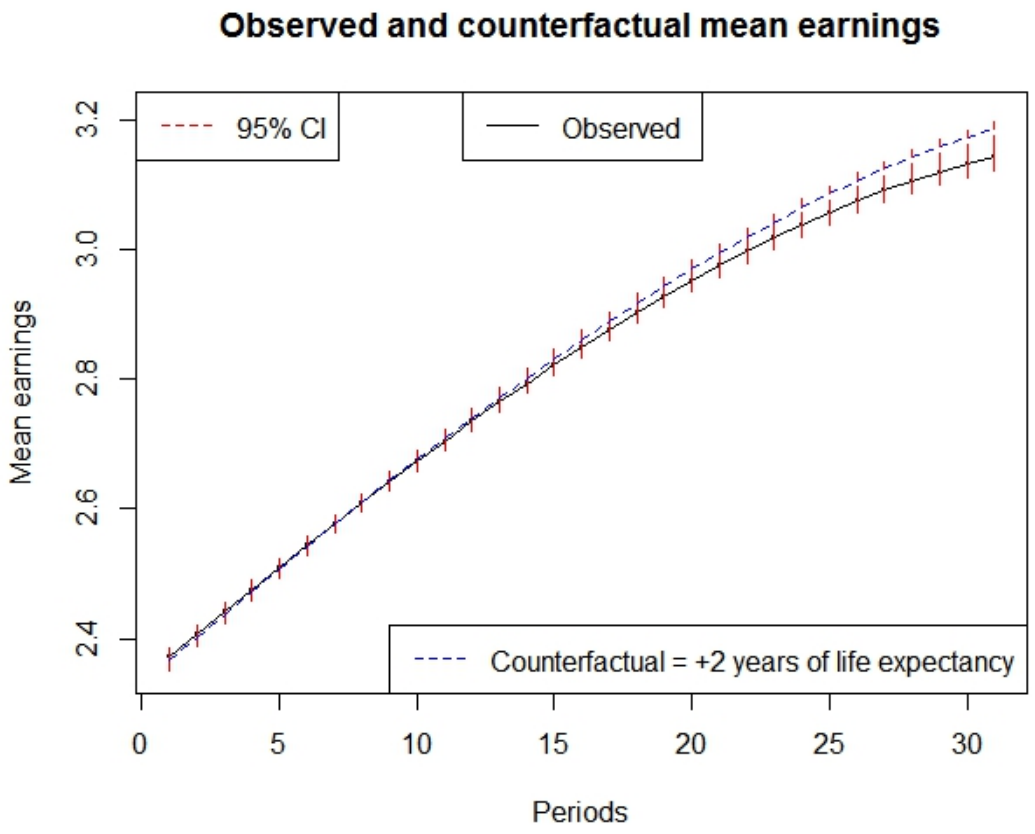
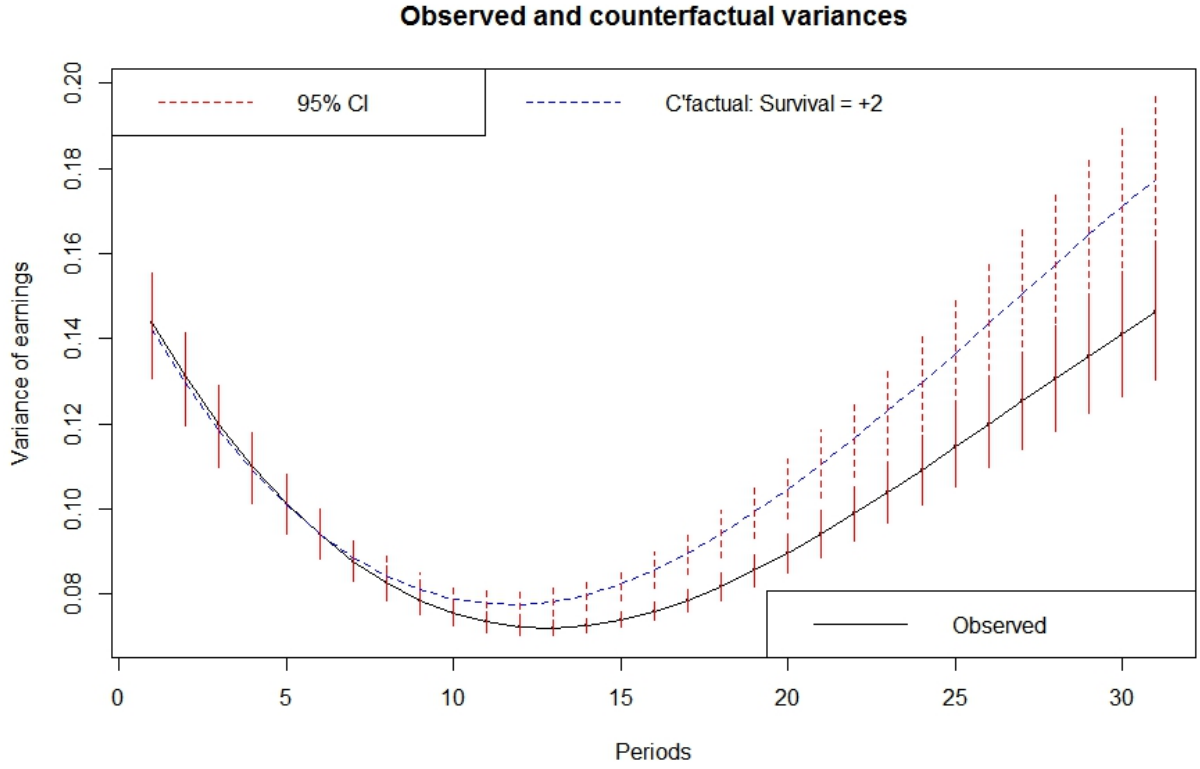


Figure 8: The density of the terminal capitalized discount rate κ





Note: Sample of 4292 observations for which observed periods>22. Standard errors are due to sampling and parameter uncertainty (30 Monte Carlo replications)

Figure 9: Counterfactual: Additional Years of Life Expectancy (K=2), Mean (Top panel) and Variance (Bottom Panel) Lower bound Impact[75]