The Public Pay Gap in Britain: Small Differences That (Don't?) Matter.

Fabien Postel-Vinay*
PSE, Paris
CREST-INSEE, Malakoff
CEPR, London
and IZA, Bonn

Hélène Turon†
Department of Economics,
University of Bristol
and IZA, Bonn

January 2005—First draft

Abstract

The existing literature on inequality between private and public sectors focuses on cross-section differences in earnings levels. A more general way of looking at inequality between sectors is to recognize that forward-looking agents will care about income and job mobility too. We show that these are substantially different between the two sectors. Using data from the BHPS, we estimate a model of income and employment dynamics over seven years. We allow for unobserved heterogeneity in the propensity to be unemployed or employed in either job sector and in terms of the income process. We then combine the results into lifetime values of jobs in either sector and carry out a cross-section comparative analysis of these values. We have four main findings. First, the public premium in the present discounted sum of future log-income flows is on average 7 percent. Second, most of the observed relative income compression in the public sector is due to a lower variance of the *transitory* component of income. Third, when taking job mobility into account, the lifetime public premium is essentially zero for workers that we categorize as "high-employability" individuals, suggesting that the UK labor market is sufficiently mobile to ensure a rapid allocation of workers into their "natural" sector. Fourth, we find some evidence of job queuing in the public sector for "low-employability" workers.
Keywords: Income Dynamics, Job Mobility, Public-Private Inequality, Selection Effects.

JEL codes: J45, J31, J62.

*Address: PSE, Ecole Normale Supérieure, 48 boulevard Jourdan 75014 Paris FRANCE. Tel. +33 1 4313 6302. E-mail fpostel@pse.ens.fr. Internet www.delta.ens.fr/postel-vinay/index.html. PSE (ParisJourdan Sciences Economiques) is a CNRS-EHESS-ENPC-ENS joint research unit.

†Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN. Tel: +44 117 928 8400. E-mail helene.turon-lacarrieu@bris.ac.uk. Internet www.efm.bris.ac.uk/economics/staff/turon/.

1 Introduction [V. preliminary]

There is an ongoing debate as to whether public sector workers experience better or worse pay conditions than their private counterparts.[1] A recurring message in the recent literature is that, although raw differences can be large, the bulk of these is likely to reflect differences in the composition of the workforce, both in terms of observed characteristics such as age and education, as unobserved characteristics. Nevertheless, a look at labor markets in the two sectors suggests that individuals do perceive some differences between the sectors: Katz and Krueger (1991) document the fact that blue collars are willing to queue to obtain public sector jobs whereas highly-skilled workers are notoriously hard to recruit and retain in the public sector.

The focus of most (if not all) of the literature on public pay gaps is on cross-section differences in earnings levels. Yet the public-private differences are arguably equally marked in terms of income mobility, income volatility and job loss risk than in terms of mean income levels. Public sector employers are widely perceived to offer more job security. There is also plentiful evidence of relative wage compression in the public sector. Moreover, as we shall point out, public sector incomes are substantially more persistent over time than private sector incomes.

The motivating point of this paper is that forward-looking agents are likely to care about income and job mobility as well as income levels, so that a more general assessment of the existence and magnitude of a "public premium" should be based on measures of expected lifetime utility derived from employment in either sector, rather than on comparisons of instantaneous income flows.

Pursuing this idea, we estimate simultaneously individual income processes, job loss risks and selection patterns into the public and private sector using a rich dynamic model that allows for unobserved heterogeneity in the propensity of individuals to work in the public sector or to become unemployed, and in individual patterns of income levels, income mobility and income volatility. We then combine the results into lifetime values of jobs in either sector and carry out a cross-section comparative analysis of these values.

[1] See for example Disney and Gosling (1998, 2003) for the UK, Moulton (1990) and Borjas (2002) for the US, and the large number of studies surveyed in Bender (1998).

We estimate this model with data from the BHPS over the period 1996-2002. We obtain a very good model fit in terms of job mobility, income distributions and income mobility, both in and out-of-sample (up to 10 years). Our results concur with the existing literature in showing that there is a small public pay premium in cross-sections, but that most of the observed "raw" difference is due to selection. There are also marked public-private differences in patterns of job and income mobility: job loss rates are lower in the public sector; income volatility and cross-sectional income variance are lower in the public sector; and returns to experience tend to be slightly smaller in the public sector.

We have four main findings. First, once selection has been accounted for, the value of a job for life is on average 7 percent higher in the public sector than in the private sector when defined as the present discounted sum of future log-income flows. Second, most of the observed relative income compression in the public sector is due to a lower variance of the *transitory* component of income, against which there is potential scope for insurance. Third, when combining all features of employment in either sector into lifetime values, the lifetime public premium is essentially zero among workers that we categorize as "high-employability" individuals based on their low unobserved propensity to become unemployed. The reason is that the UK labor market is sufficiently mobile to ensure a rapid allocation of these workers into their "natural" sector. Fourth, we find some evidence of job queuing in the public sector for "low-employability" individuals, whom we estimate to face large potential premia from public sector employment

Some related literature is reviewed in the next Section. We carry out a descriptive analysis of our data in Section 3. The statistical model to be estimated is detailed in Section 4 and results are discussed in Section 5. Finally, we compute lifetime values in Section 6 and show public-private differences taking into account wage and job mobility and compare it with cross-section wage differences. Section 7 concludes.

2 Related literature

`[Forthcoming.]`

3 Descriptive data analysis

3.1 Sample construction and basic description

We use data from waves 1996 to 2002 of the British Household Panel Survey (BHPS). We restrict our sample to males who are in the BHPS in 1996 or enter the BHPS after 1996 and who have no subsequent gap in their response history.[2] The sample is restricted to males in order to avoid issues of labor supply, such as part-time work or labor force attachment. We define three 'sectors' of activity or labor force status: employment in the public sector, employment in the private sector, and unemployment.[3] Our aim is to estimate earnings in the private and public sector as well as in unemployment. To this end we use total monthly income (reported for the month preceding the survey date), which includes incomes from labor, benefits, transfers, pensions and investment, deflated with the CPI. We trim the income data by treating income observations below the 2nd and above the 98th percentile of income within each 'education'×'job sector' cell as missing data.[4] We then keep data for individuals aged from 20 to 60 at the beginning of the panel, and we exclude retired individuals. We hence do not model the transition into (or the earnings change after) retirement. This leaves us with 3,791 men, most of whom we follow over 7 years.[5]

A breakdown of the sample by education groups (where 3 education groups are distinguished based on highest academic qualification: "low" is O-level or less, "medium" is A-level, and "high" is above A-level) goes as follows: 25.4% high, 21.3% medium and 53.2% low. Turning to job sectors, we find that, in our initial observation year of 1996, 77.3% of the individuals in our sample were holding a private sector job, 16.2% a public sector job, and 6.5% were unemployed. We should also point out the substantial amount of selection going on, as the education shares of the public (resp.

[2] There are in fact 12 BHPS waves available covering the period 1991-2002. We restrict the estimation sample to the lat 7 waves for two reasons. First, for reasons discussed at length in later sections, we want to estimate the model over a reasonably period of time which is reasonably representative of an "average" state of the business cycle. Second, we want to spare a few years of data for out-of-sample prediction testing (see Section 5).

[3] Private firms, self-employed workers and non-profit organizations are classified as private sector. Civil servants, employees in central and local government, town halls, the NHS, High Education, nationalized industries and in the armed forces are classified as public sector.

[4] That is, we only drop the corresponding *income* observations. Individuals concerned by this trimming thus still contribute to the sample and convey information about the selection process into particular job sectors. We also treat any reported monthly income below £50 (in 1996 £) as unobserved. This last bit of trimming only affects unemployed people. £50 is about 15% of the first percentile of income among employed workers.

[5] There is of course some attrition (about 10% at 4 years, 23% at 7 years), which we assume exogenous to the processes of labor market and income histories.

private) sector are 41.5% high, 21.99% medium and 36.5% low (resp. 22.9% high, 21.7 medium and 55.4% low). The public sector thus attracts markedly better educated workers.

< Figure 1 about here. >

To see how the sector composition of the British labor market evolves over time, Figure 1 reports the public sector share of total employment (left panel) and the unemployment rate (right panel) among British males aged 20-60 over the period 1991-2002 which covers our 1996-2002 observation window. Both are fairly stable over that period: the unemployment rate drops in the aftermath of the 1993 recession, then stabilizes around 6% in 1996. The employment shares of both sectors exhibit variations of a magnitude of about 2.5 percentage points, with no apparent trend over our sample period. The employment share of the public sector in the UK has decreased substantially over the past two decades, through privatizations, which took place mainly in the 1980s, and through contracting out the provision of public services to the private sector (see Disney and Gosling, 2003, p.3). While this raises several measurement issues, we shall take for granted that things are largely stabilized by the end of the 1990s.

3.2 Differences in incomes

Income levels. Table 1 contains a descriptive account of public-private differences in incomes, in the form of simple regressions. The first column of Table 1 shows that the raw public pay gap over our sample period is 13.8 log points (about 14.8%) in favor of the public sector, while unemployment incomes equal on average 21% private sector income. Conditioning on a quadratic in (potential) labor market experience and on qualifications (specification 2), the public-private sector gap becomes 4.9% and remains statistically significant. This is comparable with results found by Disney and Gosling (1998, p.354) where this conditional gap was estimated to be 5% with 1983 GHS data and 1% with BHPS data pooled over the 1990s waves. An intermediate value of 8.3% is found for the public premium when allowing for individual fixed-effects (specification 4).[6,7] Allowing for

[6] The reported fixed-effect regressions are run using within estimators. First-difference estimates are very similar.

[7] Disney and Gosling (1998) also found the raw public sector pay gap was reduced when allowing for fixed effects. They suggest, however, that the probability of moving between sectors is likely to be correlated with any difference in unobserved characteristic causing pay difference between public and private sector, so that the pay gap suggested by the fixed-effect regression result may still be misleading. In our statistical model presented in the next Section,

different effects of education by sector (specification 3), we find a significantly higher public pay premium for less educated workers. Finally allowing different returns to experience by sector, we observe non-significant such differences in an OLS regression (specification 3), which become significant again as one considers individual fixed-effects (specification 5). Believing the latter specification, we conclude that returns to experience *are* significantly smaller in the public sector, with a public sector income premium describing a U-shape with experience (with an estimated minimum at 25 years of experience). In other words, the public pay premium seems to mostly benefit less experienced workers.

< Table 1 about here. >

Income dispersion. In order to complement the description provided by Table 1, we finally give a brief account of public-private differences in income dispersion. The standard deviation of log income is smaller (0.38 *vs.* 0.51) in the public than in the private sector, and largest in unemployment (where is equals 0.63). The 90:10 percentile ratio of raw incomes are 2.79 and 3.55 respectively for public and private sector. The corresponding figures for incomes conditional on age and qualifications are 2.40 for the public sector and 3.14 for the private sector. Again, these figures are comparable to Disney and Gosling's results: they find the ratios of 90:10 percentile ratios of raw incomes to be 2.7 and 2.96 for public and private sector respectively and the corresponding conditional figures to be 2.38 and 2.61. The upcoming analysis will finally confirm that log-income variance is constantly smaller in the public sector across all levels of experience (in fact, in either sector, cross-sectional income variance varies little with experience). All this is consistent with numerous findings on wage compression in the public sector.

Income mobility. The private and public sector not only differ in terms of cross-sectional income distributions, as we just saw, but also in terms of income mobility. Differences in income mobility are illustrated in Table 2 by the transition matrices between income quintiles from one year to

we allow for unobserved heterogeneity in the propensity of individuals to work in the public sector and estimate simultaneously the income processes and the selection into sectors and employment. As discussed below, we also allow for unobserved heterogeneity in income and income mobility.

the next and at a six-year lag for the public and private sector respectively.[8] We observe that income ranks are more persistent from one year to the next in the public than in the private sector. Transition matrices of income mobility between quintiles over periods of 2 to 5 years convey the same message. The contrast is most marked in the comparison of income quintiles between the first and the last waves, for individuals who have remained in the same sector throughout: over six years, individuals in any income quintile are more likely to remain in the same quintile when continuously in the public sector than in the private sector.

< Table 2 about here. >

In terms of the dynamics of income *levels*, since the income distribution is more compressed in the public than in the private sector, transitions between quintiles in the private sector in fact suggest larger wage drops or increases than similar transitions in the private sector. To further illustrate persistence of income levels, we regressed the correlation between current normalized log income and past normalized log income on a quadratic in experience, education, and interactions of current and past sector of employment.[9] Results show that income and past income are significantly more correlated among workers employed in the public sector for two consecutive years than for workers changing sectors or staying in the private sector, thus suggesting less overall income stability in the private sector.

3.3 Differences in job mobility

The following transition matrix illustrates changes in employment sector between one wave and the next (rows refer to the initial sector and columns to the final sector):[10]

---

[8] These figures relate to individuals who are employed in the same sector in year t and in year $t+1$. This wage mobility hence abstract from wage changes caused by job transitions (from one sector to the other or to unemployment).

[9] We constructed normalized log income using specification 2 in Table 1. Specifically, we first regressed log income (say, y_{it} for individual i at date t) on our chosen set of covariates thus obtaining a predictor of mean income $\widehat{y}_{it}$, then regressed the squared residuals from this latter regression on that same set of covariates thus obtaining a predictor of income variance $\widehat{\sigma}_{it}^2$. We then constructed normalized income as $(y_{it} - \widehat{y}_{it}) / \sqrt{\widehat{\sigma}_{it}^2}$.

[10] We should emphasize that, throughout the paper, we shall be talking about these are *year-to-year, between-sector transitions*. We thus overlook multiple transitions occurring within a single year, or job changes within the same sector.

	Private	Public	Unemp.
Private	96.9	1.5	1.5
Public	8.2	90.1	0.8
Unemp.	41.0	4.1	54.9

Very few (1.5%) individuals initially employed in the private sector move jobs to the public sector while movements in the opposite direction are slightly more frequent (8.2%). The raw job loss rates are 1.5% for individuals initially in the private sector against 0.8% for individuals initially employed in the public sector. Of those initially in the public sector, 2.8% are recorded as unemployed at least once over the 6 following waves as opposed to 5% for individuals initially in the private sector.[11] Of those initially unemployed, 41% find employment in the private sector within a year and only 4.1% in the public sector. Another 54.9% are still unemployed at the next interview date. A little under 40% of the unemployed report being unemployed at more than two consecutive interview dates over our sample. These "long-term unemployed" only have a 30% probability of finding a job within a year, whereas "short-term" unemployed have a 52% chance of finding a job.[12]

These figures bring about some comments. First, unemployment persistence seems high but is comparable to results found by Stewart (2004) with similar data. Second, the average job loss rates in the public sector is just a little over half the average job loss rate of the private sector. To further investigate this issue, we ran a logit regression of the probability of being unemployed on education, a quadratic in experience, previous labor market state and interaction terms of experience and past labor market state. This exercise suggests that the public-private gap in job loss rates describes a U-shape over working life, reaching its (still positive) minimum value between 20 and 30 years of experience.[13] Hence, much like the public pay premium, the public premium in lower job loss rates mostly benefits workers who are at early stages of their careers.

3.4 Summary

The above descriptive analysis thus highlights the following facts:

[11] The former number is however to be treated with caution as it relates to only 17 individuals employed in the public sector in 1996 and experiencing at least one unemployment spell before 2002.

[12] One should however bear in mind that, of those reporting being unemployed at two consecutive interview dates, 28% have in fact found a job and lost it again in the past year.

[13] Even though these effects are of weak statistical significance. We do not report the results in any more detail (they are available upon request), yet an illustration is available on the bottom-right panel of Figure 2, which plots the raw job loss rates in both employment sectors as a function of experience.

- The public and private sector differ not only in the mean income they offer to their employees, but also in terms of income and job mobility.
- Raw figures show that the public pay premium is higher for less educated workers. Fixed-effect regressions suggest that it is also higher at low levels of labor market experience and that the returns to experience are higher in the private sector at low levels of experience and lower toward the end of the working life.
- There is substantial pay compression in the public sector at all levels of experience. There is also less income mobility in the public sector.
- The probability of loosing one's job is higher in the private sector at all levels of experience. Yet the gap is wider at both extremes of the working life than around mid-career.

Income and employment dynamics as well as income levels are hence quite different between the two sectors. All this would matter to forward-looking individuals. As was repeatedly advocated by Gottschalk and Moffitt (1993) and Gottschalk (1997), comparisons of cross-sections are not very informative in the presence of income mobility. Here we contend that their point is even more relevant in the case of cross-sector comparisons when there are cross-sector differences in income mobility. It is thus desirable to use a criterion that takes account of all aspects of the differences between sectors. This would give a more comprehensive and accurate picture of the comparison between employment in the public and the private sector. It is what the rest of this paper is devoted to.

4 The model

4.1 General structure

The data is a set of N workers indexed $i = 1, \dots N$, each of whom we follow over T_i consecutive years. Each year we observe individual job states and monthly earnings. The data also convey information about individual fixed characteristics. Thus a typical observation for an individual $i = 1, \dots, N$ can be represented as a vector[14] $\mathbf{x}_i = \left(\mathbf{y}_i, \mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^v, z_i^f\right)$, where:

[14]Throughout the paper we use boldface characters to denote vectors.

- $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT_i})$ is the observed sequence of inidividual i's log income flows.

- $\mathbf{e}_i = (e_{i1}, \ldots, e_{iT_i})$ is the observed sequence of individual i's unemployment episodes. Specifically, $e_{it} = 0$ if individual i is unemployed in period t, and 1 otherwise.

- $\mathbf{pub}_i = \left(\text{pub}_{i1}, \ldots, \text{pub}_{iT_i}\right)$ is the observed sequence of individual i's job sectors. Here $\text{pub}_{it} = 1$ if individual i is employed in the public sector in period t, and 0 if he or she is employed in the private sector. Note that pub_{it} is only defined if $e_{it} = 1$.

- $\mathbf{z}_i^v$ is a sequence of time-varying individual characteristics. In our application we only consider (polynomials in) potential labor market experience, defined as the current date less the date at which individual i left school.

- Finally, z_i^f is a set of individual fixed characteristics. It includes highest academic qualification (3 levels) and labor market cohort. By "labor market cohort" we mean the year in which the individual first entered the labor market. As a consequence, $\mathbf{z}_i^v$ is deterministic conditional on z_i^f.

On top of observed individual heterogeneity captured by $\mathbf{z}_i^v$ and z_i^f, we also recognize that unobserved individual characteristics may influence wages or selection into the various labor market states. At this point we remain general and only supplement the data vector $\mathbf{x}_i$ by appending a set k_i of such (time-invariant) unobserved characteristics. The aim of the model is to estimate simultaneously transitions between unemployment and employment, transitions between public and private sector, and income trajectories within and between employment sectors. To this end, and given the definitions above, we define individual contributions to the *complete* likelihood—i.e. the likelihood of $(\mathbf{x}_i, k_i)$, including unobserved variables—as:

$$\mathcal{L}_i\left(\mathbf{x}_i, k_i\right) = \ell_i\left(\mathbf{y}_i \mid \mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^v, z_i^f, k_i\right) \cdot \ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i\right) \cdot \ell_i\left(k_i \mid z_i^f\right) \cdot \ell\left(z_i^f\right). \tag{1}$$

The typical individual likelihood contribution is thus decomposed into four terms. The last one, $\ell\left(z_i^f\right)$ is simply the sample distribution of observed individual characteristics z_i^f.[15] This distribution is observed and is independent of any parameter. We shall therefore omit it from now on.

[15]Recall that $\mathbf{z}_i^v$ is deterministic conditional on z_i^f and therefore doesn't appear in the likelihood function.

The next to last term, $\ell\left(k_i \mid z_i^f\right)$, is the distribution of unobserved individual heterogeneity given observed characteristics z_i^f. Finally, the first two terms in (1) are the likelihood of individual earnings and labor market state paths given individual heterogeneity. We further decompose it into the likelihood of labor market states, $\ell\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i\right)$, and the likelihood of earnings sequences given labor market states $\ell\left(\mathbf{y}_i \mid \mathbf{e}_i, \mathbf{pub}_i \mathbf{z}_i^v, z_i^f, k_i\right)$. Note that, even though this dependence was kept implicit in order not to overload the equations, the first three terms in (1) depend on (various subsets of) model parameters. We shall thus obtain estimates of those parameters by maximizing the sample log-likelihood, $\sum_{i=1}^N \log\left(\int \mathcal{L}_i\left(\mathbf{x}_i, k_i\right) dk_i\right)$. Before we proceed, however, we have to go through a precise description of our modeling assumptions concerning the various components of (1).

4.2 Heterogeneity

We begin with individual heterogeneity, i.e. the third component of $\mathcal{L}_i\left(\mathbf{x}_i, k_i\right)$ in (1). As we already mentioned, we allow for both unobserved (k_i) and observed (z_i^f) heterogeneity. We next consider two types of unobserved heterogeneity, $k_i = (k_i^m, k_i^y)$. The first type, k_i^m, relates to heterogeneity in terms of propensity to be unemployed or to work in the public sector (called mobility classes hereafter). The second type, k_i^y, relates to heterogeneity in terms of income (called income classes hereafter) through its impact on both income distributions and income mobility. k_i^m conditions the parameters relating to employment and sector history, while k_i^y conditions the parameters relating to income distribution and income mobility. Both types of heterogeneity are time-invariant individual random effects, which we allow to be correlated in an arbitrary fashion. The "mobility" dimension of heterogeneity k_i^m deals with the selection problem outlined in Section 3. Concerning k_i^y, because the propensity to belong to a given income class is invariant over time, this latter type of heterogeneity increases the persistence of income ranks, which is found to be underestimated otherwise[16]. We refer to income and mobility *classes* because we use a finite mixture approach to model unobserved heterogeneity where an individual can belong to one of K^m classes of mobility

[16]This is commonly found in the literature on income mobility. Shorroks (1976) for example shows that actual earnings processes are more persistent than earnings predicted with first order Markov processes. More on this below.

and K^y income classes.[17] The total number of classes is hence $K = K^m \times K^y$. The probability of belonging to a given class depends on observed individual heterogeneity z_i^f, as follows:

$$\Pr\left\{k_i^m, k_i^y \mid z_i^f\right\} = \Pr\left\{k_i^y \mid k_i^m, z_i^f\right\} \cdot \Pr\left\{k_i^m \mid z_i^f\right\}. \tag{2}$$

The reason why we adopt this particular decomposition of the joint distribution of $\left(k_i^m, k_i^y \mid z_i^f\right)$ will become clear as we describe our likelihood maximization algorithm. As for practical matters, we model both components of (2) as multinomial logits, with respectively K^y and K^m outcomes. All formal details are gathered in Appendix A.

4.3 Job mobility

We now turn to the second component of $\mathcal{L}_i(\mathbf{x}_i, k_i)$ in (1), which pertains to job mobility. We model transitions between three distinct labor market states—namely unemployment, employment in the private sector and employment in the public sector—but not within each state. That is, an individual who has changed jobs within the private sector over the past year for example will not be analyzed as having experienced a transition. Transition probabilities are assumed to depend only on the individual's employment status in the previous period and on observed and unobserved heterogeneity—i.e. job states are assumed to follow a (conditional) first-order Markov chain. Employment histories are modeled in two stages: the probability of unemployment, $e_{it} = 0$, and the probability of working in the public sector ($\text{pub}_{it} = 1$) given employment ($e_{it} = 1$) are specified separately as follows:[18]

$$\begin{aligned}
&\Pr\left\{e_{it}, \text{pub}_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} \\
&\quad = \Pr\left\{e_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} \times \left[\Pr\left\{\text{pub}_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\}\right]^{e_{it}}.
\end{aligned} \tag{3}$$

In practice, we model both components of (3) as logits. Again, all formal details are confined to Appendix A. Finally, we face a standard initial conditions problem in that we have to specify

[17] Finite mixture approximations provide convenient and increasingly popular tools to account for unobserved heterogeneity. Heckman and Singer (1984) is the pioneering reference in the field of economics. See also Keane and Wolpin (1997) or Eckstein and Wolpin (1999) for recent applications in labor economics.

[18] Note the assumption implicit in (3) that only the date-$(t-1)$ component of $\mathbf{z}_i^v$—i.e., individual i's potential experience at date $t-1$—enters the set of conditioning variable for job mobility between dates $t-1$ and t.

the distribution of the initial labor market state, $(e_{i1}, \text{pub}_{i1})$. We let it depend on observed and unobserved heterogeneity as follows:

$$\Pr\left\{e_{i1}, \text{pub}_{i1} \mid z_i^f, k_i^m\right\} = \Pr\left\{e_{i1} \mid z_i^f, k_i^m\right\} \times \left[\Pr\left\{\text{pub}_{i1} \mid z_i^f, k_i^m\right\}\right]^{e_{i1}}. \tag{4}$$

Both components are again specified as logit models (see Appendix A). Summing up, the contribution to the likelihood of an individual job mobility trajectory is:

$$\begin{aligned}\ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i^m\right) = &\Pr\left\{e_{i1}, \text{pub}_{i1} \mid z_i^f, k_i^m\right\} \\ &\times \prod_{t=1}^{T_i-1} \Pr\left\{e_{i,t+1}, \text{pub}_{i,t+1} \mid e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^m\right\},\end{aligned} \tag{5}$$

where the components of the latter product are given by (4) and (3), respectively.

4.4 Income process

We finally turn to the more intricate derivation of the first term in $\mathcal{L}_i(\mathbf{x}_i, k_i)$ (equation (1)), which involves the modeling of individual income paths. We consider log-income y_{it} both in employment and in unemployment and assume log-income trajectories $\mathbf{y}_i$ to be the realization of a Markov process of continuous random variables Y_t. Even though we shall (briefly) experiment with various specifications, our preferred option as to the order of the Markov process is second-order (more extensive discussion of this point in Section 5). We will therefore present the earnings part of our statistical model under this particular assumption.[19] Temporarily omitting any conditioning variable or individual index, the likelihood of a given income trajectory over T periods can be written as:

$$\ell(\mathbf{y}) = \ell(y_1, y_2) \cdot \prod_{t=3}^{T} \ell(y_t \mid y_{t-1}, y_{t-2}) = \ell(y_1, y_2) \cdot \prod_{t=3}^{T} \frac{\ell(y_t, y_{t-1}, y_{t-2})}{\ell(y_{t-1}, y_{t-2})}. \tag{6}$$

A convenient and intuitive way of specializing (6) is to decompose the joint density $\ell(y_t, y_{t-1}, y_{t-2})$ as the product of marginal densities of log income y_t, denoted $f_t(y_t)$, $f_{t-1}(y_{t-1})$ and $f_{t-2}(y_{t-2})$ and a copula density of income ranks $F_t(y_t)$, denoted $c_{t,t-1,t-2}\left[F_t(y_t), F_{t-1}(y_{t-1}), F_{t-2}(y_{t-2})\right]$. The

[19]Note that the longitudinal dimension of our panel is longer than needed to identify this type of income process. Intuitively, we require three years of data to characterize the second-order Markov process, plus one additional year for unobserved heterogeneity, hence four years in total; our panel length is seven years.

copula density describes the mobility of marginal income ranks between $t-2$ and t.[20] Similarly, we can write the joint density $\ell(y_t, y_{t-1})$ appearing in the denominator of (6) as $f_t(y_t) \cdot f_{t-1}(y_{t-1}) \cdot c_{t,t-1}[F_t(y_t), F_{t-1}(y_{t-1})]$, where $c_{t,t-1}(\cdot)$ is now the joint density of income ranks at dates t and $t-1$. Using these decompositions, the likelihood in (6) now becomes:

$$\ell(\mathbf{y}) = \left[\prod_{t=1}^{T} f_t(y_t)\right] \times \left[\prod_{t=3}^{T} \frac{c_{t,t-1,t-2}[F_t(y_t), F_{t-1}(y_{t-1}), F_{t-2}(y_{t-2})]}{c_{t,t-1}[F_t(y_t), F_{t-1}(y_{t-1})]}\right] \times c_{2,1}[F_2(y_2), F_1(y_1)]. \quad (7)$$

The use of copulas thus allows us to decompose the likelihood of an income trajectory into the product of the cross-section densities at each time period and the likelihood of the trajectory of relative income ranks within these cross-section distributions. As we assume income processes to be second-order Markov processes, the expression of the likelihood only involves products of bi- or tri-variate densities. Equations (6) and (7) were derived omitting conditioning variables. We now bring them back, and to this end we need to specify the model a little further (once again, a fully detailed presentation of all our specification assumptions is available in Appendix A). We assume marginal income distributions to be normal, conditional on observed and unobserved individual heterogeneity. That is, both mean and variance are allowed to depend on observed and unobserved heterogeneity as well as on current labor market status:

$$\begin{aligned} y_t \mid e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y \quad &\sim \quad \mathcal{N}(\mu_{it}, \sigma_{it}^2) \\ \text{with } \mu_{it} = \mu\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right) \quad &\text{and} \quad \sigma_{it} = \sigma\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right), \end{aligned} \quad (8)$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are given functions (see Appendix A for a complete specification).

Given the assumed normality of the marginal distributions, we choose to specify our various copulas as Gaussian.[21] Given (7), we are interested in bi- and tri-variate copula densities involving consecutive values of log income. We denote $\widetilde{y}_{it}$ the normalized log-income, $\widetilde{y}_{it} = \frac{y_{it}-\mu_{it}}{\sigma_{it}}$. Here,

[20] Copulas are a standard element of the empirical finance toolkit. They were recently brought to labor economics by Bonhomme and Robin (2004).

[21] The (n-variate) Gaussian copula has cdf $C_n(u_1, \ldots, u_n; \underline{\tau}^n) = \Phi_n\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n); \underline{\tau}^n\right)$, where $\Phi_n(\cdot; \underline{\tau}^n)$ is the n-variate normal cdf with mean 0, unit variance and correlation matrix $\underline{\tau}^n$, and $\Phi(\cdot)$ is the standard normal cdf. The corresponding copula density is thus defined by

$$c_n(u_1, \ldots, u_n; \underline{\tau}^n) = \frac{\varphi_n\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n); \underline{\tau}^n\right)}{\varphi(\Phi^{-1}(u_1)) \cdot \ldots \cdot \varphi(\Phi^{-1}(u_n))},$$

where the φ's designate the density functions of the corresponding Φ's.

as all $\widetilde{y}_{it}$'s are themselves distributed standard normal, the expressions for the copula densities of interest become:

$$\begin{aligned} c_{t,t-1,t-2}\left[F_t\left(y_{it}\right),F_{t-1}\left(y_{i,t-1}\right),F_{t-2}\left(y_{i,t-2}\right)\right] &= c_3\left[\Phi\left(\widetilde{y}_{it}\right),\Phi\left(\widetilde{y}_{t-1}\right),\Phi\left(\widetilde{y}_{t-2}\right);\underline{\tau}_t^3\right] \\ &= \frac{\varphi_3\left(\widetilde{y}_{it},\widetilde{y}_{i,t-1},\widetilde{y}_{i,t-2};\underline{\tau}_t^3\right)}{\varphi\left(\widetilde{y}_{it}\right)\cdot\varphi\left(\widetilde{y}_{i,t-1}\right)\cdot\varphi\left(\widetilde{y}_{i,t-2}\right)} \end{aligned} \tag{9}$$

and

$$c_{t,t-1}\left[F_t\left(y_{it}\right),F_{t-1}\left(y_{i,t-1}\right)\right]=c_2\left[\Phi\left(\widetilde{y}_{it}\right),\Phi\left(\widetilde{y}_{t-1}\right);\underline{\tau}_{it}^2\right]=\frac{\varphi_2\left(\widetilde{y}_{it},\widetilde{y}_{i,t-1};\underline{\tau}_{it}^2\right)}{\varphi\left(\widetilde{y}_{it}\right)\cdot\varphi\left(\widetilde{y}_{i,t-1}\right)}. \tag{10}$$

We expand the correlation matrices as:

$$\underline{\tau}_{it}^3=\begin{pmatrix}1 & \tau_{i,t,t-1} & \tau_{i,t,t-2}\\ \tau_{i,t,t-1} & 1 & \tau_{i,t-1,t-2}\\ \tau_{i,t,t-2} & \tau_{i,t-1,t-2} & 1\end{pmatrix} \quad \text{and} \quad \underline{\tau}_{it}^2=\begin{pmatrix}1 & \tau_{i,t,t-1}\\ \tau_{i,t,t-1} & 1\end{pmatrix}. \tag{11}$$

Elements of $\underline{\tau}_{it}^n$, i.e. $\tau_{i,t,t-1}$ and $\tau_{i,t,t-2}$, are in fact the correlation coefficients between $\widetilde{y}_{it}$ and $\widetilde{y}_{i,t-1}$ (resp. $\widetilde{y}_{it}$ and $\widetilde{y}_{i,t-2}$) and are indexes of relative income mobility, a higher τ representing *lower* mobility across income ranks. The various τ's are individual-specific and are allowed to vary with observed and unobserved heterogeneity and with employment status at t, $t-1$ and $t-2$:

$$\begin{aligned} \tau_{i,t,t-1} &= \tau_1\left(e_{it},\mathrm{pub}_{it},e_{i,t-1},\mathrm{pub}_{i,t-1},z_{it}^v,z_i^f,k_i^y\right) \\ \text{and} \quad \tau_{i,t,t-2} &= \tau_2\left(e_{it},\mathrm{pub}_{it},e_{i,t-1},\mathrm{pub}_{i,t-1},e_{i,t-2},\mathrm{pub}_{i,t-2},z_{it}^v,z_i^f,k_i^y\right). \end{aligned} \tag{12}$$

Here again, $\tau_1(\cdot)$ and $\tau_2(\cdot)$ are functions specified in Appendix A.[22]

To sum up, we can now rewrite (7) using the set of assumptions just spelled out: the likelihood of the typical individual's income trajectory $\mathbf{y}_i$ now becomes:

$$\ell_i\left(\mathbf{y}_i\mid\mathbf{e}_i,\mathbf{pub}_i,\mathbf{z}_i^v,z_i^f,k_i^y\right)=\left[\prod_{t=1}^T\frac{1}{\sigma_{it}}\right]\cdot\left[\prod_{t=3}^T\frac{\varphi_3\left(\widetilde{y}_{it},\widetilde{y}_{i,t-1},\widetilde{y}_{i,t-2};\underline{\tau}_{it}^3\right)}{\varphi_2\left(\widetilde{y}_{i,t-1},\widetilde{y}_{i,t-2};\underline{\tau}_{i,t-1}^2\right)}\right]\cdot\varphi_2\left(\widetilde{y}_{i2},\widetilde{y}_{i1};\underline{\tau}_{i2}^2\right). \tag{13}$$

[22]Note in particular that the fact that $\underline{\tau}_{it}^n$ is a correlation matrix imposes a number of constraints on its elements. For instance, because $\tau_{i,t,t-1}$ is a correlation coefficient, we have to restrict its value to be in the interval $[-1;1]$.

4.5 Likelihood maximization

Using the particular structure of our statistical model, we can rewrite the individual contribution to the complete likelihood, $\mathcal{L}_i(\mathbf{x}_i, k_i)$, defined in (1), as follows:[23]

$$\mathcal{L}_i(\mathbf{x}_i, k_i) = \ell_i\left(\mathbf{y}_i \mid \mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^v, z_i^f, k_i^y\right) \cdot \ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i^m\right) \cdot \Pr\left\{k_i^y \mid k_i^m, z_i^f\right\} \cdot \Pr\left\{k_i^m \mid z_i^f\right\}. \quad (14)$$

Parameter estimates are obtained by maximization of the *sample* log-likelihood:

$$\sum_{i=1}^{N} \log\left(\int \mathcal{L}_i(\mathbf{x}_i, k_i)\, dk_i\right), \quad (15)$$

i.e. one has to integrate the individual random effects out of (14). We carry out this maximization iteratively using a slightly altered version of the EM algorithm (Dempster et al., 1977), which we now sketch.

First, in practice, we maximize the sample likelihood in two steps. The structure of (14) indeed makes it very easy to integrate income sequences ($\mathbf{y}_i$) and income classes (k_i^y) out of $\mathcal{L}_i(\mathbf{x}_i, k_i)$. We can thus recover the parameters pertaining to the job mobility process and job mobility classes by considering individual likelihood contributions of the form:

$$\mathcal{L}_i^m\left(\mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^v, z_i^f, k_i^m\right) = \ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i^m\right) \cdot \Pr\left\{k_i^m \mid z_i^f\right\}, \quad (16)$$

and maximizing $\sum_{i=1}^{N} \log\left(\int \mathcal{L}_i^m dk_i^m\right)$. Given the simple structure of our model of job mobility, this maximization can be achieved by a straightforward application of the EM algorithm for finite mixtures.

Once this first step has delivered estimates of the "job mobility" parameters, we fix those at their estimated values and turn back to the maximization of (15) in our second step, now concentrating on the "income process" part of the likelihood. Given the highly nonlinear nature of the "income part" of (14)—see subsection 4.4 —, we don't use direct maximization but rather a sequential

[23] Once again we now omit the term pertaining to the sample distribution of observed individual characteristics z_i^f, $\ell\left(z_i^f\right)$, as it doesn't involve any parameter. (So in effect we are working with the likelihood conditional on z_i^f.)

version of the EM algorithm in the spirit of Arcidiacono and Jones (2003) or Bonhomme and Robin (2004). A fully-fledged description of this second step of our algorithm is provided in Appendix B.

Before we move on to the results, we should raise two remarks about our estimation procedure. First, it is two-step, and therefore less efficient than a single-step approach. However, as is often the case, we traded off efficiency losses for gains in terms of computational tractability. Another advantage of multi-step methods is that they leave no ambiguity as to what part of the data serves for the estimation of a given set of parameters. In our case, for instance, parameters of the job mobility process are estimated for labor market state records alone, while parameters of the income process are estimated to fit income histories. Second, because we use a sequential algorithm in our second step, our estimator differs from standard ML estimator (see Appendix B for more on this). As a consequence, standard procedures such as the delta method cannot be used to obtain standard errors: we will thus have to bootstrap our model.

5 Results

We now present estimation results under the assumption that individuals sort themselves into 3 mobility classes and 2 income classes. Our approach to selecting the number of classes is very pragmatic, in that we take the minimal number of classes allowing the model to fit the data reasonably closely. This turns out to be $K^m = 3$ and $K^y = 2$.[24]

5.1 Labor market states

Worker allocation and mobility between states. We thus have 3 groups of mobility heterogeneity, each containing a non-trivial fraction of the entire population. Selection patterns into labor market states (private employment, public employment and unemployment) differ widely across mobility classes (Table 3). Class 2 (which comprises about 11% of the population) has a very high unemployment rate of 43%, and otherwise tends to be over-represented in the public sector (compared to the aggregate figure—bottom panel of Table 3). Members of class 1 (resp.

[24]Popular model selection criteria based on penalized likelihood (AIC, BIC) tend to suggest we should consider larger numbers of classes. However, increasing the number of classes only entails marginal gains in terms of fit (results with 3 income classes are available upon request), while considerably increasing computing time. Moreover, an additional advantage of limiting the number of classes to a minimum is expositional convenience.

class 3), on the other hand, overwhelmingly select themselves into the public (resp. private) sector. The latter two classes both have very low, roughly equal rates of unemployment. The largest group is class 3, comprising 72% of the sample.

< Tables 3 and 4 about here. >

Table 4 shows the average year-to-year transition probabilities between labor market states, separately for each class. Consistently with the figures shown in Table 3, we see that members of mobility class 1 (resp. 3) have very high rates of persistence in- and transition into- the public (resp. private) sector. Both of these classes have low job loss rates. Note that the job loss rate is worse in the public than in the private sector for class 3, i.e. the "private" class, and vice-versa for the "public" class $k^m = 1$. Members of mobility class 2 are rather mobile between employment sectors (compared to the aggregate figures repeated in the bottom-right panel of Table 4). Moreover, they are at relatively high risk of becoming unemployed and they tend to stay unemployed longer. Class 2 thus appears to gather "low-employability" workers tending to take public sector jobs. Note finally that the job loss rate for this class is lower in the public than in the private sector.

< Figure 2 about here. >

In the descriptive analysis of Section 3, we pointed out the potentially important role of experience in the determination of job loss rates, and public-private differences thereof. The top three panels of Figure 2 plot the experience profiles of job loss rates for the two sectors and for each mobility class. The bottom left panel shows average potential differences across sectors in these profiles (i.e. assuming that everyone is in either sector) while the bottom right panel shows these differences given the observed selection of individuals into sectors.[25] In all cases it appears that job loss rates are lower for public sector employees at all levels of experience. For all mobility classes the job loss rate is hump-shaped with experience in the public sector, whereas it is U-shaped in the private sector. The minimum cross-sector gap in job loss rates always occurs around 20 years of experience. Variations in job loss rates are far larger across mobility classes than across sectors in a given class. (Note that the scale on the vertical axis is different for class 2 than for the other

[25] All these graphs are based on predicted job loss rates. Unreported graphs show that the fit to experience profiles of observed job loss rates is excellent. (Those graphs are available upon request.)

two classes. The bottom two panels are on a common scale.)

Comparison of the bottom two graphs on Figure 2 highlights the presence of mild selection effects. Job loss rates are slightly smaller at all levels of experience in the public sector on the "whole sample with selection" graph (by 2 percentage points on average).

Determinants of unobserved mobility heterogeneity. The top panel of Table 5 reports the composition of our 3 mobility classes in terms of observed individual characteristics. Class 2 (the "unstable" class) is substantially less experienced and, surprisingly, better educated than average. Class 1 (the "public" class) has markedly less low-educated workers, and a high proportion of high-educated workers. They also tend to be marginally more experienced than average. Finally, class 3 (the "private" class) has a relatively high proportion of low-educated men. As they account for 71% of the sample, they otherwise have characteristics that are close to those of the whole sample. These results parallel the raw figures given in Section 3 about the composition of sectors by education levels.

< **Table 5 about here.** >

5.2 Income

Income dispersion and income mobility. We now turn to the analysis of income and income heterogeneity. We have 2 income classes, with total sample weights of 37.9% for class $k^y = 1$ and 62.1% for class $k^y = 2$. The top two panels of Figure 3 show mean income together with the 10th and 90th percentiles of income as a function of experience for each income class and sector. As we did for job loss rates, we also show similar graphs for the whole sample with and without selection (bottom two panels). It is immediately evident from the top two panels of Figure 3 that members of income class 2 earn higher (on average) and less dispersed incomes: looking at unconditional figures, the observed gap is about 16 log points in the means and 8 log points in the standard deviations, both in favor of income class 2. We also observe that this ranking of classes in terms of income means and variances is the same on average in both employment sectors. Also, income class 2 tends to enjoy higher returns to experience in both sectors. Yet income dispersion seems essentially invariant with experience, as the 10th and 90th percentile lines are approximately parallel on all panels.

< Figure 3 about here. >

The public gap is very different between classes: while members of income class 1 clearly benefit from public sector employment, both in terms of mean income and income dispersion, the public premium is low on both accounts for members of income class 1. In figures, the average public premium is of +22 log points for class 1 and −6 log points for class 2. In fact, both classes have roughly equal mean potential wages in the public sector, even though the income variance of income class 1 is still much higher. Looking at the bottom left panel, it appears that the average potential public premium is small at all levels of experience (5 log points on average), the only substantial difference between the two sectors pertaining to income dispersion, which is markedly lower in the public sector. Yet the income premium is positive and constant over all experience levels (at about 14 log points) in the selected sample. Comparison of the bottom two panels of Figure 3 thus confirms the prominent part played by non-random sorting of workers across employment sectors in explaining the apparent public premium. We shall discuss selection patterns at more length in a later Section.

In search of a better denomination, we shall henceforth refer to members of income class 1 (resp. income class 2) as “low income” (resp. “high income”), bearing in mind that both classes earn in fact similar incomes in the public sector.

< Table 6 and Figure 4 about here. >

Turning now to income mobility, Table 6 shows the average income quintile transition probabilities from one year to the next for each separate income class and for the whole sample. The class-specific matrices suggest that “high income” (resp. “low income”) also qualify as “income stayers” (resp. “income movers”). In other words, members of income group 1 tend to have earnings that are both slightly lower on average and less stable over time than members of income group 2. Part of this instability naturally results from the likely higher *job* instability of type $k^y = 1$ workers. As we shall see in the next paragraph, income heterogeneity is indeed correlated with mobility heterogeneity.

A different perspective on income mobility is shown on Figure 4, which shows the autocovariance at one lag of normalized income (the τ_{it}^1 coefficient in the statistical model) as a function of experience and for each sector and class. This Figure confirms the pattern observed above in Table 6, that is, individuals of income class 2 tend to be "income-stayers" while class 1 individuals exhibit more mobile incomes. It appears from the top two panels that income persistence increases somewhat with experience. As for mean incomes, being employed in the public sector has a much larger impact on members of class 1 than on members of class 2 in terms of income persistence. For both groups, public sector employment is associated with more income persistence. Finally, selection does not appear to have much impact on this measure of income persistence.

Determinants of unobserved income heterogeneity. The education- and experience-composition of our two income classes is shown in the middle panel of Table 5. Members of class 1 (the "low income" class) are somewhat older and clearly less educated than average.

Table 5 also reports the joint distribution of our latent heterogeneity classes (in the form of conditionals). Clearly, the two dimensions of heterogeneity are not independent, as being a low income type ($k^y = 1$) is likely associated with being a "low-employability" mobility type ($k^m = 2$). The distribution of income types within "private" and "public" mobility types ($k^m = 1$ and 3) is approximately the same as in the total sample. Hence the correlation between income types and mobility types is far from perfect. All but one of our $K^m \times K^y = 6$ latent classes of unobserved heterogeneity represent a non-negligible share of the sample: the only unlikely combination is for an individual to be a high income type and at the same time a low-employability mobility type (i.e. $k^y = 1$ and $k^m = 2$)—these only account for 2.4% of the total sample.

5.3 Model fit and model specification

Simulations. We end this discussion of our estimation results with a brief fit and specification analysis of the statistical model of Section 4. We thus want to simulate our model in order to compare actual and model-generated data. To this end, we replicate our sample 6 times (i.e. as many times as there are unobserved heterogeneity classes in total) and use our estimated job mobility and income processes to simulate individual labor market trajectories for each individual/class in the sample. We then produce simulated descriptive statistics, weighting each {individual i, class (k^m, k^y)} observation by the probability that individual i belongs to mobility class k^m and income class k^y, given individual i's observed characteristics $\mathbf{x}_i$, $\Pr\{k_i^m = k^m, k_i^y = k^y \mid \mathbf{x}_i\}$.[26]

Worker allocation and mobility between states. We begin with an assessment of our model's capacity to capture the observed patterns of worker allocation into the various labor market states. First looking at "cross-sectional" statistics, it appears that the model is able to fit the numbers presented in Table 3, i.e. the private employment, public employment and unemployment rates, perfectly (up to a fraction of a percentage point in each case). Turning next to labor market transitions, the top panel of Table 7 shows the observed and simulated cross-job-state transition matrices at intervals of one, two and six years (which is the maximum possible lag in a seven-year panel).[27] Here again the fit is excellent. The maximum discrepancy between observed and simulated transition matrices is less than 5%.[28]

< Table 7 about here. >

These, however, are in-sample predictions, which only show that the model does a good job at replicating the particular data that it was estimated to fit. Since we in fact have 12 available waves of BHPS data (covering the period 1991-2002—see Section 3) out of which we only used the last 7 for estimation, we can use the first 5 waves for out-of-sample prediction testing. Many (about 40%) of the 3,791 men present in our 1996-2002 estimation sample are in fact present in all 12 waves of the BHPS. For these people we can construct observed labor market state (or income quintile—see below) transition matrices at intervals of up to 11 years, and then compare these matrices their model-based counterparts. The bottom panel of Table 7 presents this comparison for a lag of 10 years.[29] The main discrepancy between the observed and predicted transition matrices is that we

[26]These probabilities are computed in the E-step of the EM algorithm. See Appendix B.

[27]Unreported matrices are available upon request and convey a similar message.

[28]One should however bear in mind that the number of observations from which the observed transition rates at long lags are computed is rather small. For instance, only 15 people in the sample are unemployed in both years 1996 and 2002.

[29]Because of a change in coding of the job sector variable in the data source from wave 5 onward, the job sector

seem to be over-predicting a bit the persistence of unemployment and employment in the public sector at long lags, the flip side of this being that we under-predict transitions into the private sector. Specifically, we are overstating somewhat the probability that someone observed to be either unemployed or holding a public sector job in 1992 be found in the same state 10 years later. What this seems to indicate is that both unemployment and public sector employment were more persistent over our 1996-2002 estimation period than they were in the first half of the 1990's.

In spite of these discrepancies, the overall performance of our statistical model at predicting labor market state transitions is arguably very satisfactory. First, the maximum absolute prediction error that we make here is in the order of 10 percentage points. Second, we should bear in mind that the number of observations on which *observed* job state transition matrices at long lags are based is very small in many cases: for instance, only 7 individuals in the sample were recorded unemployed in both extreme years 1992 and 2002. Thus, at this level of precision, our comparison exercise is only indicative of the fact that the model-based data still bears a close enough relationship to actual data even at long lags.

Income dispersion and income mobility. Now turning to the income data, we first assess the model's capacity to replicate observed cross-sectional income distributions. Figure 5 plots the observed and predicted log-income densities for the three labor market states separately, and for the sample as a whole. It also shows the income class-specific densities, normalized at the relative size of each class within each particular labor market state.

< **Figure 5 about here.** >

First looking at the whole-sample graph (bottom-right panel), we see that the overall fit is very good, meaning that a mixture of two normal densities is enough to deliver a very good approximation of the sample income density. Moreover, the class-specific densities give a visual

data prior to 1996 are possibly a bit less reliable. Moreover, yet another coding change in 1992 makes the first wave job sector data difficult to compare with similar data from other waves. This is the reason why we only present the out-of-sample prediction test at a distance of 10 (rather the maximum 11) years. We also do not report the predictions at shorter lags (to save on space). Those are available upon request. Based on numerous simulations, we can however safely consider that the prediction error we make at 10 years' distance is an upper bound on prediction errors at shorter lags.

confirmation of the characterization of classes in terms of income means of variances given above. Turning next to state-specific densities, we also find a satisfactory fit for all three states, with the qualification that the model is unable to fully capture the irregularity of the unemployment income distribution. However, this irregularity is probably mainly a consequence of the limited number of observations from which the observed distribution is estimated (736, all years pooled).[30] In any case it appears that the model generates an acceptable smoothing of the observed density of unemployment income.[31]

The model's ability to fit patterns of income *mobility* on top of cross-sectional income dispersion is of primary importance for our purposes. Following the same route as we did for job mobility, we now simulate full individual labor market histories—featuring income *and* job state transitions—and compare the thus obtained income quintile transition matrices with those directly estimated from the sample. We do this at various lags from one to ten years and report the results in Table 8.

< Table 8 about here. >

The maximum absolute discrepancy between observed and predicted quintile transition matrices is less than 10 percentage points at all lags, *including 10 years* which corresponds to an out-of-sample prediction (bottom panel of Table 8). This is a remarkable outcome. First, recall that we only have six latent classes in total, of which only two of income heterogeneity. This minimal specification of unobserved heterogeneity combined with a second-order Markov income process seems to deliver a good replication of income mobility even at long lags. Second, it turns out that the magnitude of the prediction error does not tend to increase as one looks at longer time intervals.[32] The good fit that we obtain thus constitutes strong support for our model specification.

[30] Moreover, somewhat surprisingly, the "unemployment" graph reveals that class 2 earns *less* than class 1 in unemployment (-15 log points). Yet, as we also see from this graph, members of income class 2 are unemployed much less often than members of income class 1.

[31] The "observed" density plotted on Figure 5 is produced by a Gaussian kernel estimator with a bandwidth determined using the rule of thumb $h = 0.9 \times \text{std}\,(\text{log income})\,/n^{1/5}$, which already does a certain amount of smoothing to the corresponding histogram.

[32] The error is in the order of 5 to 10 percentage points at all time intervals considered, including time intervals relating to out-of-sample predictions. The numbers are available upon request. Note also that, contrary to labor market state transition matrices, all elements of the observed income mobility matrices are based on reasonably large numbers of observations.

A discussion of possible alternative specifications. Income ranks are highly persistent, and our combined assumption of second-order Markov and time-invariant unobserved income heterogeneity goes a long way into capturing this persistence. Yet looking at the nature of prediction errors, we detect a systematic tendency of the model to over-predict income mobility somewhat. This aspect of the model could potentially be improved on by specifying the income process as a Markov process of order higher than 2 or by adding to the number of income heterogeneity classes.[33] Yet the computational cost of our estimation procedure increases very quickly as one expands the model in either dimension, often for relatively small marginal gains in terms of fit. One is therefore bound to settle for a compromise.

Now surely, there is a long list of (tractable) alternative specifications. The simplest possible "first-pass" specification is just to do away with any unobserved heterogeneity. Unfortunately, while such a "homogeneous" model is very quick and easy to estimate in practice, it generally grossly over-predicts both job and income mobility at lags beyond one or two years. Besides this first possibility, probably the most popular specification in the literature is that of a first-order Markov (e.g. AR(1) in levels) income process. In the context of the present paper, first-order Markov income processes are a bit simpler analytically and quicker to estimate given a fixed number of heterogeneity classes. Yet they don't fit the data as well: estimating our model with two classes of income heterogeneity under the assumption that $\tau_{i,t,t-2} = 0$ (see equation 12) yields income mobility in-sample (i.e. at distances of up to 6 years) prediction errors that are on average 50% larger than those reported in Table 8. Increasing the number of income classes to three, one is back with in-sample prediction errors very similar to those obtained from our baseline second-order Markov assumption. Yet the out-of-sample prediction error at a 10 year distance is on average twice as large in the {Markov(1), $K^y = 3$} than in the {Markov(2), $K^y = 2$} case. Moreover, the latter specification is less costly in computational terms than the former.

Overall, there appears to be a trade-off between the amount of "built-in" persistence resulting from the assumption made on the order of the Markov process, and the extra bit of income

[33]With three classes, the maximum prediction error is further reduced to 6%.

autocorrelation brought about by time-invariant unobserved heterogeneity. Our choice of second-order Markov with a small number of classes was mainly guided by a concern for parsimony and by the particular configuration of the data. Our estimation sample has a relatively small cross-section dimension, balanced by a long longitudinal dimension (far longer than it takes to identify a second-order Markov process). A specification with less unobserved heterogeneity and more built-in persistence is probably better suited and more computationally economical in this particular configuration.[34] Moreover, as we just mentioned, the second-order Markov assumption seems important to ensure good long-term predictions of income mobility—a valuable property for the main purpose of this paper, which is to use the model to construct lifetime values of individual labor market trajectories.

6 The public pay gap: incomes and lifetime values

In this Section we take a more systematic look at selection into the public sector and at public-private differences from the twofold perspective of income flows and lifetime values. To this end, we first construct lifetime values, then analyze differences across sectors. Finally, we examine these differences under counter-factual assumptions on sector choice.

6.1 Construction of lifetime values

Definitions. Using our estimated coefficients for income distributions, income and job mobility, we are now in a position to carry out simulations of employment and income trajectories for the individuals in our sample until retirement age. Assuming that, upon retiring (which we assume to happen at a level of experience denoted as T_R), a given individual enjoys a residual value of V_R, then the lifetime value as of experience level t of an individual's simulated future income trajectory $\mathbf{y}_{s\geq t}$ is written as:

$$V_t\left(\mathbf{y}_{s\geq t}\right) = \sum_{s=t}^{T_R} \beta^{s-t} \cdot U\left(y_s\right) + \beta^{T_R - t} \cdot V_R, \tag{17}$$

[34] The N/T ratio of the particular data set used surely plays a crucial role in the choice of a specification. To take another example, in their application of similar estimation techniques to the French labor force survey *Enquête emploi*, Bonhomme and Robin (2004) are constrained to use a first-order Markov specification as the T-dimension of their sample is only three years (compensated by a very large N-dimension).

where $\beta \in (0,1)$ is a discount factor and $U(y_t)$ is the utility flow that the individual derives from (log) income y_t. At each level of experience t, current income y_t is conditional on the individual's characteristics and labor market state, which includes experience, past income levels, current and past job sectors, and so on... as specified in our statistical model (see Section 4 and Appendix A).

We can naturally consider different functional forms and/or parameterizations for (17) to allow the individuals to exhibit various degrees of time preference or risk aversion. In the sequel we shall mainly consider the following two baseline specifications:

$$U(y_t) = y_t \tag{Logarithmic}$$

$$U(y_t) = \frac{e^{(1-1/\eta)y_t}}{1-1/\eta}. \tag{CRRA}$$

The logarithmic specification simply assimilates lifetime values $V_t(\mathbf{y}_{s\geq t})$ to present discounted sum of future log income flows. Given our modeling of the income process as log-additive in a permanent/deterministic component and a transitory shock (see equation 22 in Appendix A), simply taking the present discounted sum of future log income flows is representative of a situation where individuals are indifferent to or can insure themselves against those transitory shocks.[35]

Even though the "perfect financial market" situation underlying the logarithmic specification is an arguably useful benchmark, it nonetheless has the inconvenience of assuming away any impact of cross-sector differentials in income risk. To get round this problem, we shall also consider the more general CRRA specification shown above, where η denotes the intertemporal elasticity of substitution.[36] While admittedly, this latter specialization is theoretically a bit awkward in the absence of a fully-fledged model of consumption and income smoothing over the life cycle, it still constitutes a simple way of taking into consideration the impact of income uncertainty. In an attempt to generate clear-cut results, we will assume a fairly large income smoothing motive by giving the elasticity of intertemporal substitution a value on the low side of conventional calibrations, $\eta = 0.5$.

[35] Again given our log-additive modeling of income, considering a (perhaps more natural) specification of $U(\cdot)$ as *linear in income*, i.e. $U(y_s) = e^{y_s}$, would introduce a preference for a higher variance of the transitory component of income.

[36] Note that what we term the "CRRA" specification is also a CARA specification in log incomes. Hence the logarithmic specification confounds itself with the special case $\eta = +\infty$.

We set the discount factor to $\beta = 0.9$ *per annum*. Finally, we define the value of retirement as $V_R = \frac{1}{1-\beta} \cdot U\left(y_{T_R - 1} + \ln \rho\right)$ in both the logarithmic and the CRRA case, where ρ designates the replacement ratio. That is, we assume that after retirement, individuals receive a constant flow of income equal to ρ times their last income in activity and discount this flow over an infinite number of years.[37] We calibrated the value of ρ at 0.40, which is roughly the average ratio of the first income in retirement to the last income in activity for men in the BHPS over the 1996-2002 period. Sensitivity of our main results to changes in all these numbers will be analyzed.

Two caveats. Before we proceed, we should raise the following two caveats. First, in order to compute these lifetime values, we assume that the economic environment is stationary. That is, agents anticipate getting older and experiencing wage and job mobility given their current wage and job status, but they do not anticipate the model parameters to change over the rest of their working life. For this assumption to be credible, we need our sample period to be fairly representative of an "average" state of the business cycle. As we showed in Section 3 (Figure 1), both the share of public sector employment and the unemployment rate over our sample period are stable enough.

The second issue is that it is not straightforward to compare measures of the lifetime public premium involving different specifications of the instantaneous utility function (logarithmic *vs.* CRRA)—or, for that matter, to compare any measure of a lifetime public premium, which involves a utility function, with a public premium in terms of income. In order to produce a measure that has at least some comparability across specifications of $U(\cdot)$, we use the following transformations:

$$V(\cdot) \mapsto \widetilde{V}(\cdot) = (1-\beta) \cdot V(\cdot) \qquad \text{(Logarithmic)}$$

$$V(\cdot) \mapsto \widetilde{V}(\cdot) = \frac{\eta}{\eta - 1} \log\left[\left(\left(1 - \frac{1}{\eta}\right) \cdot V(\cdot)\right)^{\eta/(\eta-1)}\right], \qquad \text{(CRRA)}$$

and define the (relative) lifetime public premium (measured in log points) as the difference $\widetilde{V}_{\text{public}} - \widetilde{V}_{\text{private}}$. The above transformations are such that scaling all future income flows up or down by $n\%$ will result in $\widetilde{V}(\cdot)$ being scaled up or down by $n\%$ as well. Surely this particular normalization is

[37] The implicit assumption of an infinite life expectancy after retirement overestimates the value of retirement by about 10%.

arbitrary and it will condition the size of the lifetime public premia found under either specification. The reason we adopted it was to ensure scale consistency between lifetime values and permanent incomes.

6.2 The role of income mobility

The aggregate picture. We begin the analysis by running a series of counterfactual simulations in which we rule out any mobility between labor market states. Indulging in a slight misuse of language, we shall refer to these simulations as the "job for life" case. Individual trajectories are simulated under the assumption that the probabilities of moving between sectors or into unemployment are zero. The only sources of differences in lifetime values are hence cross-sectional income differences (also plotted) and differences in income mobility across sectors. Ruling out income mobility allows us to obtain a neat picture of the impact of income mobility on the public lifetime premium. We shall explore the role of job mobility in the next subsection.

Figure 6 displays the public premium in terms of income, lifetime values with a logarithmic utility and lifetime values with a CRRA utility by percentiles in their respective distributions, under the "job for life" assumption. The "whole sample, with selection" graph relates to predicted "raw" differences, i.e. it plots the difference between quantiles of income flows and lifetime values among individuals effectively observed to hold public jobs in the initial period, and corresponding quantiles of income flows and lifetime values among workers observed to hold private jobs in the initial period. The "whole sample" graph relates to predicted differences in incomes and in lifetime values across sectors for all individuals in the sample, i.e. it compares income and lifetime values that each individual could *potentially* earn in either sector.

< **Figure 6 about here.** >

First looking at the "whole sample" graph, we see that the public premium in predicted income decreases as ones goes up the income distribution, from a positive premium of about 20 log points in the first two deciles to a negative premium of 2 to 8 log points in the top two deciles. This is a reflection of the well-documented phenomenon of relative income compression in the public sector.

The picture is quite different in terms of lifetime values. Under the logarithmic utility specification, the lifetime public premium is much smaller than the income public premium in the first two deciles of the distribution (around 8 log points) and larger in the upper part (1 to 4 log points in the top two deciles). Hence there is a lot less compression in the public sector relative to the private sector in terms of lifetime values than in terms of income. Moreover, the lifetime public gap remains positive throughout lifetime value quantiles, whereas the income gap becomes negative in the top quartile of income. We interpret these phenomena as the result of income mobility offsetting differences in cross-sectional incomes: thinking of log incomes as the sum of a permanent random individual effect and a transitory random shock (both sector-specific), our results suggest that most of the observed relative income compression in the public sector is due to a lower variance of the *transitory* component of income, which is averaged out when taking lifetime values.

The CRRA case differs from the logarithmic utility case in two respects. Most noticeably, the lifetime public premium is everywhere higher in the CRRA case. This of course results from the fact that income risk matters to agents endowed with a less-than-infinite elasticity of intertemporal substitution, and we saw at various points that income risk—as measured by cross-sectional income variance or by the persistence of income ranks—was markedly smaller in the public than in the private sector. Second, we observe the gap between the lifetime premia obtained with the two specifications decreases as one goes up the distribution (from about 5 log points in the bottom two deciles to 1 log point in the top two). A tentative explanation for this decrease is that, as the conditional log-income variance is roughly constant over the income distribution, the relative log-income risk is greater at low quantiles, while we have assumed constant *absolute* log-income risk aversion.

Looking at the "whole sample, with selection" graph, we observe similar patterns of lifetime public premium as above. Both premia are larger here, averaging about 16 log points for the logarithmic utility and 19 for the CRRA utility. Comparing this with the "whole sample" graph suggests that at least half of the public-private gaps in lifetime values observed here are a product of selection effects, whereby individuals with higher potential incomes or higher employability are selected into the public sector. This echoes the composition of sectors in terms of education levels: we saw in Section 3 that employees of the public sector have substantially higher academic qualifications than average. Selection on unobservables, on the other hand, follows a more complex pattern involving unobserved determinants of income (k^y) as well as mobility (k^m). Conditional on being employed, both income classes are split between the public and private sector in roughly similar proportions (18% public for income class 1, 17% for income class 2). Yet, as one sees on Figure 3, members of income class $k^y = 1$ enjoy a much higher income premium in the public sector than members of income class $k^y = 2$. One would thus expect to see members of income class 1 go to the public sector in far *larger* proportions. However, as we pointed out in Section 5 (Table 5), being a member of income class 1 is often associated with being a member of mobility class $k^m = 2$—the "low-employability" class—and thus having high job loss rates and low job access rates altogether. This seems to point to a certain amount of job queuing in the public sector, something we shall return to in a later sub-section.

Conditioning on individual heterogeneity. Figure 7 again shows public sector premia, this time splitting the sample into separate classes of unobserved heterogeneity.

$<$ **Figure 7 about here.** $>$

Individuals of income class 1, who gain from public sector employment in terms of income would also potentially enjoy a substantial public premium in lifetime values when jobs are assumed to last for ever: this premium is positive throughout the distribution and increases from about 18 (30) log points in the bottom deciles to 30 (36) log points in the top quintile for the logarithmic case (CRRA case). For individuals in income class 2, all public-private differences are generally negative and much smaller in absolute value than for individuals of income class 1. The fact that public-private differences in income persistence (see Figure 4) and income variance (see Figure 3) are also much smaller for income class 2 may explain both that premia in income and in both types of lifetime values are confounded for individuals of this class and that these lifetime premia are small.

The bottom row of Figure 7 shows public-private differences in simulated incomes and lifetime values for jobs for life in either sector by mobility class. Differences in lifetime values are somewhat larger in the "public" class ($k^m = 3$) than in the "private" class ($k^m = 1$). They are more substantial for the “low-employability” class, averaging near 17 log points even when we use a logarithmic utility function and rising to above 24 log points when we assume a CRRA utility.

Two features of this graph attract attention. First, the public premium, be it in terms of income or in either type of lifetime values, is still positive (albeit small) for individuals who work predominantly in the private sector (mobility class 3). Second, public premia in either income or lifetime values are largest for individuals from the “low-employability” class, whose unemployment rate is 44%. This is suggestive of job queuing, whereby individuals from this class find it difficult to find and retain public sector employment. This applies to a lower extent to members of mobility class 3.

6.3 The role of job mobility

We now restore the possibility of job mobility by adopting an alternative definition of sector-specific lifetime values: we simulate income trajectories imposing that the individual be employed in a given sector *in the first period*, but allowed to move between sectors or into unemployment thereafter according to his predicted transition probabilities. Public premia based on this definition are depicted on Figure 8 for the whole sample, and on Figure 9 for separate classes of unobserved heterogeneity.

< Figures 8 and 9 about here. >

The striking point about Figures 8 and 9 is that in all panels, public-private differences in lifetime values disappear almost entirely when we allow workers to switch between sectors, regardless of the particular specialization retained for the instantaneous utility function. The public premium is less than 2 log points in absolute value at all quantiles in the “whole sample” graph (with a small positive premium of 3 to 5 log points remaining in the bottom decile). Comparing this with the “whole sample, with selection” graph where the public premium in lifetime values is positive everywhere again points to the importance of selection effects.

It thus seems that job mobility has a strong “equalizing effect” on lifetime values. An obvious driving force behind this effect is that workers may not stay long in the initial sector that we impose on them. Individuals who are observed to work, say in the private sector, are predominantly members of mobility class $k^m = 3$ (see Table 3), who have a strong propensity not only to stay in, but also to move into the private sector (see the transition matrices in Table 4). Hence, when we counter-factually place these individuals in the public sector in the initial period, it does not take them long to go back into the private sector. And of course, the fewer periods they initially spend in the public sector, the less this initial spell weighs in their lifetime value.

Workers indeed move back very quickly to their "natural" sector: among individuals of the "private" class ($k^m = 3$) whom we simulate to start out as public sector employees, about 89% have left the public sector within one year and another 10% do so within 2 years. The vast majority of these leavers (98%) go to private sector employment "directly", while the remaining 2% are observed at least once in unemployment. Individuals of the "public" class ($k^m = 1$) whom we simulate to start out in the private sector remain there for longer: "only" 44% leave within one year and another 24% within 2 years, predominantly (96%) toward a public sector job, and occasionally (4%) toward unemployment. Individuals of the "low-employability" class ($k^m = 2$) have a less clear-cut pattern of selection into sectors, as Table 4 above showed: if we place them in the public sector in the initial period, 55% of them leave within one year and another 23% within 2 years, 98% of these because they have found employment in the private sector. However, if they start off in the private sector, only about 2.5% leave every year over the first few years, and fewer thereafter, 93% of these to take up employment in the public sector while the other 7% become unemployed.

So even though it seems somewhat more difficult to move from the private to the public sector than in the opposite direction, spell durations in "counterfactual" job sectors are short enough to contend that there is sufficient mobility in the UK labor market to allow at least "high-employability" workers to select themselves effectively into their "natural" sector.

6.4 Sensitivity analysis [Incomplete.]

Time preference and the planning horizon. The annual discount rate β obviously (or, rather equivalently, the length of the planning horizon) partly condition the results described in this Section, as a smaller value of β clearly makes lifetime values stick more closely to current income flows. Yet, so long as one stays within a set of values that are not unreasonably far from standard calibrations, the impact of β is not as impressive as one might expect: while the public-private gap in lifetime values is essentially zero at all quantiles under our baseline calibration of $\beta = 0.9$ *per annum*, it rises to a range of 1 to 4 log points at the first three quartiles for $\beta = 0.6$ in the linear $U(\cdot)$ case (2 to 8 log points in the $\eta = 0.5$ case), and to 3 to 7 (resp. 5 to 10) log points in the linear (resp. $\eta = 0.5$) case as β is further reduced to a value as low as 0.3. These numbers should be compared with income gaps varying between 3 and 10 log points (see e.g. Figure 8).

The value of retirement. The set of parameters used involved in lifetime values also includes V_R, the value of retirement. The results summarized in Figures 6 to 9 are robust to essentially any reasonable assumption about V_R. `[More to come.]`

7 Conclusion

`[Forthcoming.]`

References

[1] Arcidiacono, P., and J. B. Jones (2003), "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm", Econometrica, 71(3), 933-46.

[2] Bender, K. (1998), "The Central Government-Private Sector Wage Differential", Journal of Economic Surveys, 12(2), 177-220.

[3] Bonhomme, S., and Robin, J.M. (2004), "Modeling Individual Earnings Trajectories Using Copulas with an Application to the Study of Earnings Inequality: France, 1990-2002" , mimeo CREST-INSEE.

[4] Borjas, G. (2002), "The Wage Structure and the Sorting of Workers into the Public Sector", NBER working paper 9313.

[5] Bowlus, A., and Robin, J.M. (2004), "Twenty Years of Rising Inequality in US Lifetime Labor Income Values", Review of Economic Studies, 71(3), 709-742.

[6] Dempster, A. P., N. M. Laird, and P. B. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm", Journal of the Royal Statistical Society, 39, 1-38.

[7] Disney, R., and Gosling, A. (1998), "Does it Pay to Work in the Public Sector?", Fiscal Studies, 19, 347-374.

[8] Disney, R., and Gosling, A. (2003), "A New Method for Estimating Public Sector Pay Premia: Evidence from Britain in the 1990s", CEPR Discussion Paper 3787.

[9] Dustmann, C., and van Soest, A. (1998), "Public and Private Sector Wages of Male Workers in Germany", European Economic Review, 42, 1417-1441.

[10] Eckstein, Z., and K.I. Wolpin (1999), "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities and Abilities", Econometrica, 67, 1295-339.

[11] Flinn, C. (2002), "Labor Market Structure and Inequality: A Comparison of Italy and the US", Review of Economic Studies, 69(3), 611-645.

[12] Gottschalk, P. (1997), "Inequality, Income Growth, and Mobility: The Basic Facts", Journal of Economic Perspectives, 11(2) (Spring 1997), 21-40.

[13] Gottschalk, P. and R. Moffit (1993), "The Growth of Earnings Instability in the US Labor Market", Brookings Papers on Economic Activity, vol. 1994, No. 2, 217-72.

[14] Heckman, J. J., and Singer, B. (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", Econometrica, 52, 271-320.

[15] Katz, L., and Krueger, A. (1991), "Changes in the Structure of Wages in the Public and Private Sectors", Research in Labor Economics, 12, 137-172.

[16] Keane, M., and K. Wolpin (1997), "The Career Decisions of Young Men", Journal of Political Economy, 105, 473-522.

[17] Moulton, B. (1990), "A Reexamination of the Federal-Private Wage Differential in the United States", Journal of Labor Economics, 8, 2, 270-293.

[18] Mueller, R. (1998), "Public-Private Sector Wage Differentials in Canada: Evidence from Quantile Regressions", Economics Letters, 60, 229-235.

[19] Nickell, S., and Quintini, C. (2002), "The Consequences of the Decline in Public Sector Pay in Britain: A Little Bit of Evidence", Economic Journal, 112, F107-F118.

[20] Shorrocks, A. (1976), "Income Mobility and the Markov Assumption", Economic Journal, 86, 566-578.

[21] Stewart, M. (2004), "The Inter-related Dynamics of Unemployment and Low-Wage Employment", mimeo.

[22] van Ophem, H. (1993), "A Modified Switching Regression Model for Earnings Differentials Between the Public and Private Sectors in the Netherlands", Review of Economics and Statistics, 75, 215-224.

APPENDIX

A Model specification

In this Appendix we describe our functional form assumptions in full detail. For clarity, we first recall some notation and the basic structure of our statistical model. The sample is a set of N workers indexed $i = 1, \dots N$, each of whom we follow over T_i consecutive years. A typical observation for an individual $i = 1, \dots, N$ is a vector $\mathbf{x}_i = \left(\mathbf{y}_i, \mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^v, z_i^f\right)$, to which we append a pair $k_i = (k_i^m, k_i^y)$ of unobserved individual class indexes.

As explained in Section 4 of the main text, individual i's contribution to the complete likelihood has three components (see equation (1)), pertaining respectively to unobserved heterogeneity, labor market status history and income history. We now give the full specification of each of these components separately. In all instances, our specific choice of covariates was guided by the descriptive analysis of Section 3 as well as by a concern for parsimony and numerical tractability.

Unobserved heterogeneity. As indicated by equation (2), attachment of individual i to a given latent class $k_i = (k_i^m, k_i^y)$ is modeled as the product of two terms: $\ell_i\left(k_i \mid z_i^f\right) = \Pr\left\{k_i^y \mid k_i^m, z_i^f\right\} \cdot \Pr\left\{k_i^m \mid z_i^f\right\}$, which we both specify as multinomial logits:

$$\Pr\left\{k_i^m = k^m \mid z_i^f\right\} = \frac{\exp\left(z_i^{f\prime} \cdot \kappa_{k^m}^m\right)}{\sum_{k=1}^{K^m} \exp\left(z_i^{f\prime} \cdot \kappa_k^m\right)} \text{ and } \Pr\left\{k_i^y \mid k_i^m, z_i^f\right\} = \frac{\exp\left[\begin{pmatrix} z_i^f \\ k_i^m \end{pmatrix}' \cdot \kappa_{k^y}^y\right]}{\sum_{k=1}^{K^y} \exp\left[\begin{pmatrix} z_i^f \\ k_i^m \end{pmatrix}' \cdot \kappa_k^y\right]}, \tag{18}$$

where κ_1^m and κ_1^y are both normalized at zero.

Labor market states. From equations (3), (4) and (5), we know that individual labor market histories contribute to the complete likelihood as:

$$\begin{aligned} \ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i^m\right) &= \Pr\left\{e_{i1} \mid z_i^f, k_i^m\right\} \times \left[\Pr\left\{\text{pub}_{i1} \mid z_i^f, k_i^m\right\}\right]^{e_{i1}} \\ &\times \prod_{t=2}^{T_i} \left(\Pr\left\{e_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} \cdot \left[\Pr\left\{\text{pub}_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\}\right]^{e_{it}}\right). \end{aligned} \tag{19}$$

All components are specified as logits:

$$\begin{aligned} \Pr\left\{e_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} &= \left(1 + \exp\left[-\begin{pmatrix} \begin{pmatrix} e_{i,t-1} \\ \text{pub}_{i,t-1} \end{pmatrix} * z_{i,t-1}^v \\ z_i^f \\ k_i^m \end{pmatrix}' \cdot \psi\right]\right)^{-1} \\ \Pr\left\{\text{pub}_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} &= \left(1 + \exp\left[-\begin{pmatrix} \begin{pmatrix} e_{i,t-1} \\ \text{pub}_{i,t-1} \end{pmatrix} * z_{i,t-1}^v \\ z_i^f \\ k_i^m \end{pmatrix}' \cdot \chi\right]\right)^{-1}, \end{aligned} \tag{20}$$

where the notation $x * y$ stands for all main effects and interactions of variables x and y. We finally use *mutatis mutandis* similar specifications for the initial job state probabilities:

$$\begin{aligned}
\Pr\left\{e_{i1} \mid z_i^f, k_i^m\right\} &= \left(1+\exp\left[-\begin{pmatrix} z_i^f \\ k_i^m \end{pmatrix}' \cdot \psi_0\right]\right)^{-1} \\
\Pr\left\{\mathrm{pub}_{i1} \mid z_i^f, k_i^m\right\} &= \left(1+\exp\left[-\begin{pmatrix} z_i^f \\ k_i^m \end{pmatrix}' \cdot \chi_0\right]\right)^{-1}. \qquad (21)
\end{aligned}$$

Income. Given what is explained about our modeling of individual income paths in the main text, all we need to spell out in this Appendix is the set of functions $\{\mu(\cdot), \sigma(\cdot), \tau_1(\cdot), \tau_2(\cdot)\}$ introduced in equations (8) and (12). We start with $\mu(\cdot)$:

$$\mu\left(e_{it}, \mathrm{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right) = z_i^{f\prime} \cdot \mu_0 + \left[\begin{pmatrix} e_{it} \\ \mathrm{pub}_{it} \end{pmatrix} * k_i^y * z_{it}^v\right]' \cdot \mu_1. \qquad (22)$$

As one sees from this latter equation, we allow the returns on experience to differ across job sectors and income classes.

As for log income variance, we posit:

$$\sigma\left(e_{it}, \mathrm{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right) = \exp\left(\left[\begin{pmatrix} e_{it} \\ \mathrm{pub}_{it} \end{pmatrix} * z_{it}^v\right]' \cdot \sigma_0 + k_i^{y\prime} \cdot \sigma_1\right). \qquad (23)$$

Here we force the standard deviations of log income to be positive by specifying it as an exponential. Also, the functional form used for $\sigma(\cdot)$ is somewhat more constrained than the one we use for income means. In particular, time-invariant observed individual characteristics, z_i^f, are not retained among the arguments of $\sigma(\cdot)$—that is, we assume that z_i^f only impacts income variance through its link to unobserved income classes, k_i^y.

We finally turn to the dynamics of income ranks, which are governed by the functions $\tau_1(\cdot)$ and $\tau_2(\cdot)$. The correlation between normalized income and normalized income lagged once, $\tau_1(\cdot)$, is specified as:

$$\begin{aligned}
&\tau_1\left(e_{it}, \mathrm{pub}_{it}, e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{it}^v, k_i^y\right) = -1 \\
&\quad + 2 \cdot \left(1+\exp\left[[k_i^y * z_{it}^v]' \cdot \zeta_0 + \left[\begin{pmatrix} e_{it} \\ \mathrm{pub}_{it} \end{pmatrix} * k_i^y\right]' \cdot \zeta_1 + \left[\begin{pmatrix} e_{i,t-1} \\ \mathrm{pub}_{i,t-1} \end{pmatrix} * k_i^y\right]' \cdot \zeta_2\right]\right)^{-1} \qquad (24)
\end{aligned}$$

This specification calls for some comments. First, the transformation $-1 + 2 \cdot (1+\exp[\cdot])^{-1}$ which we apply to a linear index in the explanatory variables is there to constrain $\tau_1(\cdot)$, a correlation coefficient, to lie within $[-1, +1]$. Second, we again subsume the impact of z_i^f into that of k_i^y, which in turn conditions all coefficients in the latter equation. Third, we do not allow for all possible interaction between current and past labor market states, $(e_{it}, \mathrm{pub}_{it})$ and $\left(e_{it-1}, \mathrm{pub}_{it-1}\right)$.[38]

[38] The type of constraint that we impose is that the marginal effect on $\tau_1(\cdot)$ of being observed, say, in the public sector at date $t-1$ is independent of the particular sector in which one is observed at date t. Here again, we chose to impose this restriction after a large number of trials with more elaborate specifications, the impact of which was essentially to increase computation time for very little gain in terms of fit.

The correlation between normalized income and normalized income lagged twice, $\tau_2(\cdot)$, is slightly more involved. Let us first define the notation shortcut $\tau_{it}^1 = \tau_1\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, z_{it}^v, k_i^y\right)$ to designate the date-t, one-lag autocorrelation of normalized income. Then:

$$\begin{aligned}\tau_2\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z_{it}^v, k_i^y\right) &= \tau_{it}^1 \cdot \tau_{i,t-1}^1 \\ &+ \left(\sqrt{\left(1-\left(\tau_{it}^1\right)^2\right)\cdot\left(1-\left(\tau_{i,t-1}^1\right)^2\right)}\right)\cdot \widetilde{\tau}_2\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z_{it}^v, k_i^y\right),\end{aligned} \tag{25}$$

with

$$\begin{aligned}&\widetilde{\tau}_2\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z_{it}^v, k_i^y\right) = -1 \\ &+2\cdot\left(1+\exp\left[\left[k_i^y * z_{it}^v\right]'\cdot\xi_0 + \left[\begin{pmatrix} e_{it} \\ \text{pub}_{it}\end{pmatrix} * k_i^y\right]'\cdot\xi_1 + \left[\begin{pmatrix} e_{i,t-1} \\ \text{pub}_{i,t-1}\end{pmatrix} * k_i^y\right]'\cdot\xi_2 + \left[\begin{pmatrix} e_{i,t-2} \\ \text{pub}_{i,t-2}\end{pmatrix} * k_i^y\right]'\cdot\xi_3\right]\right)^{-1}\end{aligned} \tag{26}$$

This last pair of equations requires some clarification. We have to constrain $\tau_2(\cdot)$ in such a way that, given τ_{it}^1 and $\tau_{i,t-1}^1$, the matrix:

$$\begin{pmatrix} 1 & \tau_{it}^1 & \tau_{it}^2 \\ \tau_{it}^1 & 1 & \tau_{i,t-1}^1 \\ \tau_{it}^2 & \tau_{i,t-1}^1 & 1 \end{pmatrix}$$

is a consistent covariance matrix (where we have just introduced the additional notation shortcut $\tau_{it}^2 = \tau_2\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z_{it}^v, k_i^y\right)$). This involves in particular its determinant Δ_{it} being positive. The latter is defined by $\Delta_{it} = 1-\left(\tau_{it}^1\right)^2-\left(\tau_{i,t-1}^1\right)^2-\left(\tau_{it}^2\right)^2+2\tau_{it}^1\tau_{i,t-1}^1\tau_{it}^2$. Solving for τ_{it}^2, we get:

$$\tau_{it}^2 = \tau_{it}^1\cdot\tau_{i,t-1}^1 \pm \sqrt{\left(1-\left(\tau_{it}^1\right)^2\right)\cdot\left(1-\left(\tau_{i,t-1}^1\right)^2\right)-\Delta_{it}}. \tag{27}$$

Because Δ_{it} is positive, τ_{it}^2 has to stay within the interval

$$\left[\tau_{it}^1\cdot\tau_{i,t-1}^1 - \sqrt{\left(1-\left(\tau_{it}^1\right)^2\right)\cdot\left(1-\left(\tau_{i,t-1}^1\right)^2\right)}, \tau_{it}^1\cdot\tau_{i,t-1}^1 + \sqrt{\left(1-\left(\tau_{it}^1\right)^2\right)\cdot\left(1-\left(\tau_{i,t-1}^1\right)^2\right)}\right].$$

This is achieved by the parameterization in equations (25) and (26), the latter imposing that $\widetilde{\tau}_2(\cdot)$ be within $[-1,+1]$.

B The likelihood maximization algorithm

From Appendix A and Section 4 in the main text, the list of parameters to be estimated can be divided into 2 subsets, $\Theta^m = \left\{\left(\kappa_k^m\right)_{k=1}^{K^m}, \psi_0, \chi_0; \psi; \chi\right\}$ and $\Theta^y = \left\{\left(\kappa_k^y\right)_{k=1}^{K^y}, \mu(\cdot), \sigma(\cdot), \tau_1(\cdot), \widetilde{\tau}_2(\cdot)\right\}$, where by $\{\mu(\cdot), \sigma(\cdot), \tau_1(\cdot), \widetilde{\tau}_2(\cdot)\}$ we summarize all parameters of the corresponding functions—see equations (22), (23), (24) and (26). The first subset, Θ^m, gathers all parameters from the job mobility process—i.e. involved in equations (18), (20) and (21). The subset Θ^y contains all remaining parameters, which all pertain to the income process—see equations (22) through (26).

As explained in subsection 4.5, individual contributions to the complete likelihood can be decomposed as $\mathcal{L}_i\left(\mathbf{x}_i, k_i; \Theta^m, \Theta^y\right) = \mathcal{L}_i^m\left(\mathbf{x}_i, k_i^m; \Theta^m\right)\cdot\mathcal{L}_i^y\left(\mathbf{x}_i, k_i^y; \Theta^y\right)$, where $\mathcal{L}_i^m\left(\mathbf{x}_i, k_i^m; \Theta^m\right)$ is spelled out in equation (16) (the only addition brought in this Appendix is to make the dependence on the various parameters explicit).

B.1 Step 1: Job mobility

The first step of our (2-step) estimation procedure is to get estimates of Θ^m by maximizing the sample log-likelihood of individual job state histories $\sum_{i=1}^{N} \log\left(\int \mathcal{L}_i^m dk_i^m\right)$. This maximization is carried out using a standard EM-algorithm, which consists of iterating the following two steps:

E-step. For an initial value Θ_n^m of Θ^m, for each mobility class index $k^m = 1, \ldots, K^m$, and for each individual i in the sample, compute the posterior probability that i belongs to mobility class k^m given $\mathbf{x}_i$ and Θ_n^m:

$$\Pr\left\{k_i^m = k^m \mid \mathbf{x}_i; \Theta_n^m\right\} = \frac{\mathcal{L}_i^m\left(\mathbf{x}_i, k^m; \Theta_n^m\right)}{\sum_{k=1}^{K^m} \mathcal{L}_i^m\left(\mathbf{x}_i, k; \Theta_n^m\right)}. \tag{28}$$

M-step. Update Θ_n^m into Θ_{n+1}^m by maximizing the following augmented sample log-likelihood, weighted by (28):

$$\Theta_{n+1}^m = \arg\max_{\Theta^m} \sum_{i=1}^{N} \sum_{k=1}^{K^m} \Pr\left\{k_i^m = k \mid \mathbf{x}_i; \Theta_n^m\right\} \cdot \log\left[\mathcal{L}_i^m\left(\mathbf{x}_i, k; \Theta^m\right)\right]. \tag{29}$$

This latter maximization is easily carried out running separate weighted logit regressions for ψ and χ (see equations (20) and (21), and a weighted multinomial logit for the class weight parameters κ_k^m, using (28) as weights in each case.

This algorithm converges to the MLE of Θ^m (Dempster et al., 1977). In practice we stop iterating when the maximum relative change between Θ_n^m and Θ_{n+1}^m falls below 10^{-3}, and thus obtain our estimate $\widehat{\Theta}^m$.

B.2 Step 2: Income

In our second step we obtain estimates of the subset of "income" parameters by maximization of the sample likelihood, $\mathcal{L}_i\left(\mathbf{x}_i, k_i; \widehat{\Theta}^m, \Theta^y\right)$, only fixing the "job mobility" parameters Θ^m at their estimated value $\widehat{\Theta}^m$ from step 1. This latter maximization is carried out in the same spirit as that of step 1—i.e. using an iterative, data-augmenting algorithm —, only with a sequential M-step:

E-step. For an initial value Θ_n^y of Θ^y, for each class index $k = (k^m, k^y)$, $k^m = 1, \ldots, K^m$, $k^y = 1, \ldots, K^y$, and for each individual i in the sample, compute the posterior probability that i belongs to mobility class k^m and income class k^y given $\mathbf{x}_i$, Θ_n^y and $\widehat{\Theta}^m$:

$$\Pr\left\{k_i^m = k^m, k_i^y = k^y \mid \mathbf{x}_i; \widehat{\Theta}^m, \Theta_n^y\right\} = \frac{\mathcal{L}_i\left(\mathbf{x}_i, k^m; \widehat{\Theta}^m, \Theta_n^y\right)}{\sum_{\ell^m=1}^{K^m} \sum_{\ell^y=1}^{K^y} \mathcal{L}_i\left(\mathbf{x}_i, \ell^m, \ell^y; \widehat{\Theta}^m, \Theta_n^y\right)}. \tag{30}$$

M-step. This is where our algorithm differs a bit from the standard EM. Specifically, we proceed as follows:

1. Update income mean parameters $\mu(\cdot)$ using weighted OLS regressions of y_{it} on $\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right)$, using (30) as weights. Denote the updated function $\mu(\cdot)$ as $\widehat{\mu}_{n+1}(\cdot)$.

2. Take the (log) squared residuals from the latter regression and regress those on $\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right)$, again using weighted OLS, to update income variance parameters $\sigma(\cdot)$. Denote the update as $\widehat{\sigma}_{n+1}(\cdot)$.

3. Form normalized individual log income $\widetilde{y}_{it}^{(n+1)} = \frac{y_{it}-\widehat{\mu}_{n+1}\left(e_{it},\text{pub}_{it},z_{it}^v,z_i^f,k_i^y\right)}{\widehat{\sigma}_{n+1}\left(e_{it},\text{pub}_{it},z_{it}^v,z_i^f,k_i^y\right)}$. Update $\tau_1(\cdot)$ as:

$$\widehat{\tau}_{1,n+1}\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, z_{it}^v, k_i^y\right) = \text{cov}\left(\widetilde{y}_{it}^{(n+1)}, \widetilde{y}_{i,t-1}^{(n+1)}\right), \tag{31}$$

given that $\left(\widetilde{y}_{it}^{(n+1)}, \widetilde{y}_{i,t-1}^{(n+1)}\right)$ is distributed bivariate normal with unit variances. We do this by weighted maximum likelihood, using (30) as weights. Then similarly update $\widetilde{\tau}_2(\cdot)$ knowing that $\tau_2(\cdot) = \text{cov}\left(\widetilde{y}_{it}^{(n+1)}, \widetilde{y}_{i,t-2}^{(n+1)}\right)$ is given by formula (25), and that $\left(\widetilde{y}_{it}^{(n+1)}, \widetilde{y}_{i,t-2}^{(n+1)}\right)$ is again distributed bivariate normal with unit variances. Note that $\tau_1(\cdot)$ is involved in (25), and that we replace it by $\widehat{\tau}_{1,n+1}(\cdot)$ for the update of $\widetilde{\tau}_2(\cdot)$.

4. Finally update the set of income class assignment parameters, $(\kappa_k^y)_{k=1}^{K^y}$ by running a weighted multinomial logit regression of class indexes on $\left(z_i^f, k_i^m\right)$, again using (30) as weights.

We iterate the E- and M-steps above until the maximum relative change between two consecutive updates of Θ^y becomes less than 10^{-3}. Note that, in contrast to the standard EM algorithm, we update the parameters sequentially (as opposed to simultaneously) within each iteration. The reason for doing so is that all parameters of Θ^y are involved in the copula part of the likelihood function in a very non linear way, which renders simultaneous updating of all parameters numerically cumbersome. By using a sequential approach, we throw away some information (e.g. we only use income cross-sections—and not income dynamics—to update the mean and variance parameters) and thus lose some efficiency. However, we enjoy sizable gains in terms of computational tractability and computation time.[39]

[39] Bonhomme and Robin (2004, Appendix C) show that this algorithm converges to a consistent estimator of the parameters, which nevertheless differs from the ML estimator and is not efficient. See also Arcidiacono and Jones (2003).

Dependent variable: log monthly income					
	Specification number...				
	1	2	3	4	5
Constant	7.277 (.004)	7.122 (.015)	7.141 (.017)	6.311 (.037)	6.297 (.038)
Public	0.138 (.010)	0.049 (.010)	−0.005 (.042)	0.083 (.018)	0.212 (.059)
Unemployment	−1.583 (.023)	−1.495 (.020)	−1.918 (.096)	−1.239 (.022)	−1.302 (.084)
Medium ed.	—	−0.314 (.010)	−0.329 (.023)	—	—
Low ed.	—	−0.442 (.009)	−0.471 (.010)	—	—
Experience (years/10)	—	0.423 (.015)	0.430 (.017)	0.688 (.030)	0.702 (.031)
Experience^2 (years^2/100)	—	−0.076 (.003)	−0.076 (.003)	−0.091 (.006)	−0.094 (.006)
Public×Medium ed.	—	—	0.048 (.026)	—	—
Public×Low ed.	—	—	0.119 (.023)	—	—
Unempl.×Medium ed.	—	—	0.236 (.080)	—	—
Unempl.×Low ed.	—	—	0.378 (.067)	—	—
Public×Experience	—	—	−0.009 (.038)	—	−0.130 (.056)
Public×Experience^2	—	—	0.002 (.008)	—	0.026 (.012)
Unempl.×Experience	—	—	0.101 (.079)	—	0.067 (0.086)
Unempl.×Experience^2	—	—	−0.016 (.016)	—	−0.013 (.019)
Fixed effects	No	No	No	Yes	Yes

Notes: All years pooled.
Specifications 1, 2, and 3: OLS. Specifications 4 and 5: within estimator.
Reference categories are "Private sector job" and "High education".
Standard Errors in parentheses.

Table 1: Public-private differences: mean income

	Private sector					Public sector				
	income quintile at t					income quintile at t				
income quintile at $t-1$	65.8	20.5	7.6	3.7	2.4	74.9	20.1	2.9	1.3	0.8
	19.9	50.6	19.4	7.2	2.9	14.5	55.5	23.0	5.4	1.8
	6.4	21.0	47.9	20.7	4.1	3.0	19.7	52.1	20.9	4.3
	4.1	5.9	21.1	53.1	15.8	1.5	3.9	19.7	57.8	17.2
	3.0	2.5	3.8	15.5	75.3	0.3	1.1	2.7	16.5	79.4
income quintile at $t-6$	48.7	26.7	13.4	6.4	4.8	69.2	20.5	2.6	5.1	2.6
	22.9	32.7	22.9	13.1	8.4	16.1	45.2	29.0	6.5	3.2
	13.2	26.0	32.6	19.0	9.3	4.9	19.5	46.3	24.4	4.9
	7.3	11.1	23.5	39.3	18.8	0.0	7.5	17.5	45.0	30.0
	7.4	6.2	8.6	21.0	56.8	0.0	0.0	9.1	24.2	66.7

Note: Sector-specific income quintiles.

Table 2: Public-private differences: mobility of income ranks

Sectoral composition of classes				
	% of sample	% private	% public	% unempl.
$k^m = 1$	17.28	15.9	82.0	2.1
$k^m = 2$	10.92	41.9	14.5	43.6
$k^m = 3$	71.80	97.5	0.6	1.9

Class composition of sectors				
	% of sample	% $k^m = 1$	% $k^m = 2$	% $k^m = 3$
Private	77.3	3.6	5.9	90.5
Public	16.2	87.5	9.8	2.7
Unemployment	6.5	5.6	73.7	20.7

Table 3: Mobility Classes

state at t	$k^m=1$			$k^m=2$		
	state at $t+1$			state at $t+1$		
↓	private	public	unempl.	private	public	unemp.
private	55.3	43.0	1.8	86.0	2.5	11.5
public	4.3	95.1	0.6	55.0	40.1	4.9
unempl.	27.4	46.7	25.9	38.0	5.7	56.3
	$k^m=3$			Whole sample		
private	98.6	0.5	0.9	96.9	1.5	1.5
public	88.4	11.2	0.4	8.2	90.1	0.8
unempl.	80.6	1.2	18.2	41.0	4.1	54.9

Table 4: Labor market state transition probabilitites

Mobility classes							
	% high ed.	% med. ed.	% low ed.	Mean exp.	% $k^y=1$	% $k^y=2$	
$k^m=1$	35.8	22.6	41.6	20.8	38.6	61.4	
$k^m=2$	30.1	15.7	54.2	11.8	78.3	21.7	
$k^m=3$	22.2	21.9	55.9	20.5	33.4	66.6	
Income classes							
	% high ed.	% med. ed.	% low ed.	Mean exp.	% $k^m=1$	% $k^m=2$	% $k^m=3$
$k^y=1$	19.6	20.2	60.2	20.2	17.0	21.8	61.2
$k^y=2$	29.2	22.1	48.8	19.2	17.4	3.9	78.7
Whole sample							
	% high ed.	% med. ed.	% low ed.	Mean exp.			
	25.4	21.3	53.2	19.6			

Table 5: Composition and joint distribution of unobserved heterogeneity classes

	$k^y = 1$					$k^y = 2$				
	income quintile at $t+1$					income quintile at $t+1$				
income quintile at t	64.2	18.4	8.7	4.9	3.9	74.4	21.0	3.2	0.9	0.5
	26.8	36.8	18.2	12.0	6.2	15.0	61.5	19.7	3.3	0.6
	16.5	22.1	32.6	19.2	9.7	1.8	20.3	55.9	20.5	1.6
	13.2	11.8	21.8	35.4	17.8	0.3	2.8	20.7	61.4	14.9
	9.6	7.3	10.0	18.9	54.2	0.2	0.4	1.3	14.8	83.3
	Whole sample									
	income quintile at $t+1$									
income quintile at t	68.2	19.4	6.5	3.3	2.5					
	19.1	52.8	19.2	6.3	2.6					
	6.2	20.8	48.9	20.1	4.0					
	3.8	5.2	21.0	54.4	15.7					
	2.7	2.3	3.7	15.9	75.4					

Note: Income quintiles from the unconditional sample distribution.

Table 6: Income mobility

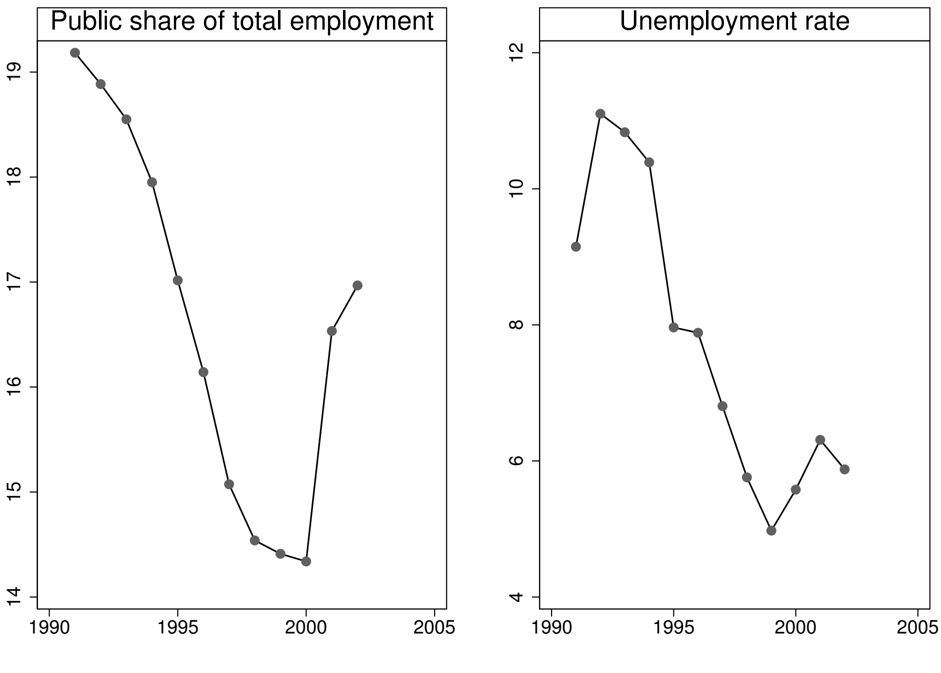

Figure 1: Sector shares and unemployment rate, males 20-55, 1991-2002

	Predicted			Observed				
state at $t-1$	state at t			state at t				
$\downarrow$	private	public	unempl.	private	public	unempl.		
private	96.7	1.5	1.7	96.9	1.5	1.5		
public	8.0	91.1	0.7	8.2	90.1	0.8	Euclidian dist.:	1.0
unempl.	40.4	3.9	55.5	41.0	4.1	54.9	Maximum dist.:	0.7
state at $t-2$								
$\downarrow$								
private	95.2	2.2	2.5	96.7	1.7	1.4		
public	11.8	86.8	1.2	11.0	88.3	0.6	Euclidian dist.:	3.1
unempl.	55.0	4.6	40.2	53.1	5.4	40.7	Maximum dist.:	1.5
$\vdots$								
state at $t-6$								
$\downarrow$								
private	96.3	3.3	2.9	95.3	2.9	1.7		
public	18.4	79.1	2.3	20.7	78.8	0.3	Euclidian dist.:	4.1
Unempl.	69.6	6.3	23.9	69.2	7.6	23.0	Maximum dist.:	2.3

Out of sample prediction

	Predicted			Observed				
state at $t-10$	state at t			state at t				
$\downarrow$	private	public	unempl.	private	public	unempl.		
private	93.8	3.9	2.1	94.1	4.0	1.7		
public	18.8	79.9	1.2	29.7	69.4	0.8	Euclidian dist.:	19.2
unempl.	69.3	10.9	19.6	75.3	14.4	10.1	Maximum dist.:	10.9

Table 7: Fit to job mobility data

	Predicted					Observed					
	income quintile at t					income quintile at t					
income quintile at $t-1$	62.9	21.6	7.9	4.7	2.6	68.2	19.4	6.5	3.3	2.5	Eucl. dist.: 18.4
	21.8	46.5	22.3	6.2	3.1	19.1	52.8	19.2	6.3	2.6	
	8.3	22.6	41.9	22.0	5.0	6.2	20.8	48.9	20.1	4.0	
	4.5	6.2	22.4	45.9	20.8	3.8	5.2	21.0	54.4	15.7	Max. dist.: 8.4
	2.4	2.9	5.2	21.0	68.2	2.7	2.3	3.7	15.9	75.4	
income quintile at $t-2$	57.2	22.9	9.7	5.6	4.4	63.5	20.2	8.9	4.1	3.1	Eucl. dist.: 13.0
	23.8	43.3	21.9	7.5	3.2	21.7	47.1	20.3	8.1	2.5	
	8.7	23.4	39.8	22.0	5.8	8.6	23.3	42.3	20.3	5.3	
	6.3	6.8	22.6	43.7	20.4	4.8	7.2	22.4	48.4	17.0	Max. dist.: 6.3
	3.8	3.3	5.8	21.0	65.9	3.3	2.6	5.4	16.9	71.6	
⋮											
income quintile at $t-6$	43.3	24.2	13.4	11.1	7.7	52.3	22.6	13.0	6.3	5.5	Eucl. dist.: 20.1
	25.3	32.8	22.2	13.9	5.6	28.2	34.2	19.9	9.9	7.6	
	13.6	23.8	29.0	22.0	11.4	11.8	27.5	30.7	20.4	9.4	
	9.7	12.3	23.7	28.8	25.3	7.1	12.8	25.4	37.4	17.0	Max. dist.: 9.0
	8.0	6.5	11.5	24.0	49.7	6.6	5.7	7.6	22.5	57.4	

Out of sample prediction

	Predicted					Observed					
	income quintile at t					income quintile at t					
income quintile at $t-10$	44.0	22.5	16.7	7.4	9.2	39.9	27.1	17.1	10.5	5.2	Eucl. dist.: 18.6
	24.8	31.4	26.4	11.0	6.1	26.1	31.8	22.5	11.5	7.9	
	16.1	23.0	24.2	25.7	10.7	13.8	26.3	31.2	19.6	8.9	
	8.1	14.9	19.3	31.5	25.9	6.1	10.9	19.6	38.4	24.8	Max. dist.: 7.7
	6.8	7.8	13.2	24.0	47.9	7.8	5.8	9.7	20.9	55.6	

Note: Income quintiles from the unconditional sample distribution.

Table 8: Fit to income mobility data

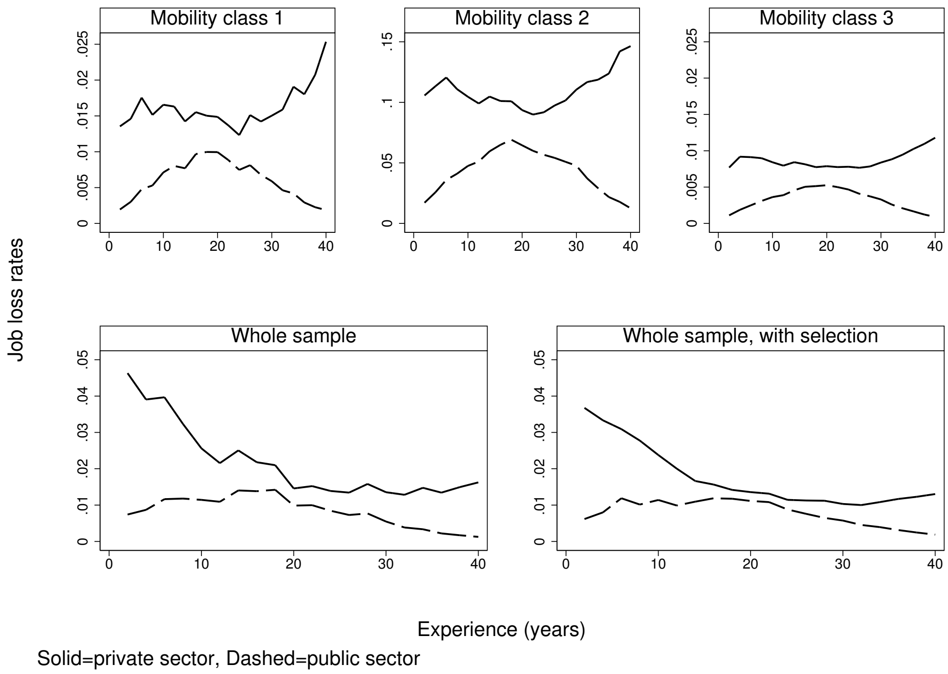

Figure 2: Job loss rates

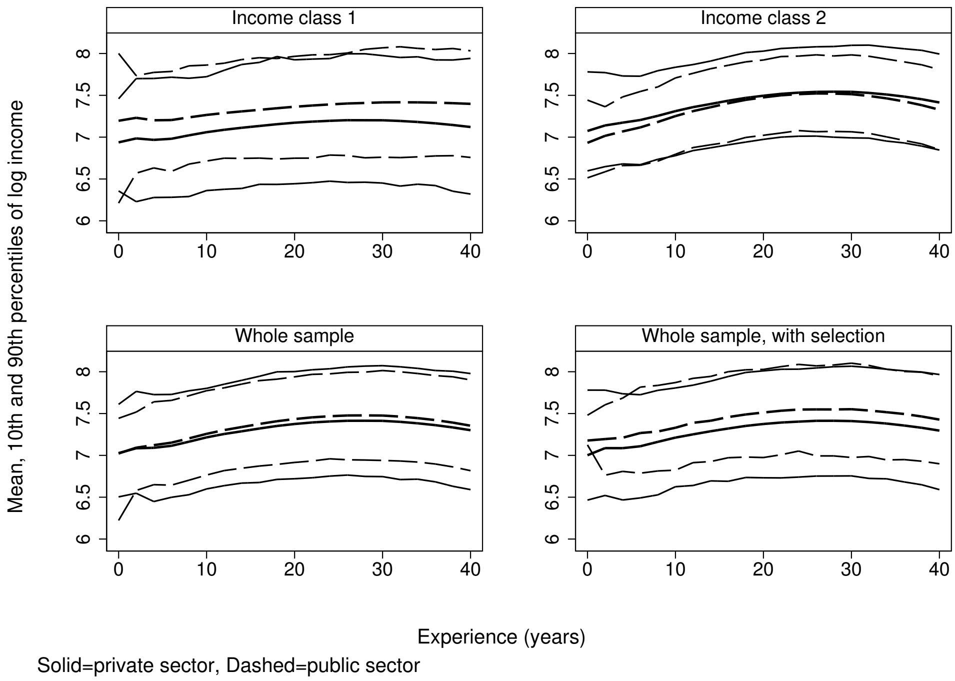

Figure 3: Income-experience profiles

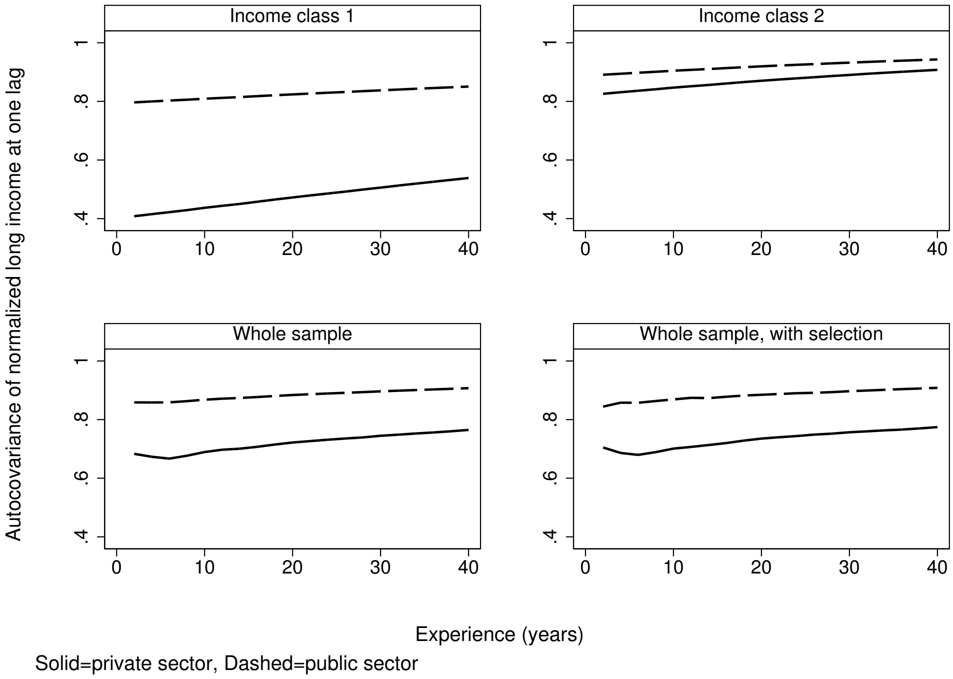

Figure 4: Autocovariance of normalized income

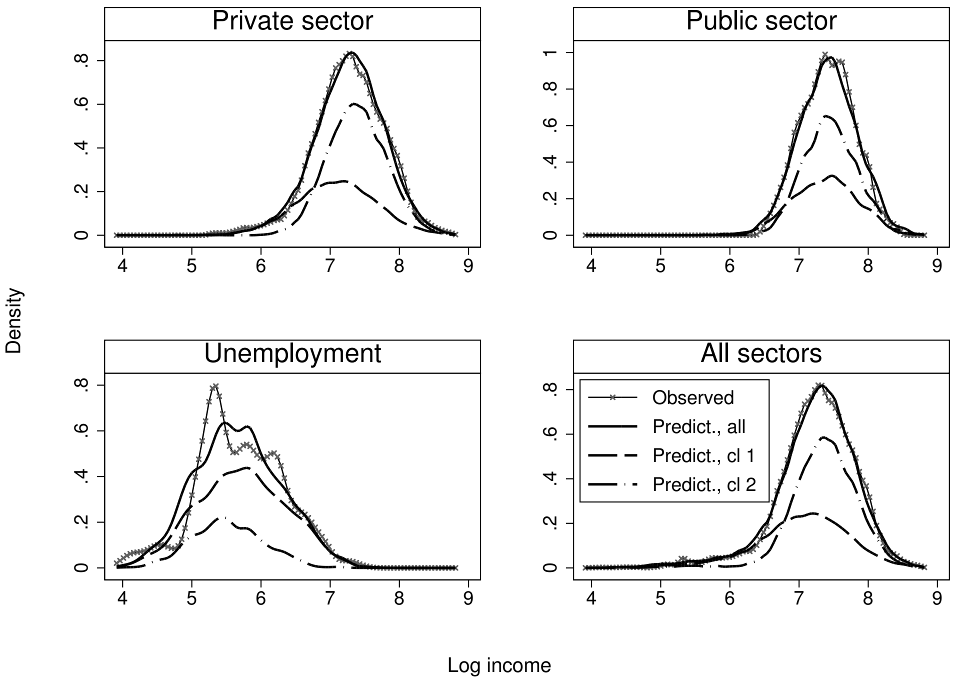

Figure 5: Income densities

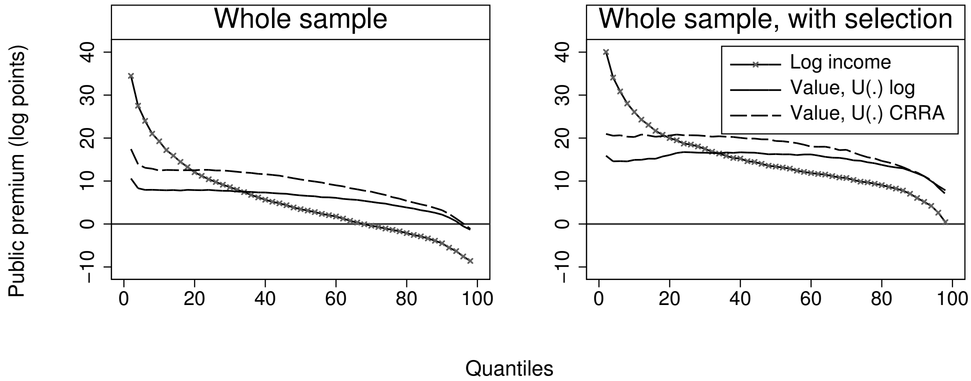

Figure 6: The public gap: jobs for life

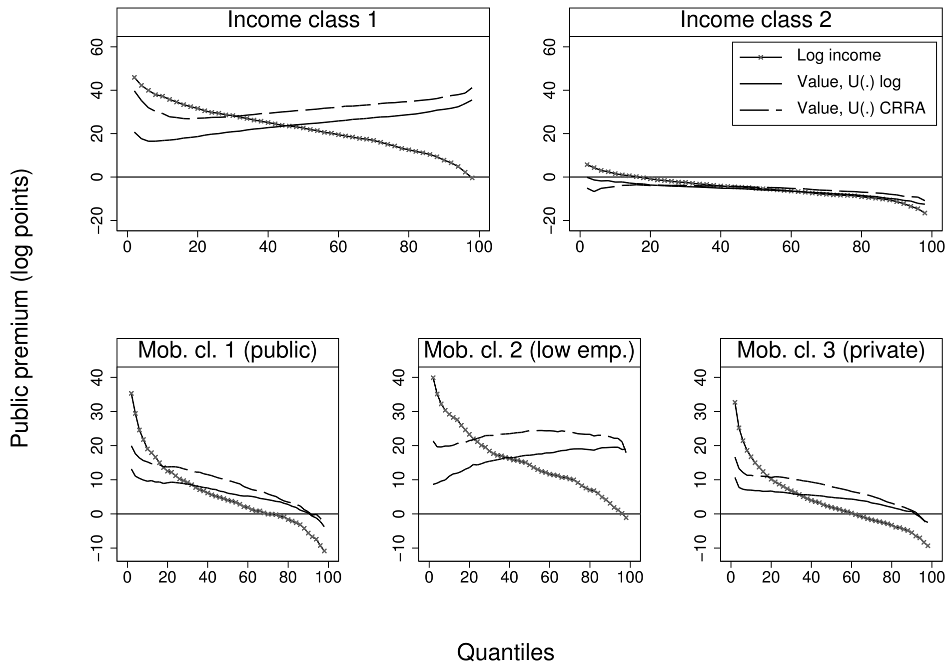

Figure 7: The public gap: jobs for life

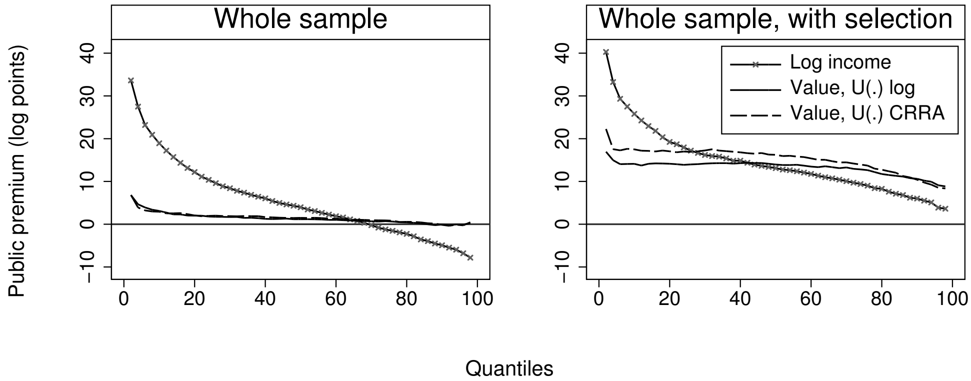

Figure 8: The public gap: income flows and lifetime values

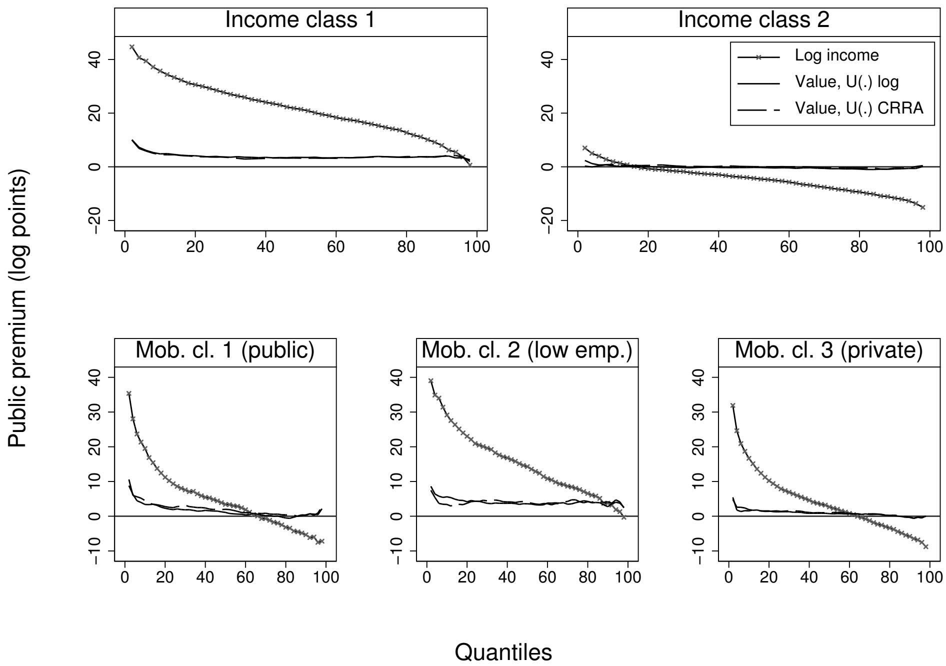

Figure 9: The public gap: income flows and lifetime values