

Heterogeneity and Wage Inequalities over the Life Cycle

Thierry Magnac* and Sébastien Roux†

This version: 19 March 2021

Abstract

Using data on French male wage workers observed over 30 years, we estimate by random and fixed effect methods a wage equation with pervasive heterogeneity. Individual wage profiles are derived from a human capital investment model and described by a level, a slope and a curvature. Among others, our empirical application delivers original results on issues like the Mincer dip, and the time-varying correlations between wage growth and levels, or between initial wages and growth. Static and long-run inequality indices can easily be compared and decomposed into their multidimensional components.

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

Keywords: human capital investment, inequality, wage dynamics, post-schooling wage growth, random and fixed effects

*Toulouse School of Economics, Université de Toulouse Capitole, *corresponding author*, thierry.magnac@tse-fr.eu

†Insee-Crest-Ined, Paris, sebastien.roux@insee.fr

1 Introduction¹

Income inequalities have returned to the center of many debates in economics, as rekindled by the work of Piketty (2013) and his co-authors on the sharp increase of top incomes over the last 50 years in most countries, although the data on labor earnings inequalities at the national level, in developed countries (Atkinson and Morelli, 2014) or at the global level (Milanovic, 2016) are slightly more contrasted. Among economic factors, the heterogeneous evolution of wage rates over the life cycle within a generation is at the heart of changes in inequalities of opportunity (Lagakos, Moll, Porzio, Qian and Shoellman, 2018, for a international comparison). Namely, wages are determined by labor market conditions facing each cohort (Jeong, Kim and Manovskii, 2015) and by individual abilities to earn and learn, while household labor earnings are ultimately derived from a sequence of decisions like marriage, labor market participation and labor supply which depend on wages and preferences (Chiappori, Costa Dias and Meghir, 2018).

Various competing theories can explain the shape of wage profiles over the life cycle and investments in human capital, learning by doing, or job search are their most popular underpinnings (Rubinstein and Weiss, 2006). It is however difficult to test them apart, especially when empirical researchers allow the parameters governing these models to be individual specific. These theories are nonetheless useful for disciplining the specification of wage profiles, for the interpretation of empirical results when estimating models of wage profiles, and therefore, for explaining the formation of inequalities over the life cycle.

In this article, we estimate the individual specific shape of (log) wage profiles – described by level, slope and curvature – as derived from a tractable model of investment in human capital à la Ben Porath (1967). We use French administrative data on a large cohort of around 7,500 men, who entered the labor market in 1977 and are followed until 2007. We focus on a single cohort, first to analyze pure life cycle issues and second, for the instrumental reason that it

¹We thank Christian Belzil, Richard Blundell, Laurent Gobillon, Jim Heckman, Nicolas Pistoiesi, Bernard Salanié, the editor and referees for helpful comments as well as participants in numerous seminars where we presented earlier versions of this research. This research has received financial support from the European Research Council under the European Community’s Seventh Framework Program FP7/2007-2013 grant agreement N°295298, funding from ANR under grant ANR-17-EURE-0010 (Investissements d’Avenir program) as well as funding from the Institut Universitaire de France. All errors remain ours.

allows this cohort to be followed over thirty years. We checked using data from surrounding cohorts that this cohort has nothing special. Another argument is that France is one of the countries in which earnings inequalities have remained stable over these years (Atkinson and Morelli, 2014) notwithstanding more unequal top incomes. Public policies have been found to be responsible for the decoupling between the evolution of wages and labor costs, the latter being unsurprisingly the same as in other industrialized countries such as the United States (Bozio, Bréda and Guillot, 2020).

Our theoretical set-up, as developed in Magnac, Pistoiesi and Roux (2018), leads to an empirical linear factor model in which a (log) wage profile is the sum of a profile of human capital stocks (in logs) and human capital (log) prices. The former is a linear function of three individual-specific parameters, one for the level, one for the slope and one for the curvature. The slope coefficient is expected to be positive, and the curvature negative, since profiles are mostly increasing and concave. Human capital (log) prices are further decomposed into aggregate prices that we identify using a "flat spot" approach (Heckman, Lochner and Taber, 1998, Bowlus and Robinson, 2012) over 1977-2007. The remaining individual-and-time (log) prices are specified as general ARMA processes. This empirical strategy allows the profile of individual human capital stocks (in logs) to be identified in expectation from wage data and compared across individuals in order to explain the formation of inequalities along the life cycle.

To estimate parameters of interest, we use an original sequence of random effect, first, and fixed effect procedures, second. A random effect method – i.e. considering the individual specific parameters as random – *(i)* deals with issues of initial conditions, likely to be quite off the stationary path in the case of wages, *(ii)* delivers an estimate of serial correlation of residual individual-and-time specific shocks and *(iii)* provides an estimate of the covariance matrix of individual specific parameters. We cannot however recover distributional characteristics of individual specific parameters other than their first and second order moments, by random effects. For this, we use fixed effect estimation methods that consist in estimating individual specific parameters although we discipline them using random effect estimates. Furthermore, missing

at random assumptions are weaker in a fixed effect setting. The price to pay with fixed effect methods, however, is the presence of an asymptotic bias in $1/T$ when estimating distributional characteristics such as variances or quantiles which can nonetheless be corrected using Jochmans and Weidner (2018).

Our analysis makes several empirical contributions.

First, modelling the curvature of wage profiles is important as forcefully argued by Lillard and Reville (1999) while most of the literature is specifying linear functions – level and slope only. The presence of the individual specific curvature, standing for horizon effects, is key in predictions about wage profiles. Its estimate is correlated negatively with both level and slope estimates so that the retraction force at the end of the life cycle, due to the shrinking horizon, is larger for those with large levels and large slopes. In this sense horizon effects act as a brake on inequalities.

Second, the correlations between estimated levels, slopes and curvatures obtained for skill aggregates, or individual specific parameters, are similar in France in contrast with the US (Heckman et al., 1998). The variance of the slope is also estimated to be smaller in France when compared to US studies suggesting that the building up of inequality over the life-cycle is smaller although it remains in the range of estimates reported by Lagakos et al. (2018) for various countries.

Fourth, our linear factor model is able to generate a Mincer dip without any role for transitory shocks whose variance is in fact slowly decreasing over time. As far as we know this is the first time a Mincer dip appears in a pervasive heterogeneity set-up. Fifth, the estimation of the correlation between wage and wage growth generated by our estimates confirms Rubinstein and Weiss’s conjecture about human capital – in contrast with learning and job search theories – that this correlation is first negative and then turns positive. Sixth, the estimation of the correlation between initial wage and subsequent wage growth, interpreted as returns, is negative at the beginning of the life cycle and then turns positive. This result stands in contrast with Gladden and Taber (2009) and Sorensen and Vejlin (2014) in which this correlation is uniformly negative.

Yet, there is no evidence of dynamic complementarities at least at the beginning of the life cycle.

Decomposing inequalities provides a final set of empirical contributions. First, we find that the variance of the long-run value of a wage profile, as accounted by discounted sums of log-wages over the (observed) life cycle, is smaller than the cross-section variance of log wages after 5 years (1982) by about 20% (.138/.113) and by around 80% (.207/.113) after 25 years. This is in line with estimates of Bonke, Corneo and Lutken (2014) using German data. The decomposition of variances by components tends to confirm the large share explained by permanent heterogeneity as was found in the literature (Guisar, 2007, Huggett, Ventura and Yaron, 2011). We confirm this result by providing further evidence on inter-decile ranges. Our next decomposition in terms of subcomponents of heterogeneity is also new to our knowledge. A single dimension heterogeneity term does not describe well the variance of log wages in cross sections and our multidimensional setting shows that curvature effects are getting much stronger at the end of the life-cycle.

Literature review It is useful to start with a brief comparison with the extensive empirical literature on earnings dynamics (see Meghir and Pistaferri, 2010, for a review). An important part of this literature aims at fitting the empirical covariance structure of (log) earnings over the life cycle using competing specifications like the one described as heterogeneous income profiles (HIP) or restricted income profiles (RIP). Up to now, there is no consensus in the literature about which specification fits the data best (see e.g. Baker, 1997, Guvenen, 2007, Hryshko, 2012 and Hoffmann, 2019). Our linear factor structure embeds both models since the permanent component includes individual specific levels and growth rates of earnings as HIP does and the stochastic component can be any mixture of permanent and transitory shocks like in RIP. Nonetheless, our three factor structure invalidates the key identifying assumption about the correlations between first differences of within shocks (for instance Blundell, 2014) because of the presence of the curvature term. Our structure is also used by Cho, Phillips and Seo (2019) to model conditional means, in a functional data set up.

Our paper also touches the estimation of the traditional homogeneous wage equation (Mincer,

1974). The state-of-the-art study is Lagakos et al. (2018) which studies an impressive set of countries and shows that experience-wage profiles are twice as steep in rich countries as in poor countries. Furthermore, more educated workers have steeper profiles. Yet, what we observe in our administrative data is similar to other studies (Engbom, 2017, using EHCP survey data) which find that wage growth in France over the life cycle is relatively small among 12 OECD countries, the same as in Germany but less than in the US and the UK.

Non-linear alternatives have been proposed in the recent literature such as Browning, Eijrnaes and Alvarez (2012), Polachek, Das and Thamma-Apiroam, (2015), Hospido (2012), Song, Price, Guvenen, Bloom, and Von Wachter (2018), Bowlus and Robin (2012) or Bonhomme and Robin (2009) as well as Pora and Wilner (2017) using French data. There are also semiparametric analyses such as Lochner and Shin (2014) and Arellano, Blundell and Bonhomme (2018) using US data. It is generally difficult to compare these alternatives with ours because our linear model is designed to capture means and covariances, while using pervasive heterogeneity. As a bridge between those methods and ours, we also estimate bias-corrected quantiles and inter-decile ranges.

In a different vein, a more economically oriented literature tries to contrast predictions derived from theories of wage growth, namely, human capital, job search or learning by doing. Rubinstein and Weiss (2006) takes stock of the literature before the 2000s, and distinguishes job search and human capital theories by their predictions. Job search models predict a negative correlation between wage and subsequent wage growth over the life cycle, while human capital models predicts that it is negative at the beginning of the life cycle but turns positive afterwards. The latter is what we find in our empirical analysis. Some recent literature explicitly models job search, in contrast with this paper in which we treat job search as a transient residual cause. Bowlus & Liu (2013) decomposes earnings growth into human capital (50%), job search (20%), the rest being their interaction. In contrast, Bagger, Fontaine, Postel-Vinay and Robin (2014) finds that job search, or "job-shopping", significantly contributes to wage growth although this seems to be mostly over the first ten years of the working life. Furthermore, Burdett, Carrillo-

Tudela and Coles (2016) finds that most of the effect of experience on wages is due to passive learning-by-doing.

Contrasting human capital investments and learning by doing is the objective of fewer papers. Heckman, Lochner and Cossa (2003) show that distinguishing job training and learning by doing might use that wage subsidies, such as EITC, provide additional incentives to work, enhance learning by doing and decrease investments in human capital. Belley (2017) contends that learning by doing does not seem prevalent since it does not predict the trade-off between current and future earnings, as observed in the data and predicted by human capital models. Blandin (2018) also points out that learning by doing does not predict a decrease in investments at the end of the working life.

Furthermore, within the human capital paradigm, Sanders and Taber (2011) reviews models with multidimensional human capital. More recently, Taber and Vejlin (2020) and Lise and Postel-Vinay (2020) estimate search and multidimensional human capital models, the former using a Ben Porath investment model, the latter a learning-by-doing framework. These papers restrict the heterogeneity dimension while our paper restricts the dimension of skills and models search frictions as residuals. (Hicks) aggregation of skills into a single index, at the individual level, would hold true if investment in skills keeps the composition of skills fixed over the life-cycle, and it is somewhat unclear whether we can tell the two settings apart. In theirs, the authors unsurprisingly find a greater importance of career shocks with respect to individual effects than we do.

Some reduced-form studies on life-cycle profiles are usefully compared to our results. Sorensen and Vejlin (2014) estimates the correlation between initial wages and later wage growth using Danish data over 20 years. They, as well as Gladden and Taber (2009), find that this correlation is negative. We find the same negative correlation over the first 20 years but it turns positive afterwards. Our approach differs from theirs in two aspects. First, we have a 3-factor linear model in which the curvature effect is key. Second, these authors use the observed initial wage, which is transient at this age, while we try to filter out these transient initial conditions. We

use the reconstructed non-transient initial log-wage by using a combination of random and fixed effects.

Finally, we only touch upon lightly the issue that returns to observable components increased dramatically, for instance in the US (Autor, Katz and Kearney, 2008) since this increase affects the cost of labor much more than wages in France (Bozio et al., 2020). Fernandez-Val, Peracchi, van Vuuren and Vella (2018) uses repeated cross-sections (US-CPS) and decomposes the evolution of quantiles, for males and females, into selection, structural and compositional changes. One of their conclusions is that selection effects seem small, as assumed here, notwithstanding a different panel data context. Our analysis also abstracts from compositional effects because of the single cohort and flat spot approach, and our analysis focuses on the structural effects affecting life cycle wage profiles.

Section 2 briefly describes the evolution of earnings inequality in France and the data we use and Section 3 details our empirical strategy. Section 4 reports estimation results and Section 5 gathers the results of various decompositions of life cycle inequalities.

Any reference to the Supplementary Appendices and Tables & Figures can be found in the working paper version (Magnac and Roux, 2021).

2 A Brief Description of the Data

We briefly summarize the evolution of wage inequality in France over the last 40 years and present stylized facts about means, variances and autocorrelations of log wages in our sample, after reporting how we constructed this sample from administrative sources.

2.1 Earnings and Wage Inequality in France

The sharp increase in earnings inequality in the UK and in the US over the last thirty years is a well known empirical fact (see for example Autor, Katz and Kearney, 2008, or Moffitt and Gottschalk, 2011, for the US and Blundell & Etheridge, 2010, for the UK). Yet, the picture is more balanced in other OECD countries and while some European countries have experienced

an increasing dispersion in earnings, others have not been affected by this trend and have had stable or decreasing dispersion. Atkinson and Morelli (2014) computes international earnings inequality comparisons over the second half, or so, of the twentieth century for 25 countries. As regards European economies, they conclude that earnings inequality has increased in Germany, Italy, Portugal, Sweden, Switzerland while in Finland, France, the Netherlands, Norway, and Spain earnings dispersion has stayed constant or decreased over this period.

In France, earnings inequality in 2010 is broadly comparable to its level in the nineteen sixties and if anything has decreased. Atkinson and Morelli (2014) report an unchanged Gini coefficient for earnings over the period. Using Labour Force Surveys (LFS), they also compute yearly measures of inequality and show a very stable inequality level. Using two different datasets, the DADS, used here, and the French LFS, Verdugo (2014) concludes that the two data sets provide strikingly similar figures of constant or decreasing earnings dispersion between 1964 and 2005. Verdugo (2014) decomposes the total earnings dispersion into upper and lower-tail earnings inequality. The dispersion at the top of the distribution remained constant, since the P90/P50 index in earnings fluctuates around 2, while the dispersion at the bottom measured by the P50/P10 index decreased from 1.9 to 1.5. Charnoz, Coudin and Gaini (2011) also use the DADS data to reach the same conclusion that earnings inequality in France has been rather stable from 1976 to 1992 and has been slightly decreasing from 1995 to 2004. This stability has been attributed, at least partly, to a strong policy driven increase in education at the end of the 1980s and labor market policy regulations at the end of the 1990s (Charnoz, Coudin and Gaini, 2014). Bozio, Breda and Guillot (2020) neatly demonstrate that inequality increases in labor costs have been the same as in the US and other major developed countries although French policy changes in payroll taxation, and in particular for low-skills, have counteracted these changes in terms of wage inequalities.

A note of caution is in order. While these studies consider changes in the cross-sectional earnings distributions, changes in the structure of the population and cohort effects that play an important role in the previous studies (e.g. Jeong, Kim and Manovskii, 2015) are neutralized

here. We follow a single cohort of males entering the labor market in 1977 to focus on inequalities unraveling over the life cycle.

2.2 Our Working Sample

Our panel dataset on wages is extracted from a French administrative source named *Déclarations Annuelles de Données Sociales* (DADS) which has been used for employee-employer studies as in Abowd, Kramarz and Margolis (1999).² DADS data are collected through a mandatory data requirement 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 every employee and payments exclude other labor costs borne by the firm. Each person is identified by a unique individual social security number which facilitates the follow-up of individuals through time although we cannot reconstruct taxes they pay. The tax system is household-based in France and the linking of this dataset with fiscal records is not authorized yet.

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

This dataset is restricted to individuals employed in the private sector, including publicly-owned companies, and we consider only males to alleviate selection issues. Observations can yet be missing for different reasons. Data were not collected in three years (1981, 1983 and 1990) for reasons specific to INSEE. It is also quite frequent that employees exit the panel and a significant fraction of those reenter it after a few years (see Table S.i in the Supplementary

²These data are available through securitized access (casd.eu). Other contributions in the earnings literature that use administrative data is Hoffmann (2019) and Daly, Hryshko and Manovskii (2016) (German and Danish data).

Appendix). Those absences might stand for spells in the public sector, as self-employed, or out-of-employment. We also code as missing, part-time employment in any given year.

We restrict our analysis to labor market entrants in 1977, a set that we call "cohort" in the following, even if their age is heterogenous. Entrants are defined as those who started working full time for more than 6 months in 1977, and are still employed the following year, possibly in a different firm. To make sure that these employees have a permanent attachment to the private sector and to mitigate the issue of missing years in 1981 and 1983, we keep only those who also worked in 1982 and 1984 and were aged between 16 and 30 in 1977. The characteristics of the distribution of individual specific parameters that we estimate in the following, refer to this subpopulation.

We define wages as the sum of all earnings during a year divided by the number of days worked. This allows employees to have within-year periods out of the private sector. Shortcomings of administrative data are that other components of income, including other sources of earnings, are missing and that few observable characteristics are available, apart from age at labor market entry and the skill level of the first job.

First, it is likely that workers delaying entry have a higher education level than the ones who entered earlier. Second, initial skills are grouped into three categories based on a two-digit codification: high-skill (managers, professionals), medium-skill (blue-collar or white-collar skilled workers) and low-skill jobs. Our 20 resulting "education" groups are defined by the interaction between these two variables when groups are not too small (see Table 2 for the definition and size of each group). Since education is defined according to characteristics recorded at labor market entry, individuals are attached to the same education group during their whole working life.

Other details completing this description can be found in Data Appendix A.

2.3 Wage dynamics over the life cycle

Table 1 reports descriptive statistics on the composition of the sample over time. The sample size is 7446 observations in 1977 and 4670 in 2007. Human capital groups defined above are

of unequal size, the groups with an early age of entry being the largest ones, and with a late age of entry, the smallest. Attrition follows a somewhat irregular pattern due in part to the original data and to our sampling design (see Supplementary Appendix, Table S.i). There are also more surprising features for instance in 1994 (or 2003 at a lesser degree) a year in which many observations are missing. This is due to the way INSEE reconstructed the data from the information in the original files.

We report in Figure 1, the evolution of average log-wages over the life cycle, in 2007-euros, for three age of entry groups (< 20 , ≥ 20 and ≤ 24 and > 24) and these profiles display the familiar increasing and concave shape.

By taking deviations of (log) wages with respect to their means in groups defined by age of entry, skills and years of observation, we compute log-wage residuals. The left panel of Figure 2 plots the change in the cross-sectional variance of those (log) wage residuals for the full sample, while the right panel graphs the variance by age of entry groups. Choosing the variance of log wages as a description of the process is adapted to the random effect specification that we estimate. Using other inequality measures (Gini , Theil or Atkinson) does not change the qualitative features of our description. The first few years witness a strong variability of wages. Until the sixth year of observation, 1982 (respectively the fourth, 1980), the variance of log wages drops for the low skill groups (resp. for the other groups) whereas it increases gradually afterwards till around 1995. The variance profile is flat after 1995 in contrast with the US where it continues to grow (e.g. Rubinstein and Weiss, 2006, using PSID). From the right panel one can notice that late entrants in the labor market experience higher levels and larger rates of growth of the variance of log-wages over the life cycle as in most countries (Lagakos et al, 2018). The covariance analysis of log-wage residuals over time is reported in Supplementary Appendix S.II. It provides strong evidence that wages are becoming more stable as employees progress in their life cycle.

3 Empirical strategy

3.1 The set-up

We specify a linear factor model of (log) wage dynamics. Our starting point is a model of human capital investments after leaving school that is developed in Magnac et al. (2018) and that provides exact theoretical foundations à la Ben Porath (1967), for a Mincer (1974) reduced form wage equation at the individual level. It assumes that consumption is intertemporally smoothed through human capital investments only, and that the accumulation equation is log linear as well as the value function at a terminal period. The assumption of decreasing returns to investment in Ben Porath is replaced by making investments subject to an increasing and convex utility cost interpreted as cost of effort, and utility is log linear in consumption. Departures from these assumptions are the object of the discussion in Magnac et al. (2018) that we do not repeat here.

In this set-up, the resulting (log) wage equation can be written as a linear factor model where the three observed factors are $f_t = (1, t, \beta^{-t})$:

$$\ln y_{it} = \eta_{i1} + \eta_{i2}t + \eta_{i3}\beta^{-t} + v_{it}, \quad (1)$$

in which y_{it} is the log wage, t is potential experience (i.e. $t = 1$ in 1977), i is the individual index and $\beta < 1$ is the discount factor, assumed homogenous. This reduced form, which is also a random coefficient model, delivers the familiar increasing and concave shape when $\eta_{i2} > 0$ and $\eta_{i3} < 0$. This specification is more tractable to estimate, than non-linear alternatives when heterogeneity is pervasive (Browning, Ejrnaes and Alvarez, 2012, Polachek et al., 2015).

The first three terms measure the logarithm of the current human capital stock (net of current investments) while v_{it} can be interpreted as the logarithm of the individual-specific price of human capital, net of accumulated depreciations. Implicitly, the influence of job search, the job ladder, or dismissals, are hidden in the latter component. This model nests the Heterogenous Income Profile (HIP) and the Restricted Income Profile (RIP) as in Guvenen (2007, 2009) since v_{it} can be any time-series process. We do not take a stand on how this would be decomposed

into persistent and transitory components since consumption data are not available so that the decomposition cannot be identified (Ejrnæs & Browning, 2014).

Factor loadings or individual specific effects η_{i1} , η_{i2} and η_{i3} have a structural economic interpretation. The first two ones are respectively related to levels of human capital — i.e. the ability to earn — and to growth rates of human capital — i.e. the ability to learn (see Browning et al., 1999). Furthermore, the ratio between η_{i3} and η_{i2} , i.e. the ratio of the "curvature" of the profile relative to the slope, can be structurally interpreted as the value given to human capital at an arbitrary terminal period. The longer the horizon of investment, the smaller the curvature (Lillard and Reville, 1999).

Estimating this very heterogenous reduced form leads to the decomposition of the heterogeneity of wage profiles in each of its structural components. Deriving the relative importance of each component is the object of Section 5 which extends the calibration exercise of Huggett et al. (2011) and other authors using US data. Moreover, it allows to predict wage profiles even if some observations are missing as in our data and we will return to this point below.

Before that, we present our empirical strategy. As a first step, we decompose log wages into aggregate and individual specific components to go beyond the usual demeaning of time and skill effects in the earnings dynamics literature (Meghir and Pistaferri, 2010). Aggregate components are constructed using the subsamples defined by the education groups that we constructed before, from skill and age at labor market entry. We adopt the equilibrium framework of Heckman et al. (1998), in which human capital stocks of different education groups are imperfect substitutes in the aggregate production function so that aggregate components can be interpreted as the sum of aggregate log human capital stocks of each skill group and the market (log) prices of skills net of depreciation. Moreover, the so called flat spot approach that Heckman et al. (1998) suggested, and that we present below using data from older cohorts, allows the identification of stocks from prices.

In contrast with aggregate components, perfect substitution holds within groups, and the remainder of individual specific shocks are interpreted as frictions orthogonal to aggregate com-

ponents and individual specific parameters. As a result, the mechanisms that underlie the specific dynamics of aggregate and individual specific components are allowed to differ, and are left unrelated. In consequence, we handle education group and individual specific effects separately and recombine them afterwards to recover the full effects.

3.2 Aggregate components: Identifying restrictions and estimation

Following the logic above, equation (1) can be linearly aggregated, in each education group, as:

$$\overline{\ln y_{gt}} = \bar{\eta}_{g1} + \bar{\eta}_{g2}t + \bar{\eta}_{g3}\beta^{-t} + \bar{v}_{gt}, \quad (2)$$

in which g denotes a group defined by skill and age of entry and $\bar{\eta}_{gk} = E(\eta_{ik} \mid i \in g)$ for $k = 1, 2$ or 3 , $\bar{v}_{gt} = E(v_{it} \mid i \in g)$. The term \bar{v}_{gt} stands for the market log-prices of human capital of group g at time t .

Denote φ_{gt} the unknown expected aggregate log prices :

$$\varphi_{gt} = E(\bar{v}_{gt} \mid f_t = (1, t, \beta^{-t})). \quad (3)$$

Identifying φ_{gt} using external sources is the key restriction that separates quantities from prices. If φ_{gt} is identified and subtracted to equation (2), the new error term is orthogonal to factors $(1, t, \beta^{-t})$ and parameters $\bar{\eta}_g$ are identified. To identify φ_{gt} (up to a constant term), Heckman et al. (1998) and Bowlus and Robinson (2012) use a flat spot method. They argue that investments and depreciation shocks for individuals close to the end of their working life (around 50) exactly balance each other so that the term $\bar{\eta}_{g2}t + \bar{\eta}_{g3}\beta^{-t} \simeq 0$ in equation (2). Estimates of φ_{gt} are therefore the mean wages of those older cohorts who are observed at the same calendar times t as our cohort of interest.³ This structural method is also a way to solving the well known impossibility of separately identifying age, cohort and time effects in a linear setting. In addition, the stability of the between and within group wage inequality in France over this period supports the credibility of this procedure.

We apply this flat spot technique and from now on, we will consider that log wages, $\overline{\ln y_{gt}}$, are net of these estimates $\hat{\varphi}_{gt}$. OLS estimation of equation (2) is therefore unbiased because of

³As t is potential experience, calendar time is $t + 1977$ and φ_{gt} is estimated using cohorts aged more than 50 at date $t + 1977$ (see end of Data Appendix A)

condition (3). We estimate equation (2) group by group using 28 observations per group. It provides consistent estimates of $\bar{\eta}_g$, say $\widehat{\bar{\eta}}_g$, and standard errors are computed using a Newey West procedure.

3.3 Individual specific components: identifying restrictions

Turning to within group variation, define centered individual effects as their deviations with respect to their skill specific means, $\eta_{ik}^c = \eta_{ik} - \bar{\eta}_{g(i)k}$, for $k = 1, 2$, or 3 and $v_{it}^c = v_{it} - \bar{v}_{g(i)t}$ where $g(i)$ is the skill group to which individual i belongs. Demeaning the log wage equation (1) with respect to their group averages yields log-wage deviations:

$$u_{it} = \ln y_{it} - \overline{\ln y}_{g(i)t} = \eta_{i1}^c + \eta_{i2}^c t + \eta_{i3}^c \beta^{-t} + v_{it}^c, \quad (4)$$

Individual specific deviations, v_{it}^c , stand for frictions (e.g. job ladder, firms' heterogeneity) in a model of search and mobility (see e.g. Postel-Vinay and Turon, 2010) as well as accidents of life. These frictions are here described by a stochastic process which is supposed to be mean independent of factors and factor loadings:⁴

$$E(v_{it}^c \mid f_t = (1, t, \beta^{-t}), \eta_i^c) = 0. \quad (5)$$

Note that it lets other moments of v_{it}^c depend freely on factors and individual effects η_i^c . The mean independence of frictions, v_{it}^c , with respect to factors, f_t , and individual effects, η_i^c , is the main assumption underpinning the identification of individual-specific structural parameters.

The issue of missing data is potentially important in our empirical application since we observe wages if and only if individuals are employed in the private sector. Missingness could be due to periods spent in the public sector, as self-employed or as non employed. Absent credible instruments or structural assumptions, we assume in the following that missingness is at random (MAR). The existence of missing data becomes unsequential for consistently estimating aggregate effects while we show below that the random effect likelihood approach we adopt, deals with missing data easily.

⁴In the empirical application below, we slightly relax the assumption of mean independence of frictions with respect to individual effects by authorizing general initial conditions in the random effect model.

3.4 Individual specific components: Estimation

At this step, our objective is to recover estimates of centered individual effects $(\eta_{i1}^c, \eta_{i2}^c, \eta_{i3}^c)$ in the linear factor structure (1) by using identifying restrictions presented in Section 3. Most of the technical details are relegated to appendices.

Stacking log wage residuals u_{it} , defined in equation (4), and the stochastic component v_{it}^c into $T \times 1$ -vectors u_i and v_i^c as well as η_{ik}^c into a 3×1 -vector η_i^c , equation (4) writes:

$$u_i = M(\beta) \eta_i^c + v_i^c. \quad (6)$$

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

We adopt a two-tier estimation strategy in which both tiers complement each other.

We first estimate a random-effect model in which the parameters of interests are the population covariance matrices of individual specific effects $V(\eta_i^c)$, and transitory errors $V(v_i^c)$, as well as allowing for some correlation between permanent and transitory terms, $E((v_i^c)' \eta_i^c)$, due to initial conditions. Identification conditions rely on the finite parameter specification of the covariance matrix $V(v_i^c)$, here time-heteroskedastic ARMAs, as shown in Arellano and Bonhomme (2012). The maximum likelihood estimation is taken from Arellano and Alvarez (2003) and is well adapted to panel data with missing observations. The details of this estimation method, and in particular all assumptions related to the time-heteroskedastic ARMA error process we estimate, and step by step construction of the likelihood function, are given in Appendix B and further completed in Supplementary Appendix S.III.

This random effect method is informative about second-order moments only. An alternative estimation method is based on fixed effects in which η_i^c are considered as parameters and the missing at random assumption is weakened to be conditional on parameters η_i^c . Those are estimated using equation (6) individual by individual. Furthermore, to improve precision, we estimate parameters by Feasible GLS in which the covariance matrix of v_i^c is the estimate obtained by the first-step random effect method.

Using estimated individual-specific parameters by a fixed effect method opens up the way to the analysis of other summaries of their distributions, like quantiles, and not only second order

moments. Yet there is a price to pay. As some data are missing, inference might be poor since there are at most 28 observations per individual in our data. If T_i is the number of observations for individual i , estimates of parameters η_i for every individual profile are consistent when $T_i \rightarrow \infty$. Nonetheless, an asymptotic bias of order $1/T_i$ arises for individual specific parameters and for most summaries like variance and quantiles (e.g. Arellano and Bonhomme, 2012) although means are exceptions. This asymptotic bias affects inference since it decenters confidence intervals. Nonetheless, correcting biases to a lower order of $1/T_i^2$ is easily obtained using methods developed by Jochmans and Weidner (2018). Appendix S.IV in the Supplementary Appendix provides a description of the estimation method as well as the correction method for asymptotic biases of variances as well as quantiles.

Arguments underpinning the two-tier estimation strategy are based on trading off the consistency properties of random effect methods when the time dimension is small and the flexibility of the fixed effect methods. On the one hand, using the covariance structure of the error terms estimated by random effects in the first step, makes the fixed effect estimation more precise. Furthermore, random effect estimation provides a benchmark against which we can assess the amount of bias in the fixed effect estimation due to the finite length of the observation period for each profile. On the other hand, fixed effect estimation brings about richer information on the underlying distribution of individual specific parameters and weakens the missing at random assumption.

Finally, we can combine aggregate and individual specific estimates yielding fixed effect estimates of the original factor loadings:

$$\hat{\eta}_i = \hat{\eta}_g + \hat{\eta}_i^c,$$

and we use these estimates to compute predictions of wage moments or quantiles in the next two Sections.

4 Results

We first describe the estimated parameters of the aggregate equation (2) in Section 4.1 and turn next to the estimation results of the within group wage equation by random effect methods in Section 4.2. In Section 4.3 we comment the estimates of summaries constructed from estimated individual effects. We wind up the section with comparisons between these results and predictions appearing in the literature.

4.1 Aggregate estimation

Details of the flat spot approach for estimating series of human capital prices, φ_{gt} , in equation (3) are presented at the end of the Data Appendix A as in Bowlus and Robinson (2012). In sum, human capital prices are estimated using a population of older males whose potential experience ranges from 25 to 40 (e.g. whose age stands between 43 and 58). These data are taken from older cohorts than our cohort of interest, at the same calendar times, and those prices are used to deflate log wages as described in Section 3.2.

As a result, Table 2 presents for each of the 20 groups the estimated aggregate group parameters from equation (2). They exhibit expected patterns. The first factor loading average $\bar{\eta}_{g1}$ ranges from 2.5 for the lowest skill groups to 3.5 for the highest skill groups – a 100% upward translation. The estimated average slope, $\bar{\eta}_{g2}$ ranges from 0 to 0.05 in the range of Mincer’s estimates (e.g. Lagakos et al., 2018). The slope is larger for the high-skilled groups than for the low-skilled ones although the evidence is weaker. The geometric factor loading average $\bar{\eta}_{g3}$ is mostly negative as expected, or non significantly different from zero.

4.2 Random effect estimates

We proceed as explained in Section 3.4. Our preferred random effect estimates are obtained when disturbances, v_{it}^c , are time-heteroskedastic ARMA(3, 1) a slightly longer dependence than usual.⁵ We start by reviewing our original results on the correlations between factor loadings,

⁵Results for general ARMA(p,q) processes in which p and q vary between 1 and 3 are reported in the Supplementary Appendix in Tables S.iv and S.v. Time-varying variances of the noise are decreasing with time (Table S.v). The values of the Akaike criterion are reported in Table S.iii. Increasing the order of the MA above 1

η_i^c , and contrast them with correlations across groups of aggregate components as reported in Table 2.

The estimated covariance matrix of the centered individual effects, η_i^c is stable across the different specifications of ARMA processes as shown in Table 3. Their standard deviations are very precisely estimated at around .30 for the "level" factor loading, η_{i1}^c , and .25 for the curvature one, η_{i3}^c , and at around .04 for the slope, η_{i2}^c . The following results are more telling. The correlation between the slope and curvature factor loadings is very strongly negative and equal to $-.95$ consistently across ARMA specifications. The magnitude of this correlation and its sign are consistent with the structural model that ties in parameters η_{i2} and η_{i3} : the larger the slope, the more curved the wage profile. The retraction force due to the horizon is stronger for high wage growth employees.

The correlation coefficient between the curvature, η_{i3}^c , and the level, η_{i1}^c , factor loadings is also significantly negative – around -0.6 – and reveals that the retraction force at the end of the life cycle is also stronger for highly skilled workers. The correlation between the level and slope factor loadings, η_{i1}^c and η_{i2}^c , is positive and around .5, and its interpretation is given further below.

Interestingly this pattern of correlations of centered factor loadings is very close to the correlation pattern of the average aggregate estimates across education groups reviewed in the previous subsection. The coefficient of correlation between $\bar{\eta}_{g1}$ and $\bar{\eta}_{g2}$ (i.e. obtained by varying g and weighting by group size) is equal to 0.64 and close to the random effect estimate of the correlation between η_{i1}^c and η_{i2}^c , which is equal to 0.5. The estimated correlation between $\bar{\eta}_{g1}$ and $\bar{\eta}_{g3}$ is negative, -0.66 and very close to the random effect estimate, -0.636 . The estimated correlation between $\bar{\eta}_{g2}$ and $\bar{\eta}_{g3}$ of -0.96 , is also very close to the random effect estimate, -0.95 . This result evinces that human capital investment patterns between and within groups are similar in France in contrast with what was found in the US (Heckman et al., 1998).

Appendix C completes this empirical analysis by describing the correlation patterns with

decreases the AIC criterion very moderately and some of the parameters are very imprecisely estimated (see Table S.iv).

initial conditions, which is an important aspect of the random effect estimation method, as well as goodness-of-fit diagnostics for variances and autocorrelations.

4.3 Fixed effect estimates

We now turn to the estimation results by FGLS of the three factor loadings for each individual wage profile. Technical details are completed in Supplementary Appendix S.IV in which bias-correction of variances and quantiles are presented.

Table 4 reports estimates of the bias-corrected covariance matrix of centered individual estimates and their random effect benchmark. Standard errors for any function of fixed effect estimates are computed using sampling variability to which is added parameter uncertainty due to the random effect pre-estimation of the covariance matrix of errors over time. We use Monte Carlo simulations to compute the latter by sampling 1,000 times in the asymptotic distribution of random effect estimates.

Our working sample to be used from now on, gathers individuals observed over more than 21 periods because small- T_i bias issues, even if corrected for first order bias seem to be lingering for observations $T_i \leq 20$.⁶ Even though random and fixed effects do not strictly refer to the same population because of this selection, discrepancies seem very moderate between random and fixed effect estimates – as Figure 3, reporting the profile of estimated variance of the interactive effects, confirms.

Table 4 also reports these estimates by subsamples indexed by a varying number of periods of observation from 21 to 28. It clearly appears that the longer the observation period, the less variable individual effects. It might be due to a subsisting small- T bias that we imperfectly control. It might also indicate that subpopulations, with exits or/and reentries, are more heterogeneous than the subpopulation of those who remain in the private sector all along. We left this point for further research since the literature is silent about higher order corrections. Random effect estimates would reflect the mixture of these subpopulations.

⁶Tables S.vii and S.viii in the Supplementary Appendix report raw and bias-corrected estimates in the full sample and provide the basis for such a selection.

Complementary results including robustness checks and goodness-of-fit diagnostics are presented in Appendix C. Further results on the consistency of the missing at random assumption is given in the Supplementary Appendix S.II.

4.4 Predictions

We now compare our results to those published in the literature. We can also assess whether predictions of human capital models conform with our estimates. Our three original points respectively develop the test and measure of the Mincer dip, the Rubinstein and Weiss (2006) prediction about the covariance between wage and wage growth and finally, the correlation between wage growth and initial wage level.

Literature estimates Table 5 completes Guvenen (2007, Table 4) with Guvenen’s own estimates, those of Huggett et al., (2011) and ours, on the estimation method, the AR(1) coefficient of transitory errors, and the covariance matrix of level and slope heterogeneity. Results in all these papers are qualitatively similar except, as expected, that the variance of the level heterogeneity is much larger when the residual process is not decomposed into a random walk and a transitory term (rows 1 and 4). To make our results comparable with others, we projected out the curvature profile onto a level and a trend and restricted our model to having heterogeneity in level and slope only. Our restricted results on the variance terms are smaller than results in Baker (1997), Haider (2001) or Huggett et al. (2011), and in particular, the variance of the slope is divided by two in the French data. The covariance of factor loadings of level and slope is negative except in the pionnering work of Lillard and Weiss (1979), and more surprisingly, in Huggett et al. (2011), a point to which we return below. Correlations, in absolute value, have the same order of magnitude although slightly larger in the US.

Mincer dip Our estimates generate a U-shaped profile of log-wage variances over the life cycle and thus, a neat Mincer dip at the beginning of the life cycle after 4 to 8 years (Mincer, 1974). This is shown by plotting, in Figure 3, the profile of predicted variances of (log) wages along the life cycle computed using random and, raw and bias-corrected, fixed effect estimates.

We display, in this graph, the life cycle profile of variances of the permanent effects (e.g. the human capital stocks), given by the combination of factors and factor loadings ($V(M(\beta)\eta_i^c)$) in which matrix $M(\beta)$ is composed of a constant, trend and geometric terms as in equation (6). Stochastic transitory earnings, v_i^c , and initial conditions are excluded from this graph because they obscure the comparison between different estimates of the permanent effects.

This graph shows that the Mincer dip is not dependent upon noisy estimates at the beginning of the life cycle and supports the structural interpretation given by Mincer. The most able workers have lower wages at the beginning of the life cycle because they are investing more in human capital and their profile of wages crosses from below the profile of those who are less able and invest less. This creates a trough in variances when potential experience is small.

Interestingly, note that the absence of a curvature parameter makes the variance profile, $V(\eta_{i1}^c + t\eta_{i2}^c)$, exhibit a Mincer dip (e.g. $t^{(Dip)} = -\frac{Cov(\eta_{i1}^c, \eta_{i2}^c)}{V(\eta_{i2}^c)} > 0$) if and only if $Cov(\eta_{i1}^c, \eta_{i2}^c) < 0$. It might explain why this estimated covariance is mostly negative in the literature as seen above in Table 5 with some exceptions. It is only when we allow for the horizon or curvature effect (i.e. our full results) that the covariance between the level and slope factor loadings, η_{i1}^c and η_{i2}^c , is positive (e.g. Table 3, fourth row). In consequence, our set-up accomodates the existence of a Mincer dip as well as the insight that the ability to earn (η_{i1}^c) is positively correlated with the ability to learn (η_{i2}^c), all other things being equal (i.e. η_{i3}^c) while a restriction to level and slope heterogeneity cannot reconcile these two insights.

Finally, wage profiles using raw and bias-corrected fixed effect estimates reproduce the random effect variance profile at a slightly higher level during the first years of working life in Figure 3. Discrepancies with random effect estimates seem however second order in particular at later periods in the life cycle and this also validates the use of the selected sample of males with more than 21 observations in the decomposition of inequalities and counterfactual variations in Section 5.

Correlation between wage and subsequent wage growth As developed in Rubinstein and Weiss (2006), a human capital model predicts that the covariance between wage levels and

subsequent wage growth should be negative when the person enters the labour market, and should turn positive after some time, in contrast to a job search model in which the correlation remains negative. This is what Table 6 neatly confirms. In the working sample, covariances are negative in years 1977 and 1982 and turn positive and significant from 1987 onwards, although the increase over time is non monotonic. This is slightly more pronounced for the high-skilled group. Rubinstein and Weiss (2006) use data on residual wages, smoothed over three years and show that this covariance increases over time although it remains negative (their Figures 8a-8e in Section 4.2). Our focus on the permanent effects that filter out transitory components allows to go a step further and produces a prediction in contradiction with a pure job-search rationale at least after ten years in the labour market.

To contrast those results with the ones in the literature, we projected out the curvature profile onto a level and a trend, as above, to restrict heterogeneity to level and slope only. In this case, the correlation between wage levels and wage growth is unambiguously negative (-0.349 s.e. 0.053). It is thus only when a curvature effect is present that the prediction of Rubinstein and Weiss holds.

Initial wage levels and returns to experience Initial skills and skills acquired during the education stage, are shown to be complementary (see Heckman, Humphries and Veramendi, 2017). It is interesting to measure dynamic complementarities over the life cycle by computing the correlation between the initial log wage and the returns at each point of the life cycle as in Gladden and Taber (2009). By filtering out the transitory components at the beginning of the life cycle, we are able to look more precisely at the complementarity between estimated initial wage levels – by equation (1), computed as the estimate of $\eta_{i1} + \eta_{i3}$ – and returns.

Returns to experience are computed as the marginal effects of potential experience, t , on the permanent component – i.e. $\eta_{i2} - \log(\beta)\beta^{-t}\eta_{i3}$. These returns are decreasing because $\log(\beta) < 0$ and estimated η_{i3} s are mostly negative. As shown in Table 7, while the correlation is significant and negative at the beginning of the life cycle (the first twenty years) as in Gladden and Taber (2009) and Sorensen and Vejlin (2014), it turns positive and significant afterwards. This result

of dynamic complementarity later in the life-cycle is true in the working sample and in various skill subsamples.

The negative correlation at the beginning of the life cycle is due to the negative correlation between the initial log wage, $\eta_{i1} + \eta_{i3}$, and the growth effect, η_{i2} . The positive correlation between the second curvature term – $-\log(\beta)\beta^{-t}\eta_{i3}$ – and the initial log wage kicks in after 20 years to reverse this foregoing negative correlation. Individuals with high initial log wages tend to have higher returns later in the life cycle. It again emphasizes the importance of including curvature as a result of horizon effects (Lillard and Reville, 1998) in the analysis of wage profiles.

5 Wage profiles and inequalities over the life cycle

We now exploit those estimates to study the impact of heterogeneity on life cycle wages and decompose wage inequalities at a point in time and over the life cycle. This evaluation starts from the log wage equation (6) that gives the following wage profile:

$$\log y_i = M(\beta)(\bar{\eta}_{g(i)} + \eta_i^e) + v_i.$$

It can be decomposed into (log) human capital stocks, $M(\beta)(\bar{\eta}_{g(i)} + \eta_i^e)$ and (log) prices, v_i , net of depreciation. As mentioned earlier, some elements in the individual profile of wages can be missing and can be imputed thanks to the assumption that the process of prices, v_i , conditional on human capital stock is known or estimated. In this sense, the estimated human capital stocks (in logs), $M(\beta)(\bar{\eta}_{g(i)} + \eta_i^e)$, are "coding in" available information about individual wage profiles that allows each individual wage profile to be recovered.

As a consequence, we can consider a longer-run measure of wages over the observed life cycle, an appropriately discounted aggregate of log wages – that happens to correspond to intertemporal utility over the observed profile of wages, normalized in such a way that it can be compared with the annual values (see Magnac et al., 2018):

$$\log y_i^{(LR)} = \frac{1 - \beta}{\beta(1 - \beta^T)} \sum_{t=1}^T \beta^t \log y_{it}. \quad (7)$$

This longer-run measure of individual welfare can then be compared with short-run ones in cross-section as in Bowlus and Robin (2012).

Yet, extrapolation to a longer time horizon (i.e after T) is not possible absent any assumption about the behaviour of prices out of the observation window. We can nonetheless still extrapolate the (log) human capital stocks, $M(\beta)(\bar{\eta}_{g(i)} + \eta_i^e)$ out-of-sample.

In this section, we present diverse counterfactual and decompositions of dispersion of log wages based on these arguments. When we use bias-corrected fixed effect estimates, we only use information on the most stable individuals ($T_i \geq 21$).

5.1 Counterfactual impacts on means

We assess the impact on average log wages of increasing each component of η_i by one tenth of its standard deviation. Those experiments can be easily expressed as a transformation of η_i into $\phi(\eta_i)$ and we have that:

$$\begin{aligned} \Delta \log y_i &= \log y_i(\phi(\eta_i)) - \log y_i, \\ &= M(\beta)(\phi(\eta_i) - \eta_i). \end{aligned}$$

In Table 8, we report mean impacts, every five years, starting in 1977 – their year of entry into the labor market – and finishing the last year of observation in 2007. Unsurprisingly because only levels of log wages are affected, the impact of increasing the level, η_{i1} , by one tenth of its standard deviation has a constant short-run and long-run impact around 0.03 – i.e. 3% on the wage level – except in the first year because initial conditions blur the impact. This is somewhat below the average slope of profiles (Table 2) and thus, somewhat below the effect of an additional year of experience at the beginning of the life cycle. In contrast, the impacts of increases of one tenth of a standard deviation in η_{i2} (returns) and η_{i3} (horizon) increase from 1977 to 2007 as they accumulate over the life cycle. At the end of the period of observation, they have sizeable magnitudes slightly above 0.15 on the log wage in the terminal year 2017. Effects on the long-run value are less sizeable and around .06 (returns) or .07 (horizon). Overall, these results mean that heterogeneities in level, slope and curvature that we consider are all economically significant.

We now address three questions related to the decomposition of the variance of log wages into its different components.

5.2 Short-run and long-run inequality decompositions

Table 9 provides decompositions of within-cohort inequalities and long-run inequality into its components: observed heterogeneity by skill groups ($\bar{\eta}_{g(i)}$), unobserved permanent heterogeneity (η_i^c) and transitory components. Because of initial conditions and their correlation with η_i in equation (6),⁷ we first need to orthogonalize permanent and transitory components. Initial conditions are first projected onto individual effects η_i^c and this projection is aggregated with the impact of individual effects η_i^c in the term, $D\eta_i^c$, and the residuals with the transitory ones in the term ξ_i so that we can write the log wage equation as:

$$\log y_i = M(\beta)\bar{\eta}_{g(i)} + D\eta_i^c + \xi_i,$$

in which the three components $\bar{\eta}_{g(i)}$, η_i^c and ξ_i are orthogonal to each other. Supplementary appendix S.IV gives more details.

We report results obtained by random effects in the first panel of Table 9 and by fixed effects in the second panel. Results obtained using random or fixed effect estimates are very close to each other and we will comment the latter only. The last column of these Tables reports the predicted variance of log-wages $V(\log y_{it})$ every five years from 1977 to 2007 as well as the predicted long-run value defined in equation (7). These results conform with the inverted U-shaped profile of variances as in Figure 2. The variance of long-run log wages is lower since transitory components over the life cycle are partly averaged out. The ratio of short-run inequalities to the long run ones is varying between 1.22 in 1982 and 1.82 in 2007, well in line with Bonke et al. (2014).

The first three columns respectively report the share of the variance due to observed heterogeneity, $V(M(\beta)\bar{\eta}_{g(i)})$, the share of the variance attributable to unobserved permanent heterogeneity, $V(D\eta_i^c)$, and the share of the variance generated by transitory components, $V(\xi_i)$.⁸

⁷Initial conditions are negatively correlated with η_{i1} and η_{i3} , and these effects play an important role but only during the very first years of the working life (Table S.iv of the Supplementary appendix).

⁸The decomposition in 1977 seems to be quite different from the one in other years because of the eventful process at the beginning of a working life. The weight in percentage terms of the transitory component is similar to other years but the variance of unobserved heterogeneity is almost absent. This is partly due to the orthogonalisation procedure that we have just discussed above, whereby negative correlations between initial conditions and fixed effects lower the contribution of permanent unobserved heterogeneity in the variance of log wages.

On average, 68% of the variance is due to the combination of the observed and unobserved heterogeneity factors in 1977 and 1982. This share displays an increase over the life cycle from 68% to 91% thirty years later. As a mirror image, the share of the variance explained by transitory components decreases sharply from 32% in 1982 to 8.5% in 2007 as well as the share due to the observed heterogeneity component, albeit at a lesser degree from 35% to 25%. This is the consequence of an increase in the importance of unobserved individual heterogeneity, which doubles its contribution to the variance of log wages from 35% to 67%. Wage profiles are thus stabilizing over the life-cycle and heterogeneity is getting less and less transitory.

As expected, these effects are weighted differently when the long-run value of log wages (i.e. equation (7)) is computed. The transitory effect is smaller (less than 6%) and observed and unobserved heterogeneity components have roughly equal contributions (45 and 49%).

These decompositions compare well with those of Guvenen (2009) given that, in that paper, observed heterogeneity is excluded from the decomposition. For the same reason, the sum of the observed and unobserved heterogeneity contributions is larger than the ones found by Huggett, Ventura and Yaron (2011), who find that the initial endowments related to human capital (initial human capital and learning ability as well as initial wealth) account for 60% of the variance of lifetime wage. Their framework however allows for less pervasive unobserved heterogeneity, than we do here, so that we capture more unobserved heterogeneity than they do. This is very much in line with the importance of unobserved heterogeneity in explaining the variance of wages (90%) found in Keane and Wolpin (1997, p508).

5.3 Interdecile ranges of log wages

We can also approach the dispersion of wages through other indicators. We here use the inter-decile range, that is the difference between the .9-quantile and the .1-quantile of the distribution of log wages that can be constructed using fixed effect estimates. The range cannot be decomposed as variances, and we thus report the inter-decile ranges of three variables: (1) observed heterogeneity, $M(\beta)\bar{\eta}_{g(i)}$, (2) the sum of observed and unobserved heterogeneity components, $M(\beta)\bar{\eta}_{g(i)} + D\eta_i^c$, (3) observed (log) wages (excluding missing observations). These statistics are

again reported every 5 years for these variables as well as the long run value, except for the "observed" long-run value since wages are not unobserved at all during three years.

As shown in column 3 of Table 10, the interdecile range of observed wages is large in the first period, decreases to 0.786 in 1982 and increases again to reach 1.04 so that the interdecile range of wages hovers around 100%. The range of the observed heterogeneity component is much lower (first column) and is varying between .46 and .57 if we exclude the large 1977 observation. Furthermore, the range of the sum of the observed and unobserved components is almost equal to the range of observed log wages and confirm the results of the variance decomposition above: the weight of unobserved heterogeneity increases over the life cycle and overcomes the weight of observed heterogeneity when potential experience is equal to 25 years; in long-run values however, the weight of observed heterogeneity remains predominant. In sum, those results confirm those we obtained for variances.

We now return to variance decompositions involving each individual specific parameter.

5.4 Components of human capital stocks

We can further decompose observed and unobserved heterogeneity into its constituent parts: level, η_{i1} , slope, η_{i2} , and curvature, η_{i3} . Denote the permanent heterogeneity variable, as defined in equation (6):

$$p_i = M(\beta)\eta_i, \tag{8}$$

which stands for the profile of log human capital stocks between 1977 and 2007 as per equation (1).

If observed heterogeneity only is present (i.e. $\eta_i = \bar{\eta}_{g(i)}$), the variance of p_i is the variance of an homogenous within-group Mincer equation. If there is as much heterogeneity as is estimated, the estimated $V(p_i)$ might be computed as previously in Table 9. Between these two benchmarks, we can compute hypothetical variances of the permanent effect by shutting down each or two of its components.

For instance, in case 2 below, we re-estimated individual-specific parameters in a model of log-wages whose individual specific parameters are restricted to one parameter only, namely a

standard level fixed effect. More comprehensively, we consider the following four estimations:

1. **Observable benchmark:** heterogeneity in level, slope and curvature restricted to observables – $\eta_{i1} = \bar{\eta}_{g(i)1}$, $\eta_{i2} = \bar{\eta}_{g(i)2}$ and $\eta_{i3} = \bar{\eta}_{g(i)3}$.
2. **Heterogeneity in levels only:** we reestimated for each individual a level fixed effect applying optimal minimum distance to the estimated $\hat{\eta}_i^c$ using its estimated variance. The restricted estimate is added to observable heterogeneity, $\bar{\eta}_{g(i)}$. In other words, this is what would be obtained in a model allowing for one fixed effect in levels only.
3. **Heterogeneity in slope only:** we reestimated for each individual a slope fixed effect applying minimum distance to the estimated $\hat{\eta}_i^c$ using its estimated variance. The restricted estimate is added to observable heterogeneity, $\bar{\eta}_{g(i)}$. In other words, this is what would be obtained in a model allowing for one fixed effect in slope only.
4. **Heterogeneity in level and slope:** we reestimated for each individual level and slope fixed effects, applying minimum distance to the estimated $\hat{\eta}_i^c$ using its estimated variance. These restricted estimates are added to observable heterogeneity, $\bar{\eta}_{g(i)}$. In other words, this is what would be obtained in a model allowing for two fixed effects in level and slope.

Table 11 reports variances of the predicted permanent profile of (log) wages, as defined in equation (8) and in each case 1 to 4. Rows report them every five years from 1977 to 2007. We also provide out-of-sample estimates for year 2013 as well as the out-of-sample long-run variance including years after the observation period between 2007 and 2013. The last column reports the estimated variance of p_i when we allow for three heterogeneity terms (e.g. describing level, slope and curvature). The first to the fourth columns report variances for every case 1 to 4 described above. We compare these predictions to the complete case using the three heterogeneity terms (last column).

First, as shown in the previous section, observables explain most of the permanent heterogeneity in 1977 while their explanatory power declines afterwards down to less than 30% in 2007 (column 1, Table 11). Second, if we restrict the heterogeneity to a level effect only (column 2),

we overestimate the variance of the permanent component in 1982 and 1987 and underestimate it after 1997. This is because this contribution remains almost constant over the years while the variance of the permanent component increases (column 5) so that we are approximating a U-shape profile (see Figure 3) by a horizontal straight line. This shows the limit of panel data analyses in which a single unobserved heterogeneity component in levels is considered. Third, if we restrict individual heterogeneity to slope only, we underpredict severely variances until 2002 and overpredict in 2007 (column 3). Considering two heterogeneity terms in level and slope combines the previous effects (column 4) and slightly underpredicts variances in intermediate years (1992 and 1997). The analogy is with an approximation of a U-shaped curve by a sloped straight line. It however does a good job at predicting long run values even if they are partly composed of out-of-sample elements. Nonetheless, all predictions restricting heterogeneity underpredicts the out-of-sample variances in 2013 although the large standard errors of these estimates make these differences insignificant.

We repeated this analysis by skill groups and those qualitative results are very similar across groups even if variance levels are different.

6 Conclusion

In this paper, we analyze wage profiles by using an observable microfounded factor model, based on a human capital investment model. Three factor loadings – level, slope and curvature – describe wage profiles. We propose an estimation method of the factor loadings based on a sequence of intertwined random and fixed effect methods. We show the importance of considering pervasive heterogeneity to model wage profiles and we revisit the questions about the Mincer dip, the correlation between wage level and wage growth or between wage growth and initial wages. These analyses clearly call for having three heterogeneity terms in our model.

We also present a set of decompositions of wage inequalities in terms of observed and unobserved heterogeneity as well as in terms of level, slope and curvature of the individual wage profiles. This allows to compute short-run and long-run inequalities and to compare them as

well as performing a set of decompositions. In those particular decompositions, heterogeneous curvature, or horizon effects are not as important as level and slope since curvature and slope are highly correlated.

Much remains of be done at the methodological level. The issue of missing data seems at the top of the agenda since the selection of balanced panels, or the missing at random assumption, might bias our view of inequalities as we select more stable subpopulations. Other ways of modelling transitory components might be in order and in particular, more attention might be paid to what can be anticipated by agents as in Cunha, Heckman and Navarro (2005). We leave these issues for further research.

REFERENCES

- Abowd, J. M., Kramarz, F., & Margolis, D. N.**, 1999, "High wage workers and high wage firms", *Econometrica*, 67(2), 251-333.
- Alvarez, J. and Arellano, M.**, 2004, "Robust likelihood estimation of dynamic panel data models", Working Paper 0421.
- Arellano, M., Blundell, R., & Bonhomme, S.**, 2017, "Earnings and consumption dynamics: a nonlinear panel data framework", *Econometrica*, 85(3), 693-734.
- Arellano M., and S., Bonhomme**, 2012, "Identifying Distributional Characteristics in Random Coefficient Panel Data Models", *The Review of Economic Studies*, 79(3): 987-1020.
- Atkinson, A.B. and S. Morelli**, 2014, "Chartbook of Economic Inequality: 25 Countries 1911-2010". *ECINEQ* WP 2014-324
- Autor, D. H., Katz, L. F., & Kearney, M. S.**, 2008, "Trends in US wage inequality: Revising the revisionists." *The Review of economics and statistics*, 90(2), 300-323.
- Bagger, J., Fontaine, F., Postel-Vinay, F., & Robin, J. M.**, 2014, "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics", *American Economic Review*, 104(6), 1551-96.
- 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.
- Belley, P.**, 2017, "Understanding Wage Growth: Estimating and Testing Learning-by-Doing", *Journal of Human Capital*, 11(2), 213-246.
- Blandin, A.**, 2018, "Learning by doing and Ben-Porath: Life-cycle predictions and policy implications", *Journal of Economic Dynamics and Control*, 90, 220-235.
- Blundell, R.**, 2014, "Income Dynamics and Life-cycle inequality: Mechanisms and Controversies", *The Economic Journal*, 124:289-318.
- Blundell, R., & Etheridge, B.**, 2010, "Consumption, income and earnings inequality in Britain", *Review of Economic Dynamics*, 13(1):76-102.
- Bonhomme, S., and J.M., Robin**, 2009, "Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990-2000", *Review of Economic Studies*, 76(1), 63-92.
- Bönke, T., Corneo, G., & Lüthen, H.**, 2014, "Lifetime earnings inequality in Germany", *Journal of Labor Economics*, 33(1), 171-208.
- Bowlus, A. J., & Liu, H.**, 2013, "The contributions of search and human capital to earnings growth over the life cycle", *European Economic Review*, 64, 305-331.
- Bowlus, A. J., & J.M., Robin**, 2012, "An international comparison of lifetime inequality: How continental Europe resembles North America", *Journal of the European Economic Association*, 10(6), 1236-1262.
- Bowlus, A. and C. Robinson**, 2012, "Human Capital Prices, Productivity, and Growth", *American Economic Review*, 102(7), 3483-3515.
- Bozio, A., Breda, T. and M. Guillot**, 2020, "The Contribution of Payroll Taxation to Wage Inequality in France", IZA DP 13317
- Browning, M., Ejrnæs, M. and Alvarez, J.**, 2012, "Modelling income processes with lots of heterogeneity", *Review of Economic Studies*, 77(4):1353-1381
- Burdett, K., Carrillo-Tudela, C., & Coles, M.**, 2016, "Wage inequality: A structural decomposition", *Review of Economic Dynamics*, 19, 20-37.
- Charnoz, P. Coudin, E. and Gaini, M.**, 2011, "Wage inequalities in France 1976-2004: a quantile regression analysis", WP DESE n°G2011/06.

- Charnoz, P. Coudin, E. and Gaini, M.**, 2014, "Forty Years of Decreasing Wage Inequality in France: The Roles of Supply and of Hidden Skill-Biased Technical Change", WP CREST n° 2014-20.
- Chiappori, P. A., Costa Dias, M., & Meghir, C.**, 2018, "The marriage market, labor supply, and education choice", *Journal of Political Economy*, 126(S1), S26-S72.
- Cho, J. S., Phillips, P. C., & Seo, J. W.**, 2019, "Parametric Inference on the Mean of Functional Data Applied to Lifetime Income Curves", working paper.
- Cunha, F., J.J., Heckman and S. Navarro**, 2005, "Separating uncertainty from heterogeneity in life cycle earnings", *Oxford Economic Papers*, 57(2):191-261.
- Daly, M., Hryshko, D., & Manovskii, I.**, 2016, "Reconciling estimates of income processes in growth rates and levels". Working paper, University of Pennsylvania.
- Ejrnæs, M., and M., Browning**, 2014, "The persistent–transitory representation for earnings processes", *Quantitative Economics*, 5(3), 555-581.
- Engbom, N.**, 2017, "Worker Flows and Wage Growth over the Life-Cycle: A Cross-Country Analysis", Available at SSRN 2976359.
- Fernández-Val, I., Peracchi, F., van Vuuren, A., & Vella, F.**, 2018, "Decomposing changes in the distribution of real hourly wages in the US", arXiv preprint arXiv:1901.00419.
- Fernández-Val, I., & Weidner, M.**, 2018, "Fixed effects estimation of large-t panel data models", *Annual Review of Economics*, 10, 109-138.
- Gladden, T., & Taber, C.**, 2009, "The relationship between wage growth and wage levels", *Journal of Applied Econometrics*, 24(6), 914-932.
- Gottschalk, P., & Moffitt, R.**, 2009, "The rising instability of US earnings", *Journal of Economic Perspectives*, 23(4), 3-24."
- 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.
- Haider, S. J.**, 2001, "Earnings instability and earnings inequality of males in the United States: 1967–1991", *Journal of Labor Economics*, 19(4), 799-836.
- Heckman, J. J., Humphries, J. E., & Veramendi, G.**, 2018, "Returns to education: The causal effects of education on earnings, health, and smoking", *Journal of Political Economy*, 126(S1), S197-S246.
- Heckman, J., Lochner, L., & Cossa, R.**, 2002, "Learning-by-doing vs. on-the-job training: Using variation induced by the EITC to distinguish between models of skill formation", NBER 9083, National Bureau of Economic Research.
- Heckman, J.J., L. Lochner and C. Taber**, 1998, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics*, 1, 1 - 58.
- Hoffmann, F.**, 2019, "HIP, RIP and the robustness of empirical earnings processes", *Quantitative Economics*, 10, 1279–1315.
- Hospido, L.**, 2012, "Modelling heterogeneity and dynamics in the volatility of individual wages", *Journal of Applied Econometrics*, 27(3), 386-414
- Hryshko, D.**, 2012, "Labor income profiles are not heterogeneous: Evidence from income growth rates", *Quantitative Economics*, 3(2)177-210.
- Huggett, M., G., Ventura and A., Yaron**, 2011, "Sources of Lifetime Inequality", *American Economic Review*, 101:2923-2954.
- Jeong, H., Kim, Y., & Manovskii, I.**, 2015, "The price of experience", *American Economic Review*, 105(2), 784-815

- Jochmans, K., & Weidner, M.**, 2018, "Inference on a distribution from noisy draws", arXiv preprint arXiv:1803.04991.
- Keane, M. P., & Wolpin, K. I.**, 1997, "The career decisions of young men", *Journal of Political Economy*, 105(3), 473-522.
- Lagakos, D., Moll, B., Porzio, T., Qian, N., & Schoellman, T.**, 2018, "Life cycle wage growth across countries", *Journal of Political Economy*, 126(2), 797-849.
- 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., & Weiss, Y.**, 1979, "Components of variation in panel earnings data: American scientists 1960-70", *Econometrica*, 47: 437-454.
- Lise, J., and F. Postel-Vinay**, 2015, "Multidimensional skills, sorting, and human capital accumulation", *American Economic Review*, 110(8), 2328–2376
- Lochner, L., and Y. Shin**, 2014, "Understanding earnings dynamics: Identifying and estimating the changing roles of unobserved ability, permanent and transitory shocks", National Bureau of Economic Research, WP-w20068.
- Magnac, T., N., Pistoiesi and S., Roux**, 2018, "Human Capital Investments and the Life Cycle of Earnings", *Journal of Political Economy*, 126(3), 1219-1249.
- Magnac, T., and S., Roux**, 2021, "Heterogeneity and Wage Inequalities over the Life-Cycle", TSE working paper, 1041.
- 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.
- Milanovic, B.**, 2016, *Global Inequality*, The Belknap Press of Harvard University Press, Cambridge.
- Mincer, J.**, 1974, *Schooling experience and earnings*, National Bureau of Economic Research, Inc.
- Moffitt, R. A., & Gottschalk, P.**, 2011, "Trends in the covariance structure of earnings in the US: 1969–1987", *The Journal of Economic Inequality*, 9(3), 439-459.
- Okui, R.**, 2009, "The Optimal Choice of Moments in Dynamic Panel Data Models", *Journal of Econometrics*, 151:1-16.
- Piketty, T.**, 2013, *Le capital au XXIe siècle*, Le Seuil: Paris.
- Polachek, S. T. Das and R. Thamma-Apiroam**, 2015, "Heterogeneity in the Production of Human Capital", *Journal of Political Economy*, 123(6), 1410-1455.
- Pora, P., & Wilner, L.** (2017). Heterogeneous Exposure to Labor Earnings Risk.
- 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.
- Sanders, C., & Taber, C.**, 2012, "Life-cycle wage growth and heterogeneous human capital", *Annual Review of Economics*, 4(1), 399-425.
- Song, J., Price, D. J., Guvenen, F., Bloom, N., & Von Wachter, T.**, 2018, "Firming up inequality", *The Quarterly Journal of Economics*, 134(1), 1-50.
- Sørensen, K. L., & Vejlín, R.**, 2014, "Return to experience and initial wage level: Do low wage workers catch up?", *Journal of Applied Econometrics*, 29(6), 984-1006.
- Taber, C., and R., Vejlín**, 2020, "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market", *Econometrica*, 88(3), 1031-1069.

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

APPENDICES

A Data Appendix

As in Le Minez and Roux (2002), we consider individuals right from their entry into the labor market and onwards. Labor market entry is defined as being employed for more than 6 months and being still employed the following year, possibly in different firms. For the entry cohort of interest which starts in 1977, this leads us to select from the administrative data 36,883 individuals who were employed more than 6 months in 1977 and at least one day in 1978. Among them, 53% have worked but not permanently before. Conversely, individuals who have worked in 1977 are not considered as entrants if their jobs are not permanent enough. They may however enter with a subsequent cohort.

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

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

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

as missing the first and last percentiles of the earnings distribution in every period. Inflation, as measured by consumer prices, leads to subtracting a profile of inflation factors to current log-wages (e.g equal to 1.17 at the end of the period of observation). The whole period can be roughly subdivided into two sub-periods between 1977 and 1986 in which the inflation factor reaches .77 in 1986 and between 1986 and 2007 during which inflation levelled off and cumulative inflation reaches .40 only.

Age at labor market entry (in 1977) can only take odd integer values from 17 to 29, i.e seven different values because of the specific sampling of the dataset. As groups formed by age at entry and skills are defined according to characteristics recorded at the entry on the labor market, individuals are attached to the same group during their whole working life.

Estimation of human capital prices by a flat spot condition and robustness checks

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

We then subtracted these prices from all log wages that lead to the estimates of Table 2 and further Tables.

B Random effect specification

B.1 An overview

Equation (6) writes

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

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

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

is a $[T, 3]$ matrix.

In the absence of correlation of the model with initial conditions (that are presented next Section), mean independence restrictions (5) lead to:

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

and:

$$V(u_i) = V(E(u_i | \eta_i^c)) + E(V(u_i | \eta_i^c)) = M(\beta)V(\eta_i^c)M(\beta)' + E(\Omega(\eta_i^c)). \quad (\text{B.1})$$

Our parameter of interest in this equation is the covariance matrix of the individual effects, $V(\eta_i^c)$ standing for the covariances between level, slope and curvature parameters. Identifying this covariance matrix requires restrictions on the average variance of the idiosyncratic errors, $E(\Omega(\eta_i^c))$. An ARMA specification is common in the earnings dynamic literature and generally low orders are used (see Guvenen, 2009, or Hryshko, 2012) whereas an alternative is a composition of permanent and transitory shocks with specific structures (Blundell, 2014, Bonhomme and Robin, 2009, Lochner and Shin, 2014). We refrain from decomposing v_i^c into its persistent and transitory components since those are not identified absent additional restrictions and consumption data (Ejrnæs and Browning, 2014).

Arellano and Bonhomme (2012) show that a finite lag ARMA specification is sufficient to identify $V(\eta_i^c)$. We use this result and proceed by specifying that the processes v_{it}^c belong in the family of time-heteroskedastic ARMA processes although we limit the orders of the AR and MA to vary between 1 and 3. This allows the robustness of the estimated covariance of individual effects, $V(\eta_i^c)$, to the orders of the ARMA process to be assessed. Moreover, we allow for time heteroskedasticity of the innovations whose importance is argued by Alvarez and Arellano (2004). What the decomposition (B.1) shows in addition, is that a restricted form of individual heterogeneity, possibly dependent on parameters, η_i , could be allowed in the ARMA process, provided that the expected value, $E(\Omega(\eta_i^c))$, remains in the ARMA family that we consider.

The most commonly used minimum distance method for estimating equation (B.1) is severely small-sample biased since the range of moments involved when the time dimension becomes large makes first order asymptotics a poor guide in empirical research. Okui (2009) derives the small sample biases not only in the mean but also in the variance of GMM estimates due to the presence of too many moments and he suggests some moment selection mechanism. This is why some researchers proposed to return to an OLS set up adding a bias correction step (Fernandez-Val and Weidner, 2018) or to maximum or pseudo-maximum likelihood methods that reduce the number of moments available (Alvarez and Arellano, 2004).

Specifically, the estimation method proposed by Alvarez and Arellano (2004) seems to dominate in Monte Carlo experiments other fixed T consistent estimators such as the maximum likelihood estimator using differenced data, and the corrected within group estimator. This method is particularly well adapted to the case in which there are missing data in wage dynamics. For any missing data configuration, it consists in deleting the rows and columns of the

covariance matrix corresponding to missing data and write the likelihood function accordingly. Under a normality assumption, the implicit moment selection underlying this estimation method is optimal, and though the method loses optimality in the non-normal cases, it is still useful for moment selection and for small-sample bias reduction (Okui, 2009).

B.2 Heteroskedastic ARMA Specification & Initial conditions

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

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

We define v_{it} as

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

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

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

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

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

$$Vu_i = \Omega_u.$$

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

The construction of the likelihood function is explained in Supplementary Appendix S.III.

C Further empirical results

Initial conditions Irrespective of the order of the ARMA process, the initial conditions are negatively correlated with the level factor loadings, η_{i1}^c , positively with the slope effect, η_{i2}^c , and negatively with the curvature one, η_{i3}^c .⁹ These initial conditions account for the strong transitory conditions that seem to affect the wage process at the beginning of the working life (as well as the impact of our data selection process).¹⁰ Even if the (log) wage process is asymptotically stationary, initial conditions are definitely not set on the stationary path that corresponds to this process. Moreover, in all ARMA specifications, the standard deviation of individual and time specific transitory shocks is decreasing over time. Individual specific frictions decrease over time and this result is found across different countries (e.g. Bagger et al., 2014, Bowlus et Liu, 2012).

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

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 the variance of wages over the life cycle using fixed effect estimates (see below), it clearly appears that the level of these profiles negatively depends on the number of periods in which we observe each individual. Variances are larger for individuals who have shorter spells in the panel. Nevertheless, correcting for non random attrition seems out of the scope of this paper and we leave it for further research. Furthermore, the conditions for consistency of fixed effect estimates described below are less stringent since the missing at random assumption can be weakened and taken as conditional on individual effects.

Robustness and other diagnostics We tested various departures from our baseline estimates to check that our results are robust. We also comment on additional goodness-of-fit diagnostics.

⁹Estimates of the covariances between the factor loadings and the initial conditions are reported in Table S.iv in the Supplementary Appendix.

¹⁰The strong decrease of the variance observed during the first years might partly be due to the very stringent selection made in the 1977 entry cohort. The very flexible initial conditions, as they are accounted for in the random effect estimation, also control for this selection.

The first issue is the estimation of the series by a flat spot approach that underpins the identifying restriction (3). The dynamics of human capital accumulation depends on whether the average wage or productivity profile is attributed to human capital only or to other factors (physical capital for instance). To control for this issue, we also repeated our procedures by deflating real wage by a series of average labor productivity. Results change only marginally with respect to the results that were presented above.

As serial correlation affects inference, we also vary the number of lags in the Newey West procedures without much impact overall.

Bias correction Another issue is related to the quality of the correction of the bias in the fixed effect estimates that we reported. In Tables S.vii and S.viii in the Supplementary Appendix which report raw and bias-corrected estimates, the magnitude of all covariance estimates decreases with the number of period of observations, as expected by the computations of Section S.IV. Bias-correction flattens these estimates by factors of 2 to 3 when the number of periods is small but it decreases to 10-20% when the number of periods is 20. There is a clear break in these Tables between the estimates below and above 20 and this is why we chose to work with the 21+ sample. As expected the bias correction becomes negligible when $T_i = 28$. The counterfactual analyses that we present below are robust to a change in this threshold.

The number of factors The comparison between random effect and fixed effect estimates implicitly relies on an homogeneity assumption of the residuals, \hat{v}_i^c as a function of η_i . When plotting the variance profiles of these residuals in groups defined either by skills or by age of entry, we find very little differences between those groups (see Figure S.v in the Supplementary appendix). The three-factor structure seems to be sufficient to describe the individual permanent heterogeneity in our data and this partly justifies ex-post the homogeneity assumption of the covariance matrix of transitory terms in the random effect specification as well as the homogeneity assumption of the discount factor which is used to measure the curvature or horizon effect.

TABLES AND FIGURES

Table 1: Sample size

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

Table 2: Group averages of individual factor loadings η_g

Skill group	Age group	Nb Obs	$\bar{\eta}_{g1}$	$\bar{\eta}_{g2}$	$\bar{\eta}_{g3}$
2	17	1268	2.5 (0.039)	0.024 (0.0073)	-0.049 (0.056)
3	17	1224	2.5 (0.04)	0.023 (0.0059)	-0.051 (0.044)
1	19	41	2.8 (0.04)	0.053 (0.0068)	-0.23 (0.051)
2	19	934	2.6 (0.035)	0.027 (0.0052)	-0.073 (0.039)
3	19	994	2.6 (0.044)	0.025 (0.0074)	-0.073 (0.054)
1	21	117	2.9 (0.085)	0.035 (0.0088)	-0.12 (0.07)
2	21	710	2.8 (0.022)	0.031 (0.0039)	-0.1 (0.03)
3	21	512	2.7 (0.019)	0.025 (0.0035)	-0.09 (0.025)
1	23	171	3.1 (0.017)	0.038 (0.0032)	-0.14 (0.023)
2	23	348	2.8 (0.03)	0.033 (0.0051)	-0.11 (0.038)
3	23	254	2.8 (0.049)	0.035 (0.0063)	-0.15 (0.048)
1	25	191	3.4 (0.053)	0.045 (0.0058)	-0.19 (0.044)
2	25	146	2.9 (0.061)	0.021 (0.0065)	-0.044 (0.046)
3	25	93	2.7 (0.025)	0.016 (0.004)	0.0083 (0.033)
1	27	114	3.5 (0.019)	0.03 (0.0042)	-0.12 (0.031)
2	27	87	3.1 (0.02)	0.044 (0.0033)	-0.22 (0.024)
3	27	63	2.7 (0.034)	0.013 (0.0048)	0.02 (0.036)
1	29	58	3.3 (0.054)	0.024 (0.0067)	-0.042 (0.049)
2	29	67	2.9 (0.089)	0.021 (0.014)	-0.097 (0.11)
3	29	55	2.7 (0.055)	-6.2e-05 (0.0079)	0.1 (0.062)

Note: Estimation of equation (2). A flat spot deflator is used. Newey West standard errors in parentheses (5 lags).

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

	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
σ_{η_1}	.302 (.001)	.302 (.003)	.301 (.003)	.310 (.003)	.306 (.003)	.304 (.003)	.306 (.003)	.300 (.003)	.298 (.004)
σ_{η_2}	.038 (.005)	.039 (.001)	.039 (.001)	.038 (.001)	.039 (.001)	.036 (.001)	.038 (.001)	.037 (.001)	.037 (.001)
σ_{η_3}	.255 (.005)	.259 (.006)	.256 (.006)	.263 (.004)	.260 (.005)	.248 (.005)	.258 (.005)	.247 (.006)	.242 (.007)
ρ_{η_1, η_2}	.473 (.016)	.413 (.021)	.454 (.021)	.571 (.013)	.486 (.017)	.610 (.013)	.505 (.017)	.485 (.020)	.365 (.030)
ρ_{η_1, η_3}	-.604 (.003)	-.548 (.020)	-.586 (.019)	-.694 (.011)	-.618 (.015)	-.729 (.012)	-.636 (.016)	-.620 (.019)	-.509 (.029)
ρ_{η_2, η_3}	-.946 (.023)	-.948 (.003)	-.947 (.003)	-.945 (.002)	-.946 (.002)	-.941 (.003)	-.946 (.002)	-.943 (.003)	-.944 (.004)

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

Table 4: Bias corrected covariance matrix of individual effects: fixed and random effect estimation

Sample periods	$Var(\eta_1)$	$Cov(\eta_1, \eta_2)$	$Cov(\eta_1, \eta_3)$	$Var(\eta_2)$	$Cov(\eta_2, \eta_3)$	$Var(\eta_3)$
21	0.18 (0.029)	0.012 (0.0032)	-0.13 (0.03)	0.0034 (0.00061)	-0.026 (0.0049)	0.22 (0.043)
22	0.15 (0.019)	0.015 (0.0031)	-0.13 (0.026)	0.0035 (0.00067)	-0.027 (0.0053)	0.22 (0.042)
23	0.16 (0.017)	0.012 (0.0024)	-0.11 (0.02)	0.0035 (5e-04)	-0.024 (0.0038)	0.19 (0.03)
24	0.14 (0.017)	0.014 (0.0027)	-0.12 (0.022)	0.0041 (0.00061)	-0.03 (0.0047)	0.23 (0.037)
25	0.13 (0.014)	0.01 (0.0023)	-0.089 (0.019)	0.0028 (0.00041)	-0.02 (0.0033)	0.16 (0.027)
26	0.097 (0.0072)	0.0066 (0.00093)	-0.059 (0.0072)	0.0025 (0.00023)	-0.017 (0.0016)	0.12 (0.012)
27	0.077 (0.0046)	0.0046 (0.00064)	-0.04 (0.0047)	0.0017 (0.00015)	-0.011 (0.001)	0.079 (0.0074)
28	0.067 (0.0049)	0.0036 (0.00067)	-0.031 (0.0047)	0.0015 (0.00016)	-0.0097 (0.0011)	0.067 (0.0074)
21+ sample	0.11 (0.0073)	0.0078 (0.00095)	-0.07 (0.0092)	0.0025 (0.00018)	-0.017 (0.0015)	0.13 (0.013)
Random effects	0.093 (0.0036)	0.0059 (0.00051)	-0.05 (0.004)	0.0015 (0.00011)	-0.0093 (0.00079)	0.066 (0.0059)

Notes: The first lines are obtained using fixed effect estimates. Sample periods = number of observed periods. Standard errors (heteroskedastic-consistent sampling and parameter uncertainty, 1000 MC simulations) between brackets. The working sample (21+) has 4873 observations

Table 5: Covariance matrix of individual effects, literature estimates

Source	ρ	$\sigma_{\eta_1}^2$	$\sigma_{\eta_2}^2$	σ_{η_1, η_2}	Stochastic process	Time varying variances
Lillard and Weiss (1979)	0.707 (0.073)	0.0305 (0.0015)	0.00018 (0.00004)	0.00076 (0.0001)	AR(1)	No
Baker (1997)	0.674 (0.050)	0.139 (0.069)	0.00039 (0.00013)	-0.004 (0.003)	ARMA(1,2)	Yes
Haider (2001)	0.639 (0.077)	0.295 (0.137)	0.00041 (0.00012)	-0.0083 (0.0036)	ARMA(1,1)	Yes
Guvenen (2007)	0.821 (0.020)	0.022 (0.074)	0.00038 (0.00008)	-0.0020 (0.0032)	AR(1)	Yes
Huggett, Ventura and Yaron (2011)	0.860 (0.010)	0.264 (0.024)	0.00006 (0.00006)	0.003 (0.001)	AR(1)	No
This paper: FE (restricted)	0.702 (0.005)	0.12 (0.017)	0.00021 ($8e - 6$)	-0.0017 (0.0002)	ARMA(3,1)	Yes
This paper: FE (complete)	<i>Idem</i>	0.11 (0.0073)	0.0025 (0.00018)	0.0078 (0.0009)	<i>Idem</i>	Yes

Note: ρ is the AR(1) coefficient and the second to fourth column report the variances of η_1 , η_2 and their covariance. Results in the first three rows are taken from Table 4, in Guvenen (2007). The fourth row is from Table 1 in Guvenen (2007) and the fifth row from Huggett, Ventura and Yaron (2011). The seventh row reports bias corrected fixed effect results taken from Table 4 in this paper. The sixth row projects out the last factor (β^{-t}) onto an intercept and a time trend and estimate a two-factor model whose results are comparable to rows (1)-(5). The value of ρ in row 6 is taken from the comparable ARMA(1,1) estimate in Table S.iv in the Supplementary Appendix.

Table 6: Covariances of wage level and subsequent growth

Years	All	Low skilled	Medium skilled	High skilled
1977	-0.00594 (0.000639)	-0.00573 (0.000564)	-0.00665 (0.000885)	-0.00688 (0.00214)
1982	-0.000972 (0.000388)	-0.00137 (0.000303)	-0.00102 (0.00052)	1.04e-05 (0.00149)
1987	0.00178 (0.000271)	0.00113 (0.000198)	0.00219 (0.000402)	0.00336 (0.00119)
1992	0.00233 (0.000207)	0.0018 (0.000162)	0.00299 (0.000357)	0.00334 (0.00089)
1997	0.00135 (0.000174)	0.00115 (0.000139)	0.00204 (0.000324)	0.00133 (0.000637)
2002	0.000812 (0.000269)	0.000838 (0.000229)	0.00141 (0.00042)	0.00108 (0.00108)
2007	0.00524 (0.000666)	0.00459 (0.000593)	0.00583 (0.00114)	0.0109 (0.00304)
Observations	4873	2942	1433	498

Notes: The coefficient is the covariance between $\eta_{i1} + \eta_{i2}t + \eta_{i3}\beta^{-t}$ and $\eta_{i2} + \beta^{-t}(1/\beta - 1)\eta_{i3}$. Standard errors (heteroskedastic-consistent sampling and parameter uncertainty, 1000 MC simulations) between brackets. The working sample (21+) has 4873 observations

Table 7: Time varying correlation of initial levels and returns

Year	All	Low	Med	High
1977	-0.498 (0.0521)	-0.674 (0.0425)	-0.561 (0.0682)	-0.379 (0.13)
1982	-0.505 (0.0529)	-0.683 (0.043)	-0.564 (0.0705)	-0.38 (0.134)
1987	-0.509 (0.0539)	-0.687 (0.0436)	-0.561 (0.0742)	-0.375 (0.139)
1992	-0.47 (0.0527)	-0.635 (0.0436)	-0.507 (0.0776)	-0.31 (0.136)
1997	-0.186 (0.0385)	-0.265 (0.042)	-0.201 (0.0664)	0.000268 (0.0796)
2002	0.156 (0.0346)	0.207 (0.0392)	0.186 (0.0458)	0.212 (0.0765)
2007	0.286 (0.0387)	0.389 (0.0387)	0.338 (0.0444)	0.276 (0.0888)

Note: The correlation is $\rho = Corr(\eta_{i1} + \eta_{i3}, \eta_{i2} - \log(\beta)\beta^{-t}\eta_{i3})$ Only observations with more than 21 periods. 4873 observations. First column reports results for centered individual effects while the other columns include aggregate effects. The last three columns for low, medium and high skills.

Table 8: Impacts of unobserved heterogeneity on mean log wages

Impact of $\{\sigma_j\}_{j=1,..,3}$ on:	Level	Slope	Curvature
	$\eta_1 \rightarrow \eta_1 + \sigma_1$	$\eta_2 \rightarrow \eta_2 + \sigma_2$	$\eta_3 \rightarrow \eta_3 + \sigma_3$
Log-wage 1977	0.0112 (0.000203)	0.0116 (0.000336)	0.0199 (0.00062)
Log-wage 1982	0.0303 (0.000549)	0.0263 (0.000765)	0.0448 (0.0014)
Log-wage 1987	0.0325 (0.000587)	0.0499 (0.00145)	0.0601 (0.00187)
Log-wage 1992	0.0326 (0.00059)	0.0747 (0.00217)	0.0778 (0.00242)
Log-wage 1997	0.0326 (0.00059)	0.0996 (0.0029)	0.101 (0.00313)
Log-wage 2002	0.0326 (0.00059)	0.124 (0.00362)	0.13 (0.00405)
Log-wage 2007	0.0326 (0.00059)	0.149 (0.00435)	0.168 (0.00523)
Long-run value	0.0287 (0.00052)	0.0573 (0.00167)	0.0672 (0.00209)

Note: Average impact on log wages of an increase of a tenth of the standard deviation of: 2nd column, unobserved heterogeneity in the initial human capital; third column, wage growth or returns; fourth column, curvature or horizon.

Table 9: Variance decomposition: random and fixed effects

Random Effects	Obs. het. %	Unobs. het. %	Transitory %	Total var.
Log-wage 1977	65.2	1.18	33.6	0.481
Log-wage 1982	34.8	33.1	32.1	0.133
Log-wage 1987	32.8	42	25.3	0.153
Log-wage 1992	29.7	49.9	20.4	0.179
Log-wage 1997	27.2	56	16.8	0.201
Log-wage 2002	26.2	60.8	13	0.206
Log-wage 2007	25.1	66.4	8.47	0.202
Long-run value (1977-2007)	50.2	44.5	5.31	0.118

Note: The variance (5th column) of each variable (1st column) is decomposed into its component shares which are reported in percentages in column 2 (observed heterogeneity), column 3 (unobserved heterogeneity) and column 4 (transitory component). The share of variance of log 1982 wage (0.133) explained by observed heterogeneity is 34.8

Fixed Effects	Obs. het. %	Unobs. het. %	Transitory %	Total var.
Log-wage 1977	64.4	1.48	34.1	0.475
Log-wage 1982	32.9	36	31.2	0.137
Log-wage 1987	33.2	40.7	26.1	0.148
Log-wage 1992	30.9	47.5	21.6	0.17
Log-wage 1997	28.7	53.3	18	0.188
Log-wage 2002	27.7	58.2	14	0.192
Log-wage 2007	24.2	67.5	8.28	0.206
Long-run value	51.6	42.8	5.58	0.112

Note: See Table above. Variances are biased corrected. Sample: 4873 individuals with more than 21 periods. Remark: 1977 is different from other years partly because of the orthogonalization procedure discussed in the text (footnote 6).

Table 10: Interdecile ranges

FE	Obs. het.	Obs. and Unobs. Heterogeneity	Observed
Log-wage 1977	1.57	1.6	1.89
Log-wage 1982	0.463	0.742	0.786
Log-wage 1987	0.478	0.798	0.883
Log-wage 1992	0.498	0.884	0.987
Log-wage 1997	0.507	0.962	1.02
Log-wage 2002	0.502	1	1.05
Log-wage 2007	0.456	1.07	1.04
Long-run value	0.567	0.801	-

Note: Quantiles at levels .9 and .1 are bias-corrected and their difference is taken to compute the inter-decile range. "Observed" uses non-missing observations only and is subject to selection issues. Sample: 4873 individuals with more than 21 periods. Fixed effect estimates.

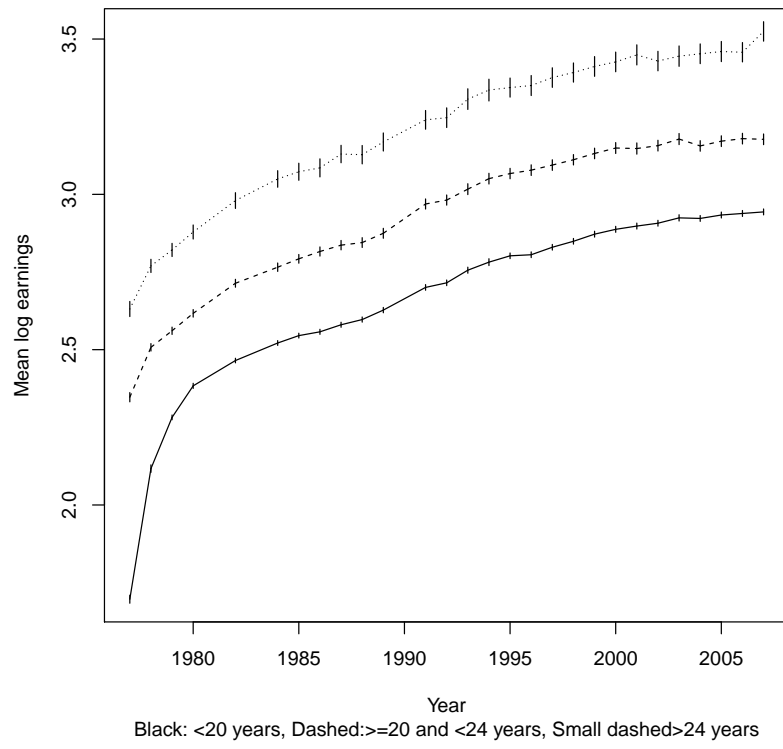
Table 11: Variance decompositions of log-wage permanent component

Years	$\eta_1 = \bar{\eta}_{1g}$ $\eta_2 = \bar{\eta}_{2g}$ $\eta_3 = \bar{\eta}_{3g}$	Heterogeneity in: Levels	Growth	Level and Growth	All
Log-wage 1977	0.0404 (0.0169)	0.05 (0.0171)	0.0414 (0.0169)	0.0479 (0.0172)	0.0474 (0.017)
Log-wage 1982	0.0449 (0.0181)	0.116 (0.0182)	0.0499 (0.0181)	0.101 (0.0183)	0.094 (0.0207)
Log-wage 1987	0.0491 (0.0197)	0.13 (0.0198)	0.0672 (0.0197)	0.111 (0.0199)	0.109 (0.0285)
Log-wage 1992	0.0523 (0.0221)	0.134 (0.0222)	0.0929 (0.0221)	0.121 (0.0222)	0.133 (0.0438)
Log-wage 1997	0.0539 (0.0257)	0.136 (0.0258)	0.126 (0.0258)	0.137 (0.0259)	0.154 (0.069)
Log-wage 2002	0.0533 (0.0313)	0.135 (0.0314)	0.166 (0.0314)	0.159 (0.0315)	0.165 (0.107)
Log-wage 2007	0.05 (0.0402)	0.132 (0.0403)	0.213 (0.0405)	0.188 (0.0405)	0.189 (0.165)
Log-wage 2013*	0.0429 (0.0591)	0.125 (0.0591)	0.277 (0.0594)	0.23 (0.0595)	0.365 (0.28)
Long-run value*	0.0474 (0.0206)	0.112 (0.0207)	0.0775 (0.0206)	0.101 (0.0207)	0.0969 (0.038)

Note: * = Out-of-sample predictions. Sample: 4873 observations with more than 21 periods. The columns are described in the text and measure the influence of each component of heterogeneity, in levels or/and growth.

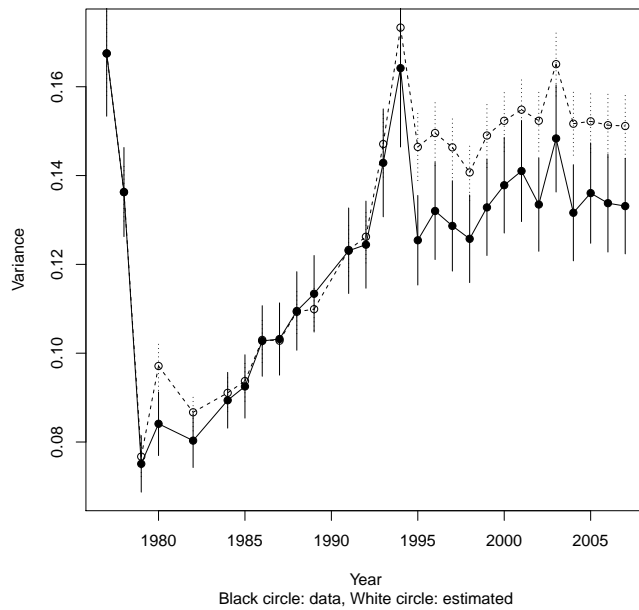
Reading: In 1977, the variance of the permanent component is 0.0474 (last column). Absent any unobserved heterogeneity, the variance is equal to 0.0404 (first column). The other columns report the predicted variance in experiments 2 to 4 in the text

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

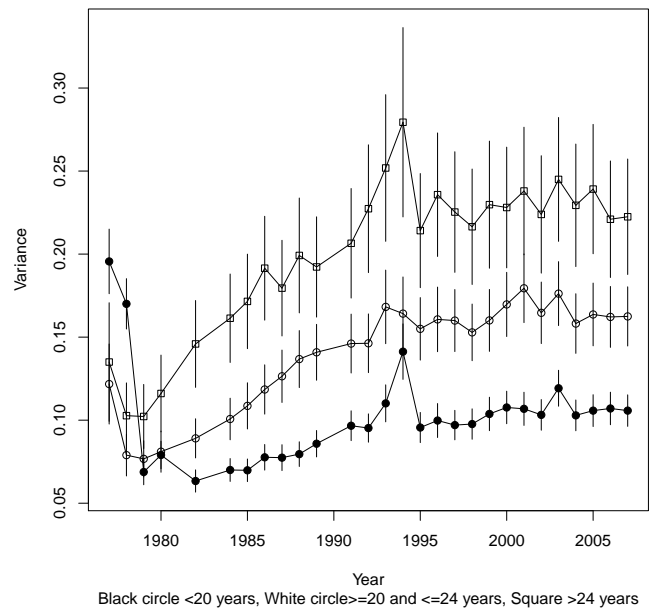


Note: The small vertical lines represent the 95% confidence intervals.

Figure 2: Cross-sectional variance of log wage residuals: 1977-2007



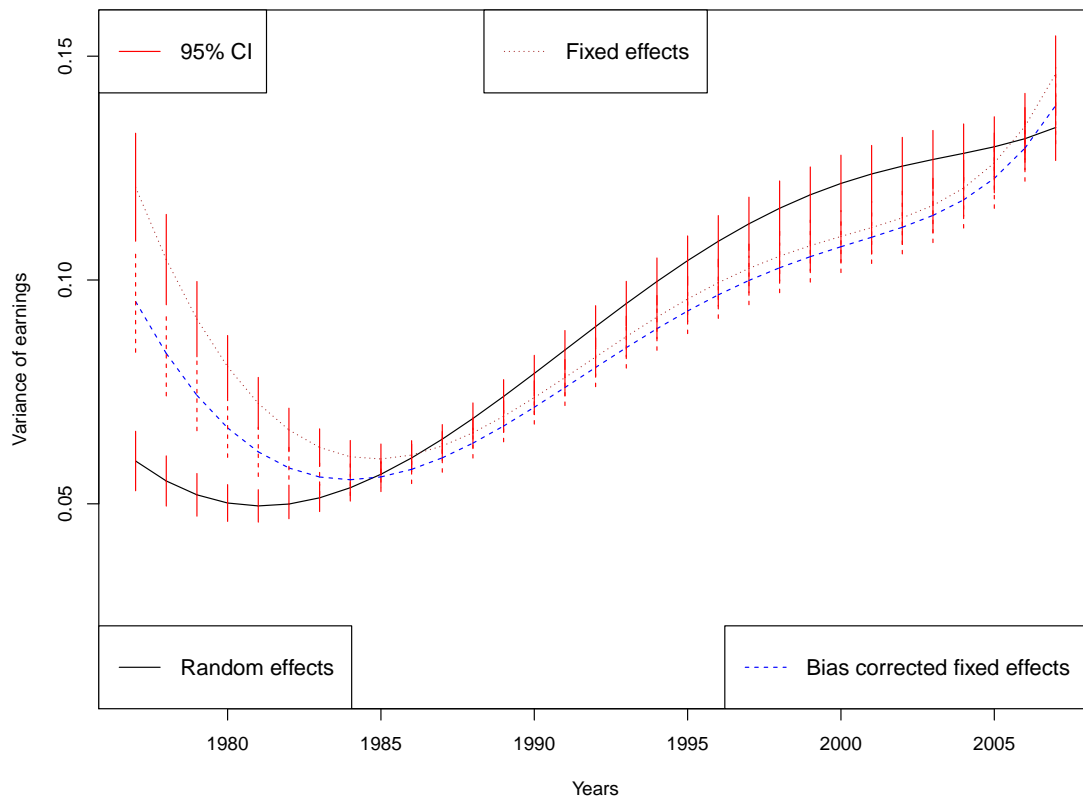
(A) full sample



(B) by age of entry

Note: The small vertical lines represent the 95% confidence intervals. Log wage residuals are obtained by regressing log wages on a saturated set of dummies for skill groups and years.

Figure 3: Variance of the permanent components



Note: The permanent component is $M(\beta)\eta_i^c$ defined in equation (6). The sample is restricted to long history profiles (more than 21 periods). "Random effects" are using estimates derived from random effect estimation. "Fixed effects" are using estimates derived from raw fixed effect estimation and "Bias corrected f.e." are the bias corrected version of them.

Supplementary Appendix to Magnac and Roux "Heterogeneity and Wage Inequalities over the Life Cycle" (for Web publication)

Thierry Magnac* and Sébastien Roux†

8th March 2021

S.I Notation

S.I.1 The model

- t : time elapsed since the entry in the labor market.
- i : index for individuals.
- β : homogenous discount rate
- T : Arbitrary date at which we examine whether individuals goes on investing in human capital, last date of observation in the empirical application.
- η_{i1} : individual-specific fixed level of log-wages.
- η_{i2} : individual-specific growth rate of log-wages.
- η_{i3} : individual-specific degree of curvature of log-wages.
- v_{it} : (log) price of human capital net of cumulative depreciation.
- g : group of workers, defined by their age at entry and their skills

*Toulouse School of Economics, Université de Toulouse Capitole, *corresponding author*, thierry.magnac@tse-fr.eu

†Insee-Crest-Ined, Paris, sebastien.roux@insee.fr

- $\overline{\ln y_{gt}}$: average of $\ln y_{it}$ over the group g
- η_{gk} : average of η_{ik} over g , for $k = 1, 2, 3$
- v_{gt} : average of v_{it} over g
- η_{ik}^c : centered individual effect of η_{ik} , for $k = 1, 2, 3$
- u_{it} : centered wages, with respect to group g
- v_{it}^c : individual-specific variations of human capital prices

S.I.2 Econometric Modeling

- $M(\beta)$: $T, 3$ matrix of factors.
- $\Omega(\eta_i^c)$: covariance matrix of centered individual fixed effects.
- $\hat{\eta}_i^c$: estimate of the centered individual fixed effect.
- B : matrix $3, T$ establishing the relationship between the centered individual fixed effects and the wages residuals.
- \hat{B} : estimate of B
- $\tilde{\eta}_i^c$: unfeasible estimator of η_i^c using B
- ξ_i : T -vector of residuals, orthogonal to η_i^c
- Ω_ξ : covariance matrix of ξ_i
- T_i : number of actual observations for the individual i .

S.II Supplementary Data Analysis

Auto-covariances The covariance matrix of log-wage residuals over time is reported Table S.ii although this is easier to use graphs to describe the main features of wage autocorrelations. Figure S.iii displays the autocorrelation of residuals of log wages in year t with residuals in an

early (resp. late) year, 1986 (resp. 2007). This Figure reveals an asymmetric pattern over time which is quite robust to the choice of these specific years (1986 and 2007). The correlation between wages in year t and in 1986 is swiftly increasing when t is before 1986 and this is also true for 2007 albeit at a lesser degree. In contrast, the correlation between wages in 1986 and in year t is only slowly decaying in t , if time t postdates 1986. Figure S.iv takes a different view that confirms the previous diagnostic by plotting the autocorrelations of order 1 and 6. Note that their shape are very similar and increase uniformly over time although at different levels. The closer to the end of the working life, the larger the autocorrelation coefficients are.

Moreover, in graph S.ii displays that this cohort has nothing specific when compared to younger cohorts entering later into the labor market.

Means and variances of individual-specific factor loadings We also estimated the correlation between mean estimates and the number of observed periods (reported in Table S.vi as well as the decreasing pattern of estimated variances with the observed number of periods (the raw version in Table S.vii and the bias-corrected one in Table S.viii).

We find that means of individual specific estimates are different when the number of missing periods vary but those differences are not significant in contrast with variances even though those are bias corrected.

S.III The random effect model : Model Specification and Likelihood function

We build up step by step the likelihood function by deriving the covariance matrix of $u_i^{[1-p,T]}$ as a function of the parameters of the different processes. We first describe the ARMA process and then include the individual heterogeneity factors.

S.III.1 Time heteroskedastic ARMA specification

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

S.III.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 of the whole innovation process $\zeta_i^{[1-q,T]}$. As for the covariance structure between initial conditions and those factors, we assume that:

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

Consider the covariance matrix of initial conditions Σ :

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

and define,

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

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

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

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

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

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

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

and

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

Hence,

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

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

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

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

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

S.IV Fixed Effect Estimation Method

S.IV.1 Individual specific parameter estimates

The main equation is again:

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

where η_i^c and $v_i^{c[1-p,T]}$ are centered by construction and where a row of $M(\beta)$ is defined as $M(\beta)^{[t]} = (1, t, 1/\beta^t)$ as in Appendix S.III (with some 0s between $1 - p$ and 0). The fixed effect method consists in estimating by Feasible GLS this equation individual by individual using the estimated covariance matrices obtained by random effect methods.

Note that even if the ARMA model is incorrect, those estimates are still consistent when $T \rightarrow \infty$ because what matters is the mean independence of individual effects with respect to factors stated in equation (??), and not the specific form of serial dependence. Their standard errors should, however, be corrected. We use Newey-West robust standard errors in the empirical section. Nonetheless, FGLS relying on serial dependence as estimated by random effects, exploits the information that we have about "aggregate" serial dependence, as opposed to a simple OLS or non-linear least square estimation (Polachek et al, 2015). It enhances the quality of the estimates if the term $\Omega(\eta_i^c)$ in equation (??), is not too heterogeneous and this will be checked after estimation.

We now construct the fixed effect estimate of η_i^c . We consider first the case with no missing values and extend it to the case with missing values.

First, to deal with the correlation between η_i^c and v_i , we can always write:

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

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

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

and:

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

This yields the estimating equation for η_i^c :

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

that we can estimate by GLS methods since D can be estimated using random effect methods.

This yields an infeasible estimator (because B is unknown and should be estimated):

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

in which:

$$B = (D'\Omega_\xi^{-1}D)^{-1}D'\Omega_\xi^{-1}.$$

Furthermore:

$$\tilde{\eta}_i^c = B(D\eta_i^c + \xi_i^{[1-p,T]}) = \eta_i^c + B\xi_i^{[1-p,T]}, \quad (\text{S.IV.2})$$

so that this infeasible estimator is unbiased:

$$E(\tilde{\eta}_i^c | \eta_i^c) = \eta_i^c.$$

The feasible estimator is now given by replacing B by the random effect estimate:

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

Finally, the case with missing values is as follows. 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,

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

and by analogy to results above, we have

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

in which \hat{B}_i is a plug-in estimate of:

$$B_i = (D'S_i'(S_i\Omega_\xi S_i')^{-1}S_i D)^{-1}D'S_i'(S_i\Omega_\xi S_i')^{-1}.$$

S.IV.2 Bias correction

Consistency properties could, however, be misleading since T_i varies in our sample between 4 and 28. To assess the magnitude of the bias, we shall compare the estimates of the covariance matrix of η_i^c that we obtained by random effect and by fixed effect methods by grouping individual profiles according to the length of the observation periods.

We start from showcasing the correction of the variance and turn to quantiles afterwards.

S.IV.2.1 Variances

Using the expression of the unfeasible estimate of individual effects as defined by equation (S.IV.2) above, we get a biased estimate of $V(\eta_i^c)$:

$$\begin{aligned} V(\tilde{\eta}_i^c) &= EV(\tilde{\eta}_i^c | \eta_i^c) + VE(\tilde{\eta}_i^c | \eta_i^c) \\ \implies V(\tilde{\eta}_i^c) &= B\Omega_\xi B' + V(\eta_i^c), \end{aligned} \tag{S.IV.3}$$

in which Ω_ξ is defined in the previous section. The bias term is $B\Omega_\xi B'$ and it is easy to show that the dominating term is of order $1/T_i$.

Our feasible estimate has an additional bias given by the measurement equation,

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

although this term is in $1/\sqrt{N}$ and thus dominated, in large N and moderate T_i samples, by the bias in $1/T_i$.

We estimate the bias in equation (S.IV.3) by replacing, in the expression, $B\Omega_\xi B'$, the unknowns by their corresponding random effect estimates and derive a bias-corrected estimate of the true variance of fixed effects, $V(\eta_i^c)$ (e.g. Arellano and Bonhomme, 2012, and Jochmans and Weidner, 2018).

S.IV.2.2 Quantiles

Let $m_{it} = e_t^\top M(\beta)\eta_i^c$ the period t value of human capital stocks, or the predicted individual specific level of log-wages (net of skill prices). Define $q(\tau) = \inf_q \{q | F(q) \geq \tau\}$ where $F(\cdot)$ is

the distribution function of variable, m_{it} . Realizations of m_{it} are not observed and are replaced by estimates $\hat{m}_{it} = e_t^\top M(\beta)\hat{\eta}_i^c$ in which by the section above we plug in

$$\hat{\eta}_i^c = B\eta_i^c = \eta_i^c + \hat{B}\xi_i^{[1,T]},$$

to derive:

$$\sqrt{T}(\hat{m}_{it} - m_{it}) \xrightarrow{T \rightarrow \infty} N(0, \sigma_{it}^2),$$

in which $\sigma_{it}^2 = \lim_{T \rightarrow \infty} (Te_t^\top M(\beta)B\Omega_\xi B^\top M(\beta)^\top e_t)$.

The estimation of $q(\tau)$ is a specific case of those studied by Jochmans and Weidner (2019). Sort observations in increasing order of values \hat{m}_{it} and define the plug-in estimator for quantiles of the distribution of \hat{m}_{it} by:

$$\hat{q}_t(\tau) = \hat{m}_{[\tau n]}$$

where $[\cdot]$ denotes the integer function and n is the number of individuals whose wage is non-missing at period t . Jochmans and Weidner (2019) argue that a direct analytical bias correction is to be avoided since it requires to non-parametrically estimate the density $f(\cdot)$ of m_{it} . They propose instead to correct the rank τ at which the quantile is evaluated because the sampling error inflates quantiles, by writing:

$$\hat{\tau}(\tau) = \tau + \frac{1}{T}\hat{b}_F(\hat{q}(\tau)),$$

and:

$$\hat{b}_F(q) = -\frac{1}{2nh^2} \sum_i \hat{\sigma}_{it}^2 \kappa' \left(\frac{\hat{m}_{it} - q}{h} \right)$$

in which κ' is the derivative of a Gaussian kernel, h a bandwidth, and $\hat{\sigma}_{it}^2$ is an estimator of σ_{it}^2 .

JW proposed estimate:

$$\tilde{q}(\tau) = \hat{m}_{[\hat{\tau}(\tau)n]},$$

and proved that:

$$\sqrt{n}(\tilde{q}(\tau) - q(\tau)) \longrightarrow N(0, \omega^2(\tau)) \tag{S.IV.4}$$

in which $\omega^2(\tau) = \tau(1 - \tau)/(f(q(\tau)))^2$ in which $f(\cdot)$ is the density of m_{it} .

Table S.i: Non Missing Values

	1977	1979	1980	1985	1986	1987	1988	1989	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	
1977	1																								
1978	1																								
1979	.882	.882																							
1980	.868	.786	.868																						
1982	1	.882	.868																						
1984	1	.882	.868																						
1985	.849	.751	.743	.849																					
1986	.834	.739	.731	.75	.834																				
1987	.804	.714	.704	.718	.737	.804																			
1988	.765	.675	.668	.694	.690	.691	.765																		
1989	.777	.689	.677	.701	.694	.694	.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	.584	.63			
2005	.634	.560	.552	.554	.558	.548	.526	.544	.536	.541	.558	.446	.557	.558	.559	.566	.570	.574	.571	.574	.574	.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 S.ii: 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	.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	.872	.874	.828	.909	.931	.952		

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

Table S.iii: AIC criterion

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

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

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

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

Table S.v: Yearly standard deviation of earnings

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

Table S.vi: Mean individual effects by groups of non-missing periods

Sample periods	$E(\eta_1)$	$E(\eta_2)$	$E(\eta_3)$
(3,13]	-0.184 (0.103)	-0.0179 (0.00916)	0.164 (0.113)
(13,20]	-0.0285 (0.0304)	-0.0115 (0.00386)	0.0607 (0.0381)
(20,25]	0.0113 (0.00932)	-0.00369 (0.00152)	0.0157 (0.0115)
(25,28]	0.0191 (0.00535)	-0.00132 (0.000877)	0.0127 (0.00599)

Note: Heteroskedastic consistent standard errors in parentheses.

Table S.vii: Raw covariance matrix by number of non-missing periods

Sample periods	$Var(\eta_1)$	$Cov(\eta_1, \eta_2)$	$Cov(\eta_1, \eta_3)$	$Var(\eta_2)$	$Cov(\eta_2, \eta_3)$	$Var(\eta_3)$
4	4.7 (889)	0.41 (68)	-5.5 (956)	0.16 (5.3)	-1.1 (74)	9.4 (1029)
5	23 (101)	0.8 (7.6)	-20 (108)	0.1 (0.59)	-1.1 (8.2)	19 (115)
6	21 (45)	1.5 (3.3)	-22 (48)	0.15 (0.25)	-1.8 (3.6)	24 (51)
7	18 (36)	1.3 (2.8)	-18 (39)	0.13 (0.21)	-1.5 (3)	20 (42)
8	20 (27)	1.6 (2.3)	-21 (30)	0.15 (0.19)	-1.8 (2.5)	24 (33)
9	16 (14)	1.3 (1.2)	-17 (16)	0.12 (0.11)	-1.5 (1.4)	19 (18)
10	15 (9.8)	1.4 (0.88)	-17 (11)	0.14 (0.081)	-1.6 (1)	20 (13)
11	9.4 (5.1)	0.86 (0.47)	-11 (5.9)	0.087 (0.044)	-1 (0.54)	12 (6.8)
12	5.6 (2.6)	0.52 (0.25)	-6.3 (3)	0.055 (0.025)	-0.63 (0.29)	7.4 (3.5)
13	3.9 (1.3)	0.37 (0.13)	-4.4 (1.5)	0.039 (0.013)	-0.44 (0.15)	5.1 (1.8)
14	2.9 (0.86)	0.33 (0.1)	-3.6 (1.1)	0.041 (0.012)	-0.43 (0.13)	4.5 (1.4)
15	2.4 (0.72)	0.25 (0.075)	-2.8 (0.86)	0.03 (0.0081)	-0.32 (0.091)	3.4 (1)
16	0.93 (0.24)	0.1 (0.026)	-1.1 (0.29)	0.016 (0.0032)	-0.15 (0.033)	1.5 (0.36)
17	1.1 (0.26)	0.12 (0.032)	-1.3 (0.33)	0.019 (0.0043)	-0.18 (0.043)	1.7 (0.44)
18	0.75 (0.17)	0.089 (0.021)	-0.9 (0.22)	0.014 (0.0029)	-0.13 (0.029)	1.2 (0.3)
19	0.53 (0.085)	0.064 (0.012)	-0.61 (0.11)	0.012 (0.0018)	-0.1 (0.017)	0.91 (0.16)
20	0.33 (0.04)	0.04 (0.0056)	-0.37 (0.052)	0.0084 (0.001)	-0.069 (0.0088)	0.59 (0.077)
21	0.22 (0.029)	0.017 (0.0032)	-0.17 (0.03)	0.0051 (0.00062)	-0.039 (0.005)	0.32 (0.044)
22	0.17 (0.019)	0.018 (0.0031)	-0.16 (0.026)	0.0048 (0.00068)	-0.037 (0.0053)	0.3 (0.042)
23	0.18 (0.017)	0.014 (0.0024)	-0.13 (0.019)	0.0047 (0.00051)	-0.033 (0.0038)	0.25 (0.03)
24	0.16 (0.017)	0.015 (0.0027)	-0.13 (0.022)	0.0049 (0.00061)	-0.036 (0.0047)	0.28 (0.037)
25	0.14 (0.014)	0.011 (0.0023)	-0.098 (0.019)	0.0035 (0.00042)	-0.025 (0.0033)	0.19 (0.027)
26	0.1 (0.0072)	0.0071 (0.00093)	-0.064 (0.0071)	0.003 (0.00023)	-0.02 (0.0016)	0.15 (0.012)
27	0.082 (0.0046)	0.0048 (0.00063)	-0.043 (0.0047)	0.0021 (0.00015)	-0.014 (0.001)	0.099 (0.0072)
28	0.071 (0.0049)	0.0037 (0.00067)	-0.033 (0.0047)	0.0018 (0.00016)	-0.012 (0.0011)	0.081 (0.0073)
Complete sample	2.6 (3.4)	0.22 (0.27)	-2.8 (3.7)	0.024 (0.022)	-0.27 (0.29)	3.3 (4)

Note: Heteroskedastic consistent standard errors in parentheses.

Table S.viii: Bias corrected covariance matrix by number of sampling periods

Sample periods	$Var(\eta_1)$	$Cov(\eta_1, \eta_2)$	$Cov(\eta_1, \eta_3)$	$Var(\eta_2)$	$Cov(\eta_2, \eta_3)$	$Var(\eta_3)$
4	-22 (744)	-1.7 (56)	23 (796)	-0.0065 (4.3)	1.2 (60)	-22 (852)
5	5.9 (69)	-0.54 (5.2)	-1.5 (73)	-0.0091 (0.4)	0.42 (5.6)	-0.98 (78)
6	6.6 (23)	0.48 (1.7)	-6.9 (24)	0.064 (0.13)	-0.62 (1.8)	7.7 (26)
7	3.9 (23)	0.14 (1.7)	-3.3 (24)	0.029 (0.13)	-0.23 (1.8)	3.3 (26)
8	4.4 (16)	0.29 (1.3)	-4.4 (18)	0.034 (0.11)	-0.36 (1.5)	4.8 (19)
9	4.1 (8)	0.29 (0.69)	-4.3 (8.9)	0.027 (0.061)	-0.33 (0.77)	4.6 (9.9)
10	6.2 (6.4)	0.54 (0.57)	-6.9 (7.3)	0.054 (0.052)	-0.64 (0.65)	8 (8.2)
11	4.2 (3.7)	0.35 (0.33)	-4.5 (4.2)	0.035 (0.03)	-0.41 (0.38)	5.2 (4.7)
12	3.3 (2.3)	0.28 (0.21)	-3.6 (2.6)	0.027 (0.02)	-0.32 (0.24)	4 (3)
13	1.6 (1.5)	0.11 (0.15)	-1.5 (1.8)	0.0089 (0.016)	-0.11 (0.18)	1.5 (2.1)
14	1.7 (0.76)	0.18 (0.088)	-2 (0.95)	0.023 (0.011)	-0.24 (0.11)	2.5 (1.2)
15	1.7 (0.65)	0.16 (0.065)	-1.9 (0.76)	0.019 (0.0068)	-0.2 (0.077)	2.3 (0.89)
16	0.73 (0.25)	0.076 (0.026)	-0.81 (0.29)	0.011 (0.0031)	-0.11 (0.033)	1.1 (0.35)
17	0.89 (0.26)	0.1 (0.032)	-1 (0.33)	0.015 (0.0042)	-0.14 (0.042)	1.4 (0.43)
18	0.62 (0.17)	0.071 (0.021)	-0.72 (0.22)	0.01 (0.0028)	-0.097 (0.028)	0.96 (0.29)
19	0.45 (0.085)	0.052 (0.012)	-0.49 (0.11)	0.0089 (0.0018)	-0.077 (0.017)	0.7 (0.16)
20	0.27 (0.04)	0.032 (0.0057)	-0.29 (0.052)	0.0062 (0.001)	-0.051 (0.0087)	0.43 (0.077)
21	0.18 (0.029)	0.012 (0.0032)	-0.13 (0.03)	0.0034 (0.00061)	-0.026 (0.0049)	0.22 (0.043)
22	0.15 (0.019)	0.015 (0.0031)	-0.13 (0.026)	0.0035 (0.00067)	-0.027 (0.0053)	0.22 (0.042)
23	0.16 (0.017)	0.012 (0.0024)	-0.11 (0.02)	0.0035 (5e-04)	-0.024 (0.0038)	0.19 (0.03)
24	0.14 (0.017)	0.014 (0.0027)	-0.12 (0.022)	0.0041 (0.00061)	-0.03 (0.0047)	0.23 (0.037)
25	0.13 (0.014)	0.01 (0.0023)	-0.089 (0.019)	0.0028 (0.00041)	-0.02 (0.0033)	0.16 (0.027)
26	0.097 (0.0072)	0.0066 (0.00093)	-0.059 (0.0072)	0.0025 (0.00023)	-0.017 (0.0016)	0.12 (0.012)
27	0.077 (0.0046)	0.0046 (0.00064)	-0.04 (0.0047)	0.0017 (0.00015)	-0.011 (0.001)	0.079 (0.0074)
28	0.067 (0.0049)	0.0036 (0.00067)	-0.031 (0.0047)	0.0015 (0.00016)	-0.0097 (0.0011)	0.067 (0.0074)
Complete sample	0.96 (2)	0.07 (0.16)	-0.95 (2.1)	0.0096 (0.013)	-0.094 (0.17)	1.1 (2.3)
Random effects	0.093 (0.0036)	0.0059 (0.00051)	-0.05 (0.004)	0.0015 (0.00011)	-0.0093 (0.00079)	0.066 (0.0059)

Note: Heteroskedastic consistent standard errors in parentheses.

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

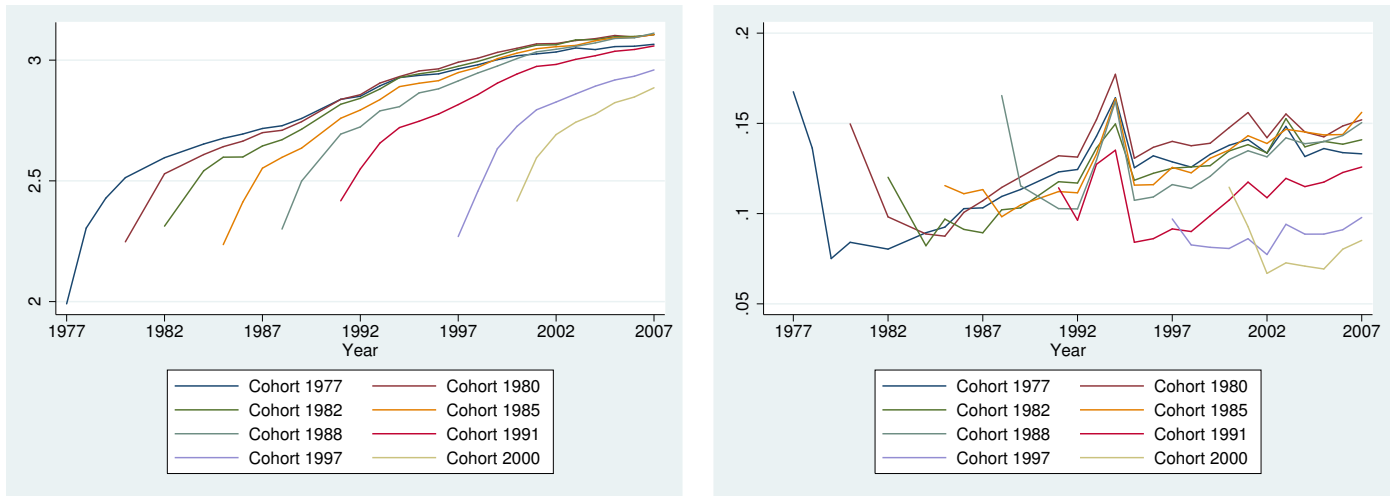


Figure S.ii: First order autocorrelation relative to potential experience for 1977, 1987 and 1997 entry cohorts

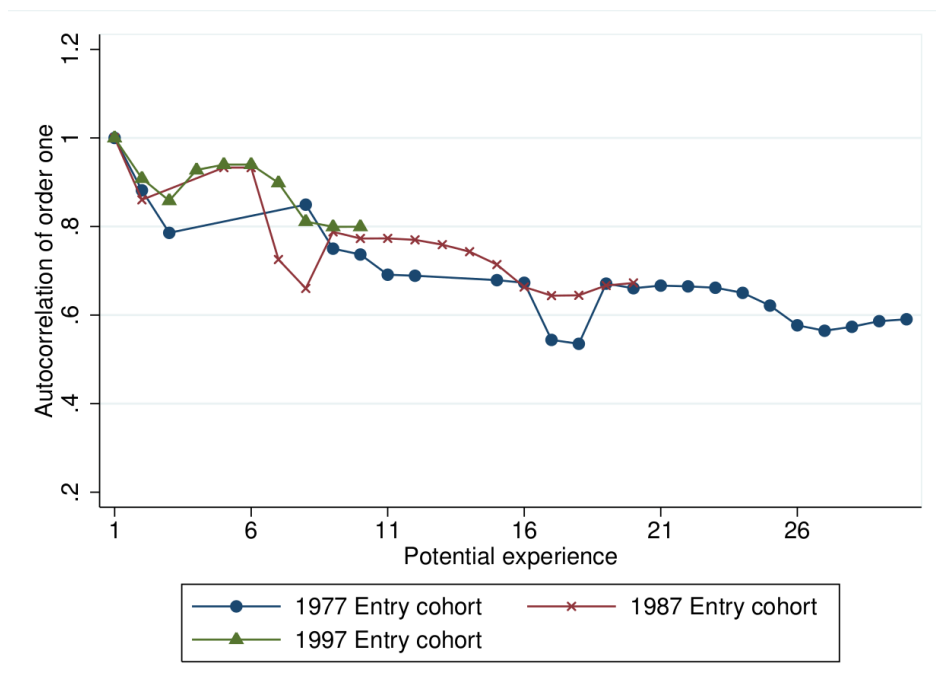


Figure S.iii: Autocorrelations with 1986 and 2007

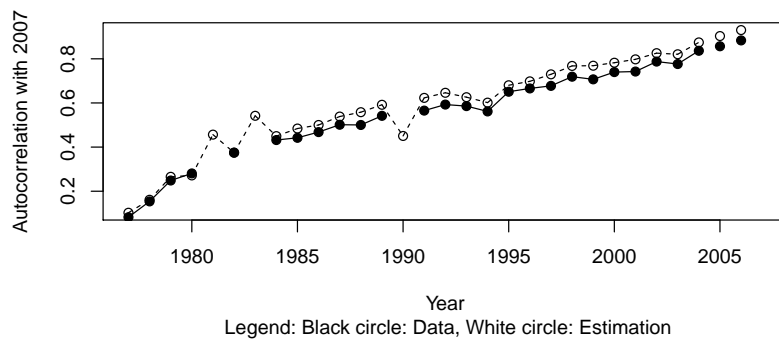
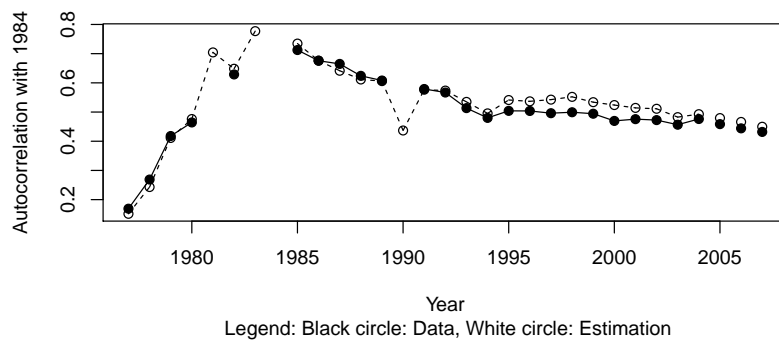


Figure S.iv: Forward autocorrelations of order 1 and of order 6

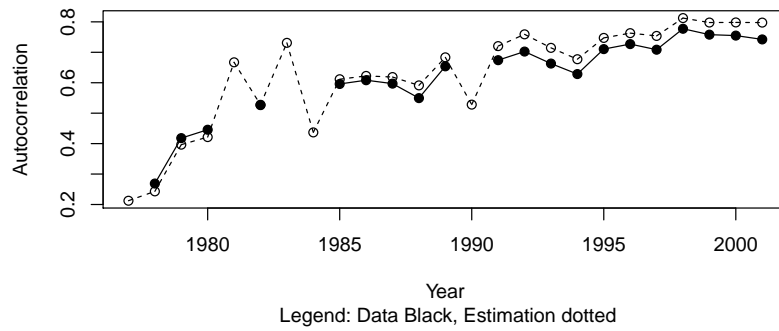
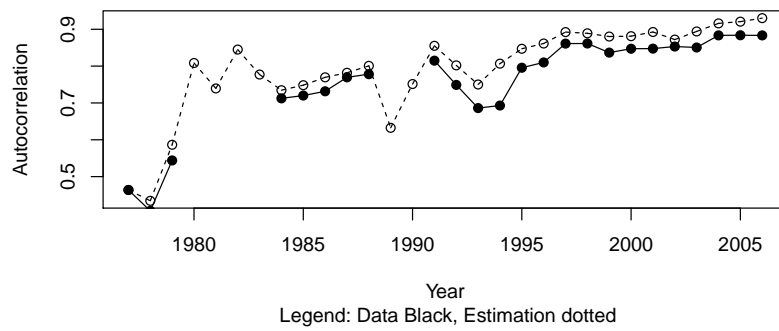


Figure S.v: Estimated variance of residuals v_{it} by potential experience and skill group

