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

March 2005

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 8 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 for public sector jobs among "low-employability" workers.

Keywords: Income Dynamics, Job Mobility, Public-Private Inequality, Selection Effects.
JEL codes: J45, J31, J62.

*We thank Sonia Bhalotra, Richard Blundell, Simon Burgess, Guy Laroque, Jean-Marc Robin, Jon Temple, David Winter and seminar participants at Bristol, CREST and UCL for useful feedback on this project.

[†]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, UK. Tel: +44 117 928 8400. E-mail `helene.turon-lacarrieu@bris.ac.uk`. Internet `www.efm.bris.ac.uk/economics/staff/turon/`.

1 Introduction

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, and 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.[2] 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

[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).

[2] Based on job satisfaction questions from multi-country household survey data (from the ECHP), Clark and Postel-Vinay (2004) show that public sector jobs are perceived as more secure and more insulated from labor market fluctuations than private sector jobs in all of the 12 countries surveyed.

comparative analysis of these values.

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 (we predict income mobility over periods of 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 there is weak evidence that returns to experience are 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 8 percent higher in the public sector than in the private sector when defined as the present discounted sum of future monthly 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. In other words, we find income inequality to be greater yet less persistent in the private then in the public sector, to the point that "long-run" inequality as measured by the cross-sectional variance of present discounted sums of log income flows is similar in both sectors. 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 for public sector jobs among "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

This paper is directly related to two different strands of literature on public-private pay differences, and income mobility and lifetime inequality.

As stressed above, most studies on the public pay gap focus on cross-section differences in wages and on the impact of selection on these differences. For example, Dustmann and van Soest (1998) find that the endogeneity of the choice between public and private sector matters crucially in the estimation of wage differentials. With German data, they estimate wages to be lower in the public sector for all age and education groups, and the gap to decrease with both age and education. They also find that the potential difference in wages is larger for employees of the private sector than those of the public sector, suggesting that workers select themselves to some extent into the sector offering them a comparative advantage. For the UK, although raw cross-section differences show a large positive public premium, Disney and Gosling (2003) find that, controlling for selection, the public sector premium becomes insignificant for men and remains positive for women. More educated men seem to select themselves into the public sector on some negative characteristics as cross-section estimates of the public pay gap is negative and becomes positive once selection is controlled for. In those studies, identification of the public wage premium in the presence of nonrandom selection is achieved with either functional form assumptions (as in Belman and Heywood (1989) for the US or van Ophem (1993) for the Netherlands) or with some instrumental variables procedure appealing to variables such as family background (Dustmann and van Soest, 1998) or variations in public sector status arising from privatization (Disney and Gosling, 2003) as instruments for selection into the public sector.

Another key feature of this literature is the evidence of pay compression in the public sector. Using quantile regressions, Disney and Gosling (1998) for the UK, Mueller (1998) for Canada and Poterba and Rueben (1994) for the US find that the 90-10 percentile ratio and the variance of the wage distribution are lower in the public sector. In these three studies, the public pay premium is estimated to be substantially higher at the 10th percentile than at the 90th. This applies to all gender and education groups. The public premium is found to be positive for all groups at the 10th percentile, while it is negative at the 90th percentile for all except women with low education in the UK and in the US. We argue in this paper, though, that much of this difference in income variance across sectors relates to transitory components of the income process, suggesting that compression of public relative to private sector incomes mainly reflects lower income mobility rather than lower permanent income inequality in the public sector.

This paper is also related to the vast literature on empirical models of income dynamics and more particularly on applications of these latter to the study of lifetime income inequality. This literature has branched out into three broad types of approaches. A majority of contributions use variedly flexible reduced form models of either absolute or relative earnings mobility to decompose the income process into a permanent and a transitory component, interpreting inter-individual differences in the former as measures of lifetime inequality (e.g. Gottschalk and Moffit, 1993; Gottschalk, 1997; Bunchinsky and Hunt, 1999; Bonhomme and Robin, 2004). Our paper—to which Bonhomme and Robin (2004) is most closely related—fits into this first category. A second set of papers take a more structural approach derived from job search theory to analyze inequality in lifetime values and its changes over time (see Bowlus and Robin, 2004 for the US or Flinn, 2002 for a US-Italy comparison). Whereas all these studies measure inequality using data on individual earnings, a third approach consists of looking at consumption inequality, which, under the Permanent Income/Life Cycle Hypothesis conveys information on long-run income inequality (e.g. Blundell and Preston, 1997; Blundell, Pistaferri and Preston, 2004; Jappelli and Pistaferri, 2004). None of these, however, examine lifetime differences between job sectors.

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.[3] 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.[4] 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.[5] 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.[6]

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. A substantial amount of selection is going on, as the education shares of the public (resp. 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

[3]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 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).

[4]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.

[5]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.

[6]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.

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.[7] 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).[8,9]

[7] Throughout the paper, we will be looking at monthly total income. While the part-time shares are equal across sectors in our sample of males (3.6% vs 3.8%), there are slight systematic differences in hours worked between the public and private sector: the median number of weekly hours worked is 37 in the public sector and 39 in the private sector, the standard deviation is larger in the private sector by about 20 minutes. By focusing on monthly incomes, we thus slightly understate any positive public premium in hourly wages, as we do differences in cross-sectional wage variances. The use of monthly income raises a second issue, which is that individuals who have experienced a job/employment transition within the month preceding the interview will report an income that relates in part to their previous employment/sector status. It is, however, difficult to overcome this imprecision by looking at data on the elapsed duration of the current spell as this variable is missing for a substantial fraction of respondents.

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

[9] 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,

Allowing for 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 slightly 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. 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.[10] 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: 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 one-lag autocorrelation of normalized log income on a quadratic in experience, education, and interactions of current and past sector of employment.[11] 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):[12]

	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 to the public sector while movements in the opposite direction are slightly more frequent (8.2%). The raw job loss rates are

[10] These figures relate to individuals who are employed in the same sector in year t and in year $t+1$ (resp. in all years between t and $t+6$). These measures of income mobility hence abstract from wage changes caused by job transitions (from one sector to the other or to unemployment).

[11] 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 $\widehat{\sigma}^2_{it}$. We then constructed income disturbances as $\left(y_{it} - \widehat{y}_{it}\right) / \sqrt{\widehat{\sigma}^2_{it}}$.

[12] 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.

1.5% for individuals initially in the private sector against 0.8% for individuals initially 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.[13]

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.[14] Overall, unemployment persistence thus seems high but is comparable to results found by Stewart (2004) with similar data.

3.4 Summary

The above descriptive analysis thus highlights the following facts:

- 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 average job loss rate in the public sector is just a little over half the average job loss rate of the private sector.

[13]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.

[14]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.

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 e.g. Gottschalk and Moffitt (1993), Gottschalk (1997) and Buchinsky and Hunt (1999) (among many others), comparisons of cross-sections are not very informative in the presence of income mobility. Here we contend that this 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 in order to 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 $i = 1, \ldots N$, 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 $i = 1, \ldots, N$ can be represented as a vector[15] $\mathbf{x}_i = \left(\mathbf{y}_i, \mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^v, z_i^f\right)$, where:

- $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT_i})$ 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 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.

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

- 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:

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

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.[16] This distribution is observed and is independent of any parameter. We shall therefore omit it from now on. 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, $\sum_{i=1}^N \log\left(\int \mathcal{L}_i(\mathbf{x}_i, k_i)\, dk_i\right)$. Before we proceed, however, we have to give a precise account of our modeling assumptions for the various components of (1).

[16] Recall that $\mathbf{z}_i^v$ is deterministic conditional on z_i^f and therefore does not appear in the likelihood function.

4.2 Heterogeneity

We begin with individual heterogeneity, i.e. the third component of $\mathcal{L}_i(\mathbf{x}_i, k_i)$ in (1). As we already mentioned, we allow for both unobserved (k_i) and observed (z_i^f) heterogeneity. We 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[17]. 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 and K^y income classes.[18] 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.

[17]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.

[18]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. See Bonhomme and Robin (2004) for an application more closely related to the present paper.

4.3 Mobility between labor market states

We now turn to the second component of $\mathcal{L}_i(\mathbf{x}_i, k_i)$ in (1). 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:[19]

$$\begin{aligned}
&\Pr\left\{e_{it}, \text{pub}_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z^v_{i,t-1}, z^f_i, k^m_i\right\} \\
&\quad = \Pr\left\{e_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z^v_{i,t-1}, z^f_i, k^m_i\right\} \times \left[\Pr\left\{\text{pub}_{it} \mid e_{i,t-1}, \text{pub}_{i,t-1}, z^v_{i,t-1}, z^f_i, k^m_i\right\}\right]^{e_{it}}. \quad (3)
\end{aligned}$$

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

$$\Pr\left\{e_{i1}, \text{pub}_{i1} \mid z^f_i, k^m_i\right\} = \Pr\left\{e_{i1} \mid z^f_i, k^m_i\right\} \times \left[\Pr\left\{\text{pub}_{i1} \mid z^f_i, k^m_i\right\}\right]^{e_{i1}}. \quad (4)$$

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:

$$\begin{aligned}
\ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}^v_i, z^f_i, k^m_i\right) &= \Pr\left\{e_{i1}, \text{pub}_{i1} \mid z^f_i, k^m_i\right\} \\
&\quad \times \prod_{t=1}^{T_i-1} \Pr\left\{e_{i,t+1}, \text{pub}_{i,t+1} \mid e_{it}, \text{pub}_{it}, z^v_{it}, z^f_i, k^m_i\right\}, \quad (5)
\end{aligned}$$

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

[19] Note the assumption implicit in (3) that only the date-$(t-1)$ component of $\mathbf{z}^v_i$—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.

4.4 Income process

Modeling assumptions. 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.[20]

Temporarily omitting any conditioning variable or individual index, the likelihood of a given income trajectory over T periods can be written as:

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

Because we assume incomes to follow second-order Markov processes, the likelihood function only involves products of bi- or tri-variate densities. Now bringing back conditioning variables, 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:

$$\begin{aligned} y_{it} \mid e_{it}, \mathrm{pub}_{it}, z_{it}^{v}, z_i^{f}, k_i^{y} \quad &\sim \quad \mathcal{N}(\mu_{it}, \sigma_{it}^2) \\ \text{with } \mu_{it} = \mu\left(e_{it}, \mathrm{pub}_{it}, z_{it}^{v}, z_i^{f}, k_i^{y}\right) \quad &\text{and} \quad \sigma_{it} = \sigma\left(e_{it}, \mathrm{pub}_{it}, z_{it}^{v}, z_i^{f}, k_i^{y}\right), \end{aligned} \tag{7}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are given functions (see Appendix A for a fully detailed presentation of all our specification assumptions).

We next introduce the normalized log-income as $\widetilde{y}_{it} = \frac{y_{it} - \mu_{it}}{\sigma_{it}}$. The triple $(\widetilde{y}_{it}, \widetilde{y}_{i,t-1}, \widetilde{y}_{i,t-2})$ and the pair $(\widetilde{y}_{it}, \widetilde{y}_{i,t-1})$ are Gaussian vectors with covariance matrices $\underline{\tau}_{it}^{(3)}$ and $\underline{\tau}_{it}^{(2)}$ respectively, which

[20] 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.

we expand as:

$$\underline{\tau}_{it}^{(3)} = \begin{pmatrix} 1 & \tau_{i,t,t-1} & \tau_{i,t,t-2} \\ \tau_{i,t,t-1} & 1 & \tau_{i,t-1,t-2} \\ \tau_{i,t,t-2} & \tau_{i,t-1,t-2} & 1 \end{pmatrix} \qquad \text{and} \qquad \underline{\tau}_{it}^{(2)} = \begin{pmatrix} 1 & \tau_{i,t,t-1} \\ \tau_{i,t,t-1} & 1 \end{pmatrix}. \tag{8}$$

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 $t-2$:

$$\begin{aligned} \tau_{i,t,t-1} &= \tau_1 \left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, z_{it}^{v}, z_i^{f}, k_i^{y} \right) \\ \text{and} \quad \tau_{i,t,t-2} &= \tau_2 \left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z_{it}^{v}, z_i^{f}, k_i^{y} \right). \end{aligned} \tag{9}$$

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

With the assumptions and notation introduced in equations (7) to (9), the likelihood of the typical individual's income trajectory $\mathbf{y}_i$ defined in (6) now becomes:

$$\ell_i \left(\mathbf{y}_i \mid \mathbf{e}_i, \mathbf{pub}_i, \mathbf{z}_i^{v}, z_i^{f}, k_i^{y} \right) = \left[\prod_{t=1}^{T} \frac{1}{\sigma_{it}} \right] \times \left[\frac{\prod_{t=3}^{T} \varphi_3 \left(\widetilde{y}_{it}, \widetilde{y}_{i,t-1}, \widetilde{y}_{i,t-2}; \underline{\tau}_{it}^{(3)} \right)}{\prod_{t=3}^{T-1} \varphi_2 \left(\widetilde{y}_{it}, \widetilde{y}_{i,t-1}; \underline{\tau}_{it}^{(2)} \right)} \right], \tag{10}$$

where $\varphi_n\left(\cdot; \underline{\tau}^{(n)}\right)$ is the n-variate normal pdf with mean 0 and covariance matrix $\underline{\tau}^{(n)}$.

Three comments. A first comment brought about by our specification of the income process is that in effect we assume (normalized) incomes to follow a familiar AR(2) process:

$$\widetilde{y}_{it} = \rho_{i,t,t-1} \cdot \widetilde{y}_{i,t-1} + \rho_{i,t,t-2} \cdot \widetilde{y}_{i,t-2} + \varepsilon_{it}, \tag{11}$$

where the innovations ε_{it} are normal with zero mean and serially uncorrelated and where the AR coefficients are related to the autocorrelation coefficients of income disturbances as:

$$\rho_{i,t,t-1} = \frac{\tau_{i,t,t-1} - \tau_{i,t,t-2} \cdot \tau_{i,t-1,t-2}}{1 - \tau_{i,t-1,t-2}^2} \qquad \text{and} \qquad \rho_{i,t,t-2} = \frac{\tau_{i,t,t-2} - \tau_{i,t,t-1} \cdot \tau_{i,t-1,t-2}}{1 - \tau_{i,t-1,t-2}^2}. \tag{12}$$

Thus, while staying close to popular linear Markov models of income dynamics, we allow for more flexibility than the standard linear AR(2) model by letting the various τ's (and thus the ρ's) depend on individual (observed and unobserved) attributes as in (9). This will prove useful in fitting the observed mobility of income ranks. But most importantly, this is necessary to inform one of the key questions of this paper, that is how income mobility varies across sectors (and individuals). Our model offers at least two simple indices of income mobility (the τ's and the ρ's) which can be used to address this question.[21]

The second comment relates our paper to the recent literature on the use of *copulas* in models of income dynamics and sheds additional light on the status of the the τ's as measures of relative income mobility. An alternative way of approaching the likelihood of individual income sequences consists in decomposing the joint density $\ell(y_t, y_{t-1}, y_{t-2})$ appearing in the numerator of (6) 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 $c_{t,t-1,t-2}[F_t(y_t), F_{t-1}(y_{t-1}), F_{t-2}(y_{t-2})]$. The copula density is the joint density of successive income ranks and thus describes *relative* income mobility between $t-2$ and t.[22] Similarly, we can write the joint density $\ell(y_t, y_{t-1})$ appearing in the denominator of (6) as $f_t(y_t) \times f_{t-1}(y_{t-1}) \times c_{t,t-1}[F_t(y_t), F_{t-1}(y_{t-1})]$, 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:

$$\ell(\mathbf{y}) = \left[\prod_{t=1}^{T} f_t(y_t)\right] \times \left[\frac{\prod_{t=3}^{T} c_{t,t-1,t-2}[F_t(y_t), F_{t-1}(y_{t-1}), F_{t-2}(y_{t-2})]}{\prod_{t=3}^{T-1} c_{t,t-1}[F_t(y_t), F_{t-1}(y_{t-1})]}\right]. \tag{13}$$

The use of copulas thus allows decomposition of the likelihood of an income trajectory into the product of cross-section densities at all time periods and the likelihood of the trajectory of relative income ranks within these cross-section distributions.

In the special case of our model, given our normality assumptions, the marginal distributions of income levels are all normal (e.g. $f_t(y_t) = \frac{1}{\sigma_{it}} \cdot \varphi(\widetilde{y}_{it})$ and $F_t(y_t) = \Phi(\widetilde{y}_{it})$ where $\Phi(\cdot)$ is the

[21]We choose to parameterize the autocovariances of income disturbances (the τ's—see equation (9)) rather than the AR coefficients (ρ's) because the former are the moments that we end up matching in the estimation (see Appendix B). In fact, the τ's are easier to handle as they come up "naturally" in the likelihood function (10) and also because the fact that $\underline{\tau}_{it}^{(n)}$ is a correlation matrix imposes a number of natural restrictions on its elements (for instance, because $\tau_{i,t,t-1}$ is a correlation coefficient, it has to lie within the interval $[-1;1]$). The corresponding restrictions on the AR coefficients are more cumbersome.

[22]Copulas are a standard element of the empirical finance toolkit. They were recently brought to labor economics by Bonhomme and Robin (2004). The (unique) decomposability of any (continuous) joint density into the product of marginal densities and a copula density is a result known as Sklar's theorem.

standard normal cdf) and the copulas of income ranks are Gaussian with parameters $\underline{\tau}_{it}^{(n)}$.[23] Yet the formulation (13) of the likelihood of income trajectories holds true independently of any functional form assumption on either the marginal income distributions or the copulas. So we could use this specification to estimate income trajectories under more general distributional assumptions. Its equivalence to the AR(2) specification of income processes, however, rests on our normality assumptions for both the marginal and copula distributions.

The third and final comment bears on our overall specification of the joint income and job mobility process. An implicit assumption in the model is that innovations to the income process (the ε_{it} shocks in equation (11)) are independent of the transitory shocks governing mobility across labor market states (the underlying error terms in the various logit models of (3)). This amounts to assuming that the individual income process only influences individual mobility between labor market states through either observed individual attributes (education, experience...) or through the unobserved individual random effects k_i^m and k_i^y (the latter being time-invariant), and not through transitory income shocks. While feasible in principle, relaxing this restriction would involve great numerical complexity as it would destroy the separability property of the likelihood function into a block relating to labor market mobility and another relating to income.

4.5 Likelihood maximization

Given the specifications of individual contributions to the complete likelihood, $\mathcal{L}_i(\mathbf{x}_i, k_i)$ defined in the previous three sub-sections, we obtain parameter estimates by maximization of the *sample* log-likelihood:

$$\sum_{i=1}^{N} \log \left(\sum_{k_i^m=1}^{K^m} \sum_{k_i^y=1}^{K^y} \mathcal{L}_i \left[\mathbf{x}_i, (k_i^m, k_i^y) \right] \right), \tag{14}$$

where individual random effects $k_i = (k_i^m, k_i^y)$ were integrated out of the complete likelihood (1). The procedure that we use to carry out the maximization of (14) is fully described in Appendix

[23]The (n-variate) Gaussian copula has cdf $C_n(u_1, \dots, u_n; \underline{\tau}^n) = \Phi_n\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n); \underline{\tau}^n\right)$. The corresponding copula density is thus defined by

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

It is straightforward to check that using these normal margins and Gaussian copulas in (13) delivers exactly (10).

B.[24]

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 $K^y = 2$.[25]

Given our choice of six classes, the number of parameters to estimate is 87. Tables of coefficient estimates and standard errors are reported in Appendix C. In this section, rather than commenting on each coefficient directly, we chose to focus on more readily interpretable statistics pertaining to the various building blocks of our model.

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 14% of the population) has a very high unemployment rate of 39%, 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. 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 70% of the sample.

< Tables 3 and 4 about here. >

[24] The highly nonlinear nature of (14) and the relatively large number of parameters to be estimated render any frontal attack on this maximization problem numerically cumbersome. The main numerical issue is to find initial parameter values to feed into the optimization routine that are not "too far" from the optimal values. We thus proceed in two stages, by first "calibrating" the model using a sequential version of the EM algorithm (Dempster et al., 1977), and then using the parameter estimates obtained from this first stage as initial values for the full maximization of (14).

[25] 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.

Table 4 shows the average year-to-year transition probabilities between labor market states, separately for each class and for the whole sample. These transition probabilities are "clean" of selection effects in that they are all based on the whole sample rather than on the subsamples of individuals effectively observed in a given sector. Comparison of the whole sample panel of Table 4 with the raw transition probabilities in subsection 3.3 points to nonrandom selection of individuals into labor market states. Most remarkably, unemployment is almost 3.5 times as persistent in the selected population of unemployed workers as in the whole sample. Also, the public-private gap in job loss rates, while still negative, is much smaller (divided by four as a proportion) once selection effects are removed. All this confirms the importance of accounting for individual heterogeneity.

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 private than in the public sector for classes 1 and 2, and are roughly similar for the private class ($k^m = 3$). Members of mobility class 2 are otherwise 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.

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 exhibits an education composition concentrated in the high and low groups. 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 workers. As they account for 70% 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 58.3% for class $k^y = 1$ and 41.7% for class $k^y = 2$. The top two panels of Figure 2 show mean income together with the 10th and 90th percentiles of income as a function of experience for each income class and sector. The bottom left panel shows similar profiles for the whole sample correcting for selection effects (i.e. assuming that everyone is in either sector), while the bottom right panel shows these profiles given the selection of individuals into sectors. The top two panels of Figure 2 show that members of income class 1 tend to earn higher incomes. Their income-experience profile is also "more concave", meaning that the mean returns to experience are higher for members of income class 1 at early stages of their working lifes, and taper off faster as they grow more experienced. Finally, income dispersion varies very little with experience, as the 10th and 90th percentile lines are approximately parallel on all panels.[26]

< Figure 2 about here. >

The striking fact about Figure 2, however, is that the public gap is very different between classes: while members of income class 2 clearly benefit from public sector employment, both in terms of mean income and income dispersion, the public premium is virtually inexistent on both accounts for members of income class 1. In figures, the average public premium is of 15 log points for class 2 and slightly under 1 log point for class 1. The cross-sector differences in income standard deviations are of 20 log points for class 2 and less than 1 log point for class 1. Returns to experience look very similar in both sectors for each income class. In search of a better denomination, we shall henceforth refer to members of income class 1 (resp. income class 2) as "low public premium" (resp. "high public premium"), bearing in mind that class 1 also tends to earn overall higher incomes than class 2.

Looking at whole sample figures, the average potential public premium is of 7 log points. The

[26] In figures, the cross-sectional standard deviation of income conditional on experience increases significantly by 2.7 log points every additional ten years of experience in the private sector and by 1.1 log points in the public sector. Table C2 in Appendix C also gives direct information on the experience profile of the standard deviation of income disturbances. This increases by 4.1 log points every ten years of experience, both sectors pooled.

cross-sector gap in income dispersion is more substantial, with a markedly lower (10 log points) standard deviation of income in the public sector. This squares in well with the recurring finding of income compression in the public sector. Finally, the average income public premium in the *selected* sample (bottom-right panel) is higher at about 17 log points. Comparison of the bottom two panels of Figure 2 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.

< Table 6 and Figure 3 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 members of income class 1 (resp. class 2) also qualify as "income stayers" (resp. "income movers"). In other words, members of income group 2 tend to have earnings that are both slightly lower on average and less stable over time than members of income group 1. Part of this instability naturally results from the likely higher *job* instability of type $k^y = 2$ 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 3, which shows the autocovariance at one lag of income disturbances (the τ_{it}^1 coefficient in the 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 1 tend to be "income-stayers" while class 2 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 2 than on members of class 1 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 2 (the "high public premium" 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 "high public premium" type ($k^y = 2$) is likely associated with being a "low-employability" mobility type ($k^m = 2$). "High public premium" income types are also slightly under-represented in the "private" mobility class ($k^m = 1$). 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 least frequent combination is for an individual to be a "low public premium" type and at the same time a low-employability mobility type (i.e. $k^y = 1$ and $k^m = 2$)—these only account for 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 $\mathbf{x}_i$, $\Pr\{k_i^m = k^m, k_i^y = k^y \mid \mathbf{x}_i\}$.[27]

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

[27] These probabilities are computed in the E-step of the EM algorithm which we use in the calibration stage of our estimation procedure. See Appendix B.

matrices at intervals of one, two and six years (which is the maximum possible lag in a seven-year panel).[28] Here again the fit is excellent. The maximum discrepancy between observed and simulated transition matrices is less than 2%.[29]

< 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 a third) 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 to their model-based counterparts. The bottom panel of Table 7 presents this comparison for a lag of 10 years.[30] The maximum absolute prediction error that we make here is in the order of 10 percentage points. The main discrepancy between the observed and predicted transition matrices is that we seem to be over-predicting 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 satisfactory enough. We should indeed bear in mind that

[28] Unreported matrices are available on request and convey a similar message.

[29] 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.

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

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 4 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 4 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 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 (253, all years pooled).[31] In any case it appears that the model generates an acceptable smoothing of the observed density of unemployment income.[32]

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

[31] Moreover, somewhat surprisingly, the "unemployment" graph reveals that class 1 earns *less* than class 2 in unemployment (-25 log points). Yet, as we also see from this graph, members of income class 1 are unemployed much less often than members of income class 2. Moreover, the variance of unemployment incomes is a lot smaller among members of income class 1 than for income class 2.

[32] The "observed" density plotted on Figure 4 is produced by a Gaussian kernel estimator with a bandwidth determined using the rule of thumb $h = 0.9 \times \mathrm{std}\,(\text{log income})\,/n^{1/5}$, which already does a certain amount of smoothing to the corresponding histogram.

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.[33] The good fit that we obtain thus constitutes strong support for our model specification.

A discussion of possible alternative specifications. Incomes 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. Yet the computational cost of our estimation procedure increases very quickly as one expands the model in either dimension, often for relatively modest 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,

[33]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.

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 one of 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 quite as well: estimating our model with two classes of income heterogeneity under the assumption that $\rho_{i,t,t-2} = 0$ (see equation (11)) 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 $\{\text{Markov}(1),\ K^y = 3\}$ as in the $\{\text{Markov}(2),\ K^y = 2\}$ 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 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.

[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).

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 counterfactual assumptions on sector choice.

6.1 Construction of lifetime values

Definitions. Using our estimated coefficients for income distributions, income and job mobility, we can 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:

$$V_t\left(\mathbf{y}_{s\geq t}\right) = \sum_{s=t}^{T_R} \beta^{s-t} \cdot U\left(y_s\right) + \beta^{T_R - t} \cdot V_R, \tag{15}$$

where $\beta \in (0,1)$ is a discount factor and $U\left(y_t\right)$ 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 as specified in our statistical model (see Section 4 and Appendix A).

We can naturally consider different functional forms and/or parameterizations for (15) 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:

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

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

The logarithmic specification simply assimilates lifetime values $V_t\left(\mathbf{y}_{s\geq t}\right)$ to present discounted sums 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 A5 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 insurance 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.[37] 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, $\eta = 0.5$.

We set the discount factor to $\beta = 0.9$ *per annum*. Finally, we define the value of retirement as $V_R = \frac{1}{1-\beta} \cdot U\left(y_{T_R-1} + \ln RR\right)$ in both the logarithmic and the CRRA case, where RR designates the replacement ratio. That is, we assume that after retirement, individuals receive a constant flow of income equal to RR times their last income in activity and discount this flow over an infinite number of years.[38] We calibrated the value of RR 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,

[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 $\eta = +\infty$.

[37] In the end, the question of which specification is more adequate boils down to that of how closely consumption tracks income. The recently growing literature on consumption inequality seems to conclude that the truth lies somewhere in between the perfect insurance and the consumption equals income cases. See e.g. Gourinchas and Parker (2002), Blundell, Pistaferri and Preston (2004) for work on US data, or Jappelli and Pistaferri (2004) on Italian data.

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

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:

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

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

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 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 and differences in income mobility across sectors. Ruling out inter-sector mobility (whose role we shall explore in the next subsection) allows us to obtain a neat picture of the impact of income mobility on the public lifetime premium.

Figure 5 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 5 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 9 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 smaller than the income public premium in the first two deciles of the distribution (10 to 15 log points) and larger in the upper part (5 to 6 log points in the top two deciles). Moreover, the lifetime public gap remains positive throughout lifetime value quantiles, whereas the income gap becomes negative in the top quartile of income. The CRRA case differs from the logarithmic utility case in that 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.

The striking result is thus that there is considerably less compression in the public sector relative to the private sector in terms of lifetime values than in terms of income. We interpret this phenomenon as the result of income mobility offsetting differences in cross-sectional incomes: intuitively, 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. To further investigate this intuitive statement, let us consider some generic variable X_i representing either incomes or lifetime values in the initial 1996 cross-section of our sample. Given permanent employment in either the public or private sector (i.e. $e_i = 1$ and $\text{pub}_i = 0$ or 1), we can define the permanent component of X_i as the projection of X_i on time-invariant individual attributes, i.e. $E\left(X_i \mid e_i, \text{pub}_i, z_i^f, k_i^m, k_i^y\right)$ and consider the following variance decomposition for X_i:

$$\text{Var}\left(X_i \mid e_i, \text{pub}_i\right) = \text{Var}\left[E\left(X_i \mid e_i, \text{pub}_i, z_i^f, k_i^m, k_i^y\right)\right] + E\left[\text{Var}\left(X_i \mid e_i, \text{pub}_i, z_i^f, k_i^m, k_i^y\right)\right]. \quad (16)$$

The first term in the r.h.s. of (16) reflects the contribution of the permanent component of X_i to $\text{Var}\left(X_i \mid e_i, \text{pub}_i\right)$, while the second term relates to the transitory component of X_i.

< **Table 9 about here.** >

Table 9 reports the variance decomposition in (16) for income and lifetime values (still under the "job for life" assumption) in the public and private sectors.[39] The first row relates to income and confirms the well-documented public sector income compression: cross-sectional log income variance is 50 percent larger in the private sector. The decomposition further indicates that this differential is mainly driven by a greater variance of the transitory component of income in the private sector (both in absolute terms and as a fraction of the total variance). Next turning to lifetime values (second row), we no longer see much compression in the public sector, as total cross-sectional variance of lifetime values is only 10 percent smaller in the public than in the private sector. Moreover, the variance of the transitory component of lifetime values is the same in both sectors and now explains a larger share of total variance of lifetime values in the public than in the private sector. Comparison of the two rows consistently reveals that going from income to lifetime

[39] All reported numbers relate to the whole sample and thus they do not reflect endogenous selection of individuals into sectors. Moreover we focus on the logarithmic specification of utility functions, as log-income flows are more directly comparable with present discounted values of log income flows. Results for the CRRA specification, which brings in an additional nonlinear transformation of log income, are available upon request.

values substantially reduces the relative importance of the transitory component in total variance in the private sector, while it leaves the variance decomposition essentially unaffected in the public sector.

To understand those decompositions more clearly, let us recall that log income flows y_{it} are modeled as the sum of a predictable mean μ_{it} and a (normal) persistent disturbance $\sigma_{it} \cdot \widetilde{y}_{it}$ that follows an AR(2) process—see equations (7) and (11). This latter is transitory by its nature, and surely constitutes the main part of the transitory component in equation (16) and Table 9.[40] We found evidence above (see Figures 2 and 3) that these income disturbances exhibit greater variance but less persistence in the private than in the public sector. This is consistent with the observation from Table 9 that the variance of the transitory component is reduced when taking lifetime values to a much greater extent in the private than in the public sector.

The overall conclusion that we can draw from these results is twofold. First, income inequality is greater yet less persistent in the private sector than in the public sector. Second, we find that there is no sizeable cross-sector difference in "long-term" inequality as measured by the variance of present discounted sum of log incomes.

Selection into sectors. 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 17 log points for the logarithmic utility and 21 for the CRRA utility. Comparing this with the "whole sample" graph suggests that about 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).

[40] Even though the two concepts are not fully confounded. Given our maintained assumption of permanent employment in either the public or private sector (i.e. $e_{it} \equiv 1$ and $\text{pub}_{it} \equiv 0$ or 1 for all t), the deterministic component of income μ_{it} still varies over time because of the returns to experience. Yet these returns differ little between the two sectors. And of course part of the transitory component in Table 9 may also reflect measurement error, the extent of which may potentially differ between sectors. Our results are subject to this latter qualification.

Conditional on being employed, both income classes are split between the public and private sector in roughly similar proportions (17% public for income class 1, 18% for income class 2). Yet, as one sees on Figure 2, members of income class $k^y = 2$—which we labeled the "high public premium" workers—enjoy a much higher income premium in the public sector than members of income class $k^y = 1$. One would thus expect to see members of income class 2 go to the public sector in far *larger* proportions. However, as we pointed out in Section 5 (Table 5), being a "high public premium" worker 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 phenomenon of job queuing in the public sector, something we shall return to at various stages of the discussion.

Conditioning on individual heterogeneity. Figure 6 displays means of lifetime public sector premia by groups of a given education level, experience range and unobserved heterogeneity class, together with the corresponding 95% confidence intervals. The top row shows a breakdown of the sample by mobility classes while the bottom row shows a breakdown by income classes. Lifetime values are computed under the logarithmic specification.[41]

< Figure 6 about here. >

Looking at all 9 panels, one first sees that lifetime public premia vary widely across groups, from a low of a negative few log points for members of income class 1 at early stages of their working lives to a high of 15 to 20 log points for low-educated, senior members of the "low employability" mobility class ($k^m = 2$) or the "high public premium" income class ($k^y = 2$). Also it appears that those conditional lifetime public premia tend to decrease somewhat with education and increase with experience, with the exception of the low-educated subgroup of the "public" mobility class ($k^m = 1$), for which lifetime public premia are highest at the beginning of the working life (i.e. under 4 years of experience).

The top row further shows that members of the "low employability" mobility class ($k^m = 2$)

[41]Graphs for the CRRA specification are available on request. Also note that, to save on space, we do not report lifetime premia conditional on interactions of both dimensions of unobserved heterogeneity, k^m ans k^y.

generally benefit from larger lifetime public premia than other mobility classes by around 5 log points. This combined with the fact that the unemployment rate of this mobility class is highest (at 39%) is suggestive of job queuing for public sector jobs, whereby individuals from this class find it difficult to find and retain public sector employment. Members of the "public" mobility class ($k^m = 1$) who are found to work predominantly in the public sector enjoy lifetime value premia in that sector that are equal or greater than those enjoyed by members of the "private" mobility class ($k^m = 3$). These differences are significant for low-educated groups at all levels of experience and for the highly educated in the experience range of 5 to 14 years. While these differences in lifetime public premia between the "private" and "public" mobility classes are admittedly tenuous, they still suggest that expected lifetime premia are related to the selection pattern of individuals into job sectors.[42]

Now turning the bottom panel of Figure 6, we see that lifetime public premia are consistently and significantly higher for groups of individuals in the "high public premium" income class ($k^y = 2$) than for groups of individuals in the other income class, as the name suggests. This difference in premia is around 10 log points at all levels of education and experience.

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 allowing him to move between sectors or into unemployment thereafter according to his predicted transition probabilities. Public premia based on this definition are depicted on Figure 7 for the whole sample.[43]

< Figure 7 about here. >

The striking point about Figure 7 is that in all panels, most public-private differences in lifetime values disappear 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

[42] An additional qualification that applies to this argument is that lifetime public premia are significantly positive even for groups of individuals in the "private" class ($k^m = 3$).

[43] Again to save on space, we do not display the corresponding analysis by groups of individual heterogeneity.

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 few percentiles). 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 88% have left the public sector within one year and another 10% do so within 2 years. Almost all of these leavers (99%) go to private sector employment "directly", while the remaining 1% 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" 45% leave within one year and another 24% within 2 years, again almost exclusively (99%) toward a public sector job. 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, 28% of them leave within one year and another 21% within 2 years, 80% of these because they have found employment in the private sector. However, if they start off in the private sector, only about 6% leave every year over the first few years, and fewer thereafter, 54% of these to take up employment in the public sector while the other 46% 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

The above developments involve calibrated values for the annual discount rate β and the income replacement ratio at the point of retirement, RR. The particular values of $\beta = 0.9$ and $RR = 0.4$ that were adopted obviously condition the results described in this Section.

The impact of a change in the rate of time discounting is obvious enough, as a smaller value of β clearly makes lifetime values stick more closely to current income flows. Hence for brevity we choose not to report any quantitative sensitivity analysis on β.[44]

As for the sensitivity of our results to the replacement ratio, the following can be said. First, due to time discounting, increasing the replacement ratio only has a sizeable (positive) impact on lifetime values of individuals who are close to retirement age. Second, a uniform increase in the value of the replacement ratio has little discernible impact on the public premium in lifetime values. (It only magnifies the impact of the income gap at the time of retirement.) Yet, there are institutional differences in pension schemes between the public and the private sector, which might potentially translate into differences in replacement ratios across sectors. While such differences do not appear in our data, recent policy changes regarding pension schemes in the private sector are likely to impact on these differences.

7 Conclusion

The current debate on the existence and size of differences between public and private sector pay usually focuses on cross-section differences in earnings. In a dynamic environment, individuals will however anticipate changes in their employment status, sector of employment and income level within a given sector. This will matter in their assessment of the lifetime value of potential employment in either sector. In this paper we take a view on public-private differences beyond cross-section to account for job and income mobility. To this end, we construct a flexible model of income

[44] Such quantitative analysis is available on request.

dynamics and of selection of individuals into employment sectors, where individual income and employment processes are conditioned by observed as well as unobserved individual heterogeneity.

Estimating this model over the last 7 waves of the BHPS, we are able to replicate very well worker mobility across labor market states and income trajectories, both in and out of sample. We hence use our estimated model to simulate individual working lives for all workers in our sample and are able to compute lifetime values of their income streams and compare these across sectors.

Our main results are the following. The average public premium in terms of present discounted sum of future log-incomes is positive at 8 percent, were individuals planning to remain employed all their working life in either sector. If, however, job mobility is taken into account in that individuals are allowed to switch employment sector in the simulations, the public premium in lifetime values becomes essentially zero for "highly employable" individuals. This suggests that the UK labor market is sufficiently mobile to ensure a rapid allocation of these workers into their "natural" sector. Finally, although we document the well-known income compression in the public sector, we find that this compression phenomenon mostly disappears when considering lifetime values. We thus argue that the greater variance of private sector incomes relates to the transitory component of income.

As to general leads for future research, our model of employment and income dynamics is flexible enough and can be fairly easily estimated and used to simulate individual income and employment trajectories. It hence can be a useful tool for a variety of applications in the study of inequality between groups or over time in the labor market.

Pointing back to the public gap, for the sake of brevity, we have restricted the analysis in the present paper to male individuals. Looking at the female dataset has been left for future research and raises additional issues. First, average hours worked and the extent of part-time work vary substantially across sectors, and would require a finer description of the labor supply behavior. Second, non-wage job characteristics, such as different provision of maternity benefits and flexible hours of work are likely to influence the selection of women across sectors.

Another interesting extension of this paper would be to carry out this analysis with data for other countries. Studies on pay differences between public and private sectors document the phenomenon of income compression for numerous countries. Our model provides a way of assessing whether this pertains mainly to the transitory component of income, as we find for the UK, or whether it reflects some more permanent inequality between sector-specific income distributions, though estimation of our model would require panel datasets with similar characteristics to the BHPS (long enough longitudinal dimension and still a reasonable cross-sectional width).

References

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

[2] Belman, D., and J.S. Heywood (1989), “Government Wage Differentials: A Sample Selection Approach”, Applied Economics, 21, 427-438.

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

[4] Blundell, R., L. Pistaferri, and I. Preston (2004), “Consumption Inequality and Partial Insurance”, University College London and Stanford University mimeo.

[5] Blundell, R., and I. Preston (1997), “Consumption Inequality and Income Uncertainty”, Quarterly Journal of Economics, 113(2), 603-640.

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

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

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

[9] Buchinsky, M., and J. Hunt (1999), "Wage Mobility in the United States", Review of Economics and Statistics, 81(3), 351-68.

[10] Clark, A. E., and F. Postel-Vinay (2004), "Job Security and Job Protection", CEPR Discussion Paper 4927.

[11] 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.

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

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

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

[15] 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.

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

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

[18] 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.

[19] Gourinchas, P.-O., and J. A. Parker (2002), "Consumption Over the Life Cycle", Econometrica, 70(1), 47-89.

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

[21] Jappelli, T., and L. Pistaferri (2004), "Intertemporal Choice and Consumption Mobility", Stanford University mimeo.

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

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

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

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

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

[27] Poterba, J.M., and K.S. Rueben (1994), "The Distribution of Public Sector Wage Premia: New Evidence Using Quantile Regression Methods", NBER Working Paper, 4734.

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

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

[30] 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(2), 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 individual observation i 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 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:

$$\Pr\left\{k_i^m = k^m \mid z_i^f\right\} = \frac{\exp\left(z_i^{f\prime} \cdot \kappa_{k^m}^m\right)}{\sum_{k=1}^{K^m} \exp\left(z_i^{f\prime} \cdot \kappa_k^m\right)} \quad \text{and} \quad \Pr\left\{k_i^y \mid k_i^m, z_i^f\right\} = \frac{\exp\left[\begin{pmatrix} z_i^f \\ k_i^m \end{pmatrix}' \cdot \kappa_{k^y}^y\right]}{\sum_{k=1}^{K^y} \exp\left[\begin{pmatrix} z_i^f \\ k_i^m \end{pmatrix}' \cdot \kappa_k^y\right]}, \tag{A1}$$

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:

$$\begin{aligned}
\ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i^m\right) &= \Pr\left\{e_{i1} \mid z_i^f, k_i^m\right\} \times \left[\Pr\left\{\mathrm{pub}_{i1} \mid z_i^f, k_i^m\right\}\right]^{e_{i1}} \\
&\times \prod_{t=2}^{T_i} \left(\Pr\left\{e_{it} \mid e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} \cdot \left[\Pr\left\{\mathrm{pub}_{it} \mid e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\}\right]^{e_{it}}\right).
\end{aligned} \tag{A2}$$

All components are specified as logits:

$$\begin{aligned}
\Pr\left\{e_{it} \mid e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} &= \Lambda\left(\left[e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{i,t-1}^v{}', z_i^{f\prime}, k_i^{m\prime}\right] \cdot \psi\right), \\
\Pr\left\{\mathrm{pub}_{it} \mid e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\} &= \Lambda\left(\left[e_{i,t-1}, \mathrm{pub}_{i,t-1}, z_{i,t-1}^v{}', z_i^{f\prime}, k_i^{m\prime}\right] \cdot \chi\right),
\end{aligned} \tag{A3}$$

where $\Lambda(x) = (1+e^{-x})^{-1}$ designates the logistic cdf.[45] We finally use *mutatis mutandis* similar specifications for the initial job state probabilities:

$$\Pr\left\{e_{i1} \mid z_i^f, k_i^m\right\} = \Lambda\left(\left[z_i^{f\prime}, k_i^{m\prime}\right] \cdot \psi_0\right) \quad \text{and} \quad \Pr\left\{\text{pub}_{i1} \mid z_i^f, k_i^m\right\} == \Lambda\left(\left[z_i^{f\prime}, k_i^{m\prime}\right] \cdot \chi_0\right). \tag{A4}$$

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 (7) and (9). We start with $\mu(\cdot)$:

$$\mu\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right) = z_i^{f\prime} \cdot \mu_0 + \left[\begin{pmatrix} e_{it} \\ \text{pub}_{it} \end{pmatrix} * k_i^y * z_{it}^v\right]' \cdot \mu_1, \tag{A5}$$

where the notation $x * y$ stands for all main effects and interactions of variables x and y. 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:

$$\sigma\left(e_{it}, \text{pub}_{it}, z_{it}^v, k_i^y\right) = \sqrt{\exp\left(\left[\begin{pmatrix} e_{it} \\ \text{pub}_{it} \end{pmatrix} * k_i^y\right]' \cdot \sigma_0 + z_{it}^{v\prime} \cdot \sigma_1\right)}. \tag{A6}$$

Here we force the variance 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 also omit interactions of employment state or income class and experience.

We finally turn to the dynamics of income, which are governed by the functions $\tau_1(\cdot)$ and $\tau_2(\cdot)$. The first-order autocorrelation of income, $\tau_1(\cdot)$, is specified as:

$$\begin{aligned} &\tau_1\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, z_{it}^v, k_i^y\right) \\ &\quad = -1 + 2 \cdot \Lambda\left(\left[z_{it}^v * k_i^y\right]' \cdot \zeta_0 + \left[\begin{pmatrix} e_{it} \\ \text{pub}_{it} \end{pmatrix} * k_i^y\right]' \cdot \zeta_1 + \left[\begin{pmatrix} e_{i,t-1} \\ \text{pub}_{i,t-1} \end{pmatrix} * k_i^y\right]' \cdot \zeta_2\right). \end{aligned} \tag{A7}$$

This specification calls for some comments. First, the transformation $-1 + 2 \cdot \Lambda(\cdot)$ 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, $(e_{it}, \text{pub}_{it})$ and $(e_{it-1}, \text{pub}_{it-1})$.[46]

[45] Note that unobserved mobility heterogeneity k_i^m only affects the constant term in this specification. The number of observed sector transitions in our data set does not seem to be sufficient for a precise estimation of less restrictive specifications. For instance we tried allowing the effect of experience on sector mobility to differ between mobility classes by considering interaction terms of k_i^m and $z_{i,t-1}^v$. The corresponding parameters turned out to be very poorly estimated.

[46] 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 recall the notation shortcuts designating the date-t, one- and two-lag autocorrelations of normalized income used in the main text (equation (9)):

$$\begin{aligned} \tau_{i,t,t-1} &= \tau_1\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, z^v_{it}, k^y_i\right) \\ \text{and} \quad \tau_{i,t,t-2} &= \tau_2\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z^v_{it}, k^y_i\right). \end{aligned}$$

Then we write:

$$\begin{aligned} \tau_2\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, e_{i,t-2}, \text{pub}_{i,t-2}, z^v_{it}, k^y_i\right) &= \tau_{i,t,t-1} \cdot \tau_{i,t-1,t-2} \\ &+ \left(\sqrt{\left(1-\tau^2_{i,t,t-1}\right)\cdot\left(1-\tau^2_{i,t-1,t-2}\right)}\right)\cdot\widetilde{\tau}_2\left(k^y_i\right), \end{aligned} \tag{A8}$$

with $\widetilde{\tau}_2\left(k^y_i\right) = -1 + 2\cdot\Lambda\left(k^{y\prime}_i \cdot \xi\right)$ simply specified as an income class-specific constant within $[-1,+1]$.

This last pair of equations requires some clarification. We have to constrain $\tau_2(\cdot)$ in such a way that, given $\tau_{i,t,t-1}$ and $\tau_{i,t-1,t-2}$, the matrix:

$$\underline{\tau}^{(3)}_{it} = \begin{pmatrix} 1 & \tau_{i,t,t-1} & \tau_{i,t,t-2} \\ \tau_{i,t,t-1} & 1 & \tau_{i,t-1,t-2} \\ \tau_{i,t,t-2} & \tau_{i,t-1,t-2} & 1 \end{pmatrix}$$

is a consistent covariance matrix. This is the case provided that its determinant Δ_{it} is positive (and that the various τ's lie in $[-1,+1]$). Δ_{it} is defined by $\Delta_{it} = 1-\tau^2_{i,t,t-1}-\tau^2_{i,t-1,t-2}-\tau^2_{i,t,t-2}+2\tau_{i,t,t-1}\tau_{i,t-1,t-2}\tau_{i,t,t-2}$. Solving for $\tau_{i,t,t-2}$, we get:

$$\tau_{i,t,t-2} = \tau_{i,t,t-1}\cdot\tau_{i,t-1,t-2} \pm \sqrt{\left(1-\tau^2_{i,t,t-1}\right)\cdot\left(1-\tau^2_{i,t-1,t-2}\right)-\Delta_{it}}. \tag{A9}$$

Because Δ_{it} is positive, $\tau_{i,t,t-2}$ has to stay within the interval

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

This is achieved by the parameterization in equation (A8) given the constraint $\widetilde{\tau}_2(\cdot) \in [-1,+1]$.

B Estimation details

In this Appendix we fully describe the procedure that we use for the maximization of the sample log-likelihood (14). As briefly explained in the main text (sub-section 4.5, footnote 24), this procedure consists of a first "calibration" stage, from which we get parameter estimates to use as initial values in a second "likelihood maximization" stage.

B.1 Stage 1: Model calibration using a sequential EM algorithm

B.1.1 General description

From Appendix A and Section 4 in the main text, the list of parameters to be estimated can be divided into 2 subsets, $\Theta^m = \left\{\left(\kappa^m_k\right)^{K^m}_{k=1}, \psi_0, \chi_0; \psi; \chi\right\}$ and $\Theta^y = \left\{\left(\kappa^y_k\right)^{K^y}_{k=1}, \mu(\cdot), \sigma(\cdot), \tau_1(\cdot), \widetilde{\tau}_2(\cdot)\right\}$, where by $\{\mu(\cdot), \sigma(\cdot), \tau_1(\cdot), \widetilde{\tau}_2(\cdot)\}$ we summarize all parameters of the corresponding functions—see equations (A5), (A6), (A7) and (A8). The first subset, Θ^m, gathers all parameters from the job mobility process—i.e. involved in equations (A1), (A3) and (A4). The subset Θ^y contains all remaining parameters, which all pertain to the income process—see equations (A5) through (A8).

The structure of (1) implies that individual contributions to the complete likelihood can be decomposed as $\mathcal{L}_i(\mathbf{x}_i, k_i; \Theta^m, \Theta^y) = \mathcal{L}_i^m(\mathbf{x}_i, k_i^m; \Theta^m) \cdot \mathcal{L}_i^y(\mathbf{x}_i, k_i^m, k_i^y; \Theta^y)$, where:

$$\mathcal{L}_i^m(\mathbf{x}_i, k_i^m; \Theta^m) = \ell_i\left(\mathbf{e}_i, \mathbf{pub}_i \mid \mathbf{z}_i^v, z_i^f, k_i^m; \Theta^m\right) \cdot \Pr\left\{k_i^m \mid z_i^f; \Theta^m\right\} \tag{B1}$$

(note that we now make the dependence on the various parameters explicit). This particular structure makes it very easy to integrate income sequences $(\mathbf{y}_i)$ and income classes (k_i^y) out of $\mathcal{L}_i(\mathbf{x}_i, k_i; \Theta^m, \Theta^y)$. We can thus recover the parameters pertaining to the job mobility process and job mobility classes by separately considering the likelihood of observed job sector mobility, $\sum_{i=1}^N \log\left(\sum_{k_i^m=1}^{K^m} \mathcal{L}_i^m(\mathbf{x}_i, k_i^m; \Theta^m)\right)$. Given the simple structure of our job mobility model, this maximization can be achieved by a straightforward application of the EM algorithm for finite mixtures, which we describe below in paragraph B.1.2.

Once this first sub-stage has delivered estimates of the "job mobility" parameters, we fix those at their estimated values $\widehat{\Theta}^m$ and turn back to the maximization of the sample likelihood, $\mathcal{L}_i\left(\mathbf{x}_i, k_i; \widehat{\Theta}^m, \Theta^y\right)$, now concentrating on the "income process" part of this likelihood, $\mathcal{L}_i^y(\mathbf{x}_i, k_i^m, k_i^y; \Theta^y)$, and the related parameters Θ^y. We give the technical details of this second sub-stage below in paragraph B.1.3.

B.1.2 Stage 1.1: Calibration of the job mobility parameters Θ^m

The standard EM-algorithm 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:

$$\Pr\{k_i^m = k^m \mid \mathbf{x}_i; \Theta_n^m\} = \frac{\mathcal{L}_i^m(\mathbf{x}_i, k^m; \Theta_n^m)}{\sum_{k=1}^{K^m} \mathcal{L}_i^m(\mathbf{x}_i, k; \Theta_n^m)}. \tag{B2}$$

M-step. Update Θ_n^m into Θ_{n+1}^m by maximizing the following augmented sample log-likelihood, weighted by (B2):

$$\Theta_{n+1}^m = \arg\max_{\Theta^m} \sum_{i=1}^N \sum_{k=1}^{K^m} \Pr\{k_i^m = k \mid \mathbf{x}_i; \Theta_n^m\} \cdot \log\left[\mathcal{L}_i^m(\mathbf{x}_i, k; \Theta^m)\right]. \tag{B3}$$

This latter maximization is easily carried out running separate weighted logit regressions for ψ and χ (see equations (A3) and (A4), and a weighted multinomial logit for the class weight parameters κ_k^m, using (B2) 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.1.3 Stage 1.2: Calibration of the income parameters Θ^y

We then want to calibrate the subset of "income" parameters, Θ^y. The natural "limited information" approach would be to maximize the sample likelihood, $\mathcal{L}_i\left(\mathbf{x}_i, k_i; \widehat{\Theta}^m, \Theta^y\right)$, only fixing Θ^m at its estimated value $\widehat{\Theta}^m$ from stage 1.1. Yet again the highly nonlinear nature of $\mathcal{L}_i^y(\cdot)$—see subsection 4.4—renders even this maximization numerically difficult. At this calibration stage we thus choose not to use direct maximization but rather a sequential, limited information version of the EM algorithm in the spirit of Arcidiacono and Jones (2003) or Bonhomme and Robin (2004). This algorithm goes as follows:

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 $\widehat{\Theta}^m$:

$$\Pr\left\{k_i^m = k^m, k_i^y = k^y \mid \mathbf{x}_i; \widehat{\Theta}^m, \Theta_n^y\right\} = \frac{\mathcal{L}_i\left(\mathbf{x}_i, k^m, k^y; \widehat{\Theta}^m, \Theta_n^y\right)}{\sum_{\ell^m=1}^{K^m}\sum_{\ell^y=1}^{K^y}\mathcal{L}_i\left(\mathbf{x}_i, \ell^m, \ell^y; \widehat{\Theta}^m, \Theta_n^y\right)}. \tag{B4}$$

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 (B4) 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 log income disturbances $\widetilde{y}_{it}^{(n+1)} = \frac{y_{it} - \widehat{\mu}_{n+1}\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right)}{\widehat{\sigma}_{n+1}\left(e_{it}, \text{pub}_{it}, z_{it}^v, z_i^f, k_i^y\right)}$. Update $\tau_1(\cdot)$ as:

$$\widehat{\tau}_{1,n+1}\left(e_{it}, \text{pub}_{it}, e_{i,t-1}, \text{pub}_{i,t-1}, z_{it}^v, k_i^y\right) = \text{cov}\left(\widetilde{y}_{it}^{(n+1)}, \widetilde{y}_{i,t-1}^{(n+1)}\right), \tag{B5}$$

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 (B4) 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 (A8), 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 (A8), 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 (B4) 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}.

B.2 Stage 2: likelihood maximization

The big advantage of the sequential algorithm used in stage 1 is that it is computationally more stable (given arbitrary starting values) and more tractable than direct, "full information" maximization of the total sample likelihood (14). Moreover, it can be shown (under the combined assumption of identification of the model parameters and numerical convergence of the algorithm) to converge to a consistent estimator of the parameters (see Arcidiacono and Jones, 2003, or Bonhomme and Robin ,2004, Appendix C).[47] Its drawback, however, is that it converges to an estimator which differs from the ML estimator and which is not efficient.[48]

In spite of this inefficiency, the estimates of Θ^m and Θ^y obtained from stage 1 turn out to be close enough to full-information ML estimates to at least constitute very good initial values to feed into a standard optimization routine applied to the total sample likelihood (14). Pursuing this idea, stage 2 simply consists of maximizing (14) starting from the values of Θ^m and Θ^y obtained from stage 1.

[47]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

[48]It is inefficient for two reasons. First, it is two-step. Second, the "income" step 1.2 itself is sequential, i.e. in that step we update the parameters sequentially (as opposed to simultaneously) within each iteration. By using such a sequential approach, we throw away some information (in effect, we are only matching a number of selected means and covariances, e.g. we only use income cross-sections—and not income dynamics—to update the mean and variance parameters) and thus lose some efficiency.

C Parameter estimates

Initial unemployment probability: $\Pr\left\{e_{i1}=0 \mid z_i^f, k_i^m\right\}$			
Experience (years/10)	0.030 (0.348)	Experience^2 (years^2/100)	0.100 (0.083)
High education	−0.995 (0.343)	Medium education	−0.192 (0.385)
$k^m = 2$	6.026 (1.247)	$k^m = 3$	1.530 (1.280)
Constant	−6.566 (1.325)		
Subsequent unemployment probability: $\Pr\left\{e_{it}=0 \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\}$, $t \geq 2$			
Experience (years/10)	−0.670 (0.263)	Experience^2 (years^2/100)	0.180 (0.056)
High education	−1.058 (0.234)	Medium education	−0.136 (0.219)
Public last period: $\text{pub}_{i,t-1} = 1$	−0.209 (0.448)	Unemployed last period: $e_{i,t-1} = 0$	2.561 (0.179)
$k^m = 2$	3.827 (0.960)	$k^m = 3$	0.739 (0.971)
Constant	−5.284 (1.025)		
Initial probability of public sector: $\Pr\left\{\text{pub}_{i1}=0 \mid z_i^f, k_i^m\right\}$			
Experience (years/10)	0.402 (0.409)	Experience^2 (years^2/100)	−0.087 (0.096)
High education	1.442 (0.307)	Medium education	0.412 (0.335)
$k^m = 2$	−3.608 (0.524)	$k^m = 3$	−6.386 (0.344)
Constant	1.153 (0.502)		
Subsequent prob. of public sector: $\Pr\left\{\text{pub}_{it}=0 \mid e_{i,t-1}, \text{pub}_{i,t-1}, z_{i,t-1}^v, z_i^f, k_i^m\right\}$, $t \geq 2$			
Experience (years/10)	−0.221 (0.326)	Experience^2 (years^2/100)	0.016 (0.066)
High education	0.863 (0.233)	Medium education	0.249 (0.274)
Public last period: $\text{pub}_{i,t-1} = 1$	3.610 (0.178)	Unemployed last period: $e_{i,t-1} = 0$	0.784 (0.368)
$k^m = 2$	−2.710 (0.296)	$k^m = 3$	−5.382 (0.354)
Constant	−0.174 (0.461)		

Table C1: Parameters of job sector mobility (logit models)

Income means: $\mu\left(e_{it}, \mathrm{pub}_{it}, z^v_{it}, z^f_i, k^y_i\right)$			
High education	0.440 (0.015)	Medium education	0.134 (0.015)
Experience (years/10)	0.529 (0.023)	Experience^2 (years^2/100)	−0.084 (0.005)
Public: $\mathrm{pub}_{it} = 1$	−0.012 (0.034)	Experience×Public	−0.007 (0.031)
Experience^2×Public	0.005 (0.006)	Unemployed: $e_{it} = 0$	−1.513 (0.095)
Experience×Unemployed	−0.340 (0.082)	Experience^2×Unemployed	0.084 (0.015)
$k^y = 2$	0.066 (0.058)	$(k^y = 2)$ ×Experience	−0.182 (0.053)
$(k^y = 2)$ ×Experience^2	0.019 (0.011)	$(k^y = 2)$ ×Public	0.207 (0.089)
$(k^y = 2)$ ×Public×Experience	−0.063 (0.080)	$(k^y = 2)$ ×Public×Experience^2	0.013 (0.015)
$(k^y = 2)$ ×Unemployed	−0.069 (0.191)	$(k^y = 2)$ ×Unemployed×Experience	0.721 (0.184)
$(k^y = 2)$ ×Unemployed×Experience^2	−0.167 (0.038)	Constant	6.568 (0.028)
Income standard deviations: $\sigma\left(e_{it}, \mathrm{pub}_{it}, z^v_{it}, k^y_i\right)$			
Experience (years/10)	0.082 (0.016)	Public: $\mathrm{pub}_{it} = 1$	−0.055 (0.048)
Unemployed: $e_{it} = 0$	−0.577 (0.115)	$k^y = 2$	0.753 (0.044)
$(k^y = 2)$ ×Public	−0.868 (0.099)	$(k^y = 2)$ ×Unemployed	0.984 (0.137)
Constant	−2.055 (0.046)		

Table C2: Parameters of cross-sectional income means and standard deviations

First-order income autocorrelation: $\tau_1\left(e_{it}, \mathrm{pub}_{it}, e_{i,t-1}, \mathrm{pub}_{i,t-1}, z^v_{it}, k^y_i\right)$			
$(k^y = 1)$ ×Experience	0.170 (0.021)	$(k^y = 1)$ ×Public	1.106 (0.119)
$(k^y = 1)$ ×Unemployed	−0.245 (0.220)	$(k^y = 1)$ ×(Public last period)	−0.149 (0.116)
$(k^y = 1)$ ×(Unemployed last period)	−2.450 (0.197)	$k^y = 1$	2.817 (0.061)
$(k^y = 2)$ ×Experience	0.083 (0.024)	$(k^y = 2)$ ×Public	0.418 (0.104)
$(k^y = 2)$ ×Unemployed	−0.362 (0.106)	$(k^y = 2)$ ×(Public last period)	0.403 (0.093)
$(k^y = 2)$ ×(Unemployed last period)	−0.462 (0.111)	$k^y = 2$	0.900 (0.066)
Second-order income autocorrelation: $\widetilde{\tau}_2\left(k^y_i\right)$			
$k^y = 1$	0.517 (0.029)	$k^y = 2$	0.404 (0.031)

Table C3: Parameters of income mobility

Mobility heterogeneity: $\Pr\left\{k_i^m = 2 \mid z_i^f\right\}$			
Experience (years/10)	-0.072 (0.011)	High Education	-1.023 (0.286)
Medium education	-1.127 (0.321)	Constant	1.774 (0.291)

Mobility heterogeneity: $\Pr\left\{k_i^m = 3 \mid z_i^f\right\}$			
Experience (years/10)	-0.020 (0.006)	High Education	-0.914 (0.140)
Medium education	-0.472 (0.143)	Constant	2.302 (0.181)

Income heterogeneity: $\Pr\left\{k_i^y = 2 \mid k_i^m, z_i^f\right\}$			
Experience (years/10)	0.012 (0.004)	High Education	-0.624 (0.127)
Medium education	-0.225 (0.118)	$k^m = 2$	1.371 (0.297)
$k^m = 3$	-0.292 (0.151)	Constant	-0.382 (0.188)

Table C4: Parameters of unobserved heterogeneity (multinomial logit models)

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.
$k^m = 1$	15.45	12.17	87.58	0.25
$k^m = 2$	14.41	49.28	11.95	38.77
$k^m = 3$	70.14	97.46	1.34	1.19

Class composition of sectors				
	% of sample	% $k^m = 1$	% $k^m = 2$	% $k^m = 3$
Private	77.34	2.43	9.18	88.39
Public	16.20	83.55	10.63	5.82
Unemployment	6.46	0.60	86.47	12.93

Table 3: Mobility Classes

	$k^m = 1$			$k^m = 2$		
state at t	state at $t+1$			state at $t+1$		
$\downarrow$	private	public	unempl.	private	public	unemp.
private	54.0	45.7	0.3	84.1	5.1	10.8
public	3.3	96.5	0.2	30.4	60.7	8.9
unempl.	35.2	61.5	3.3	36.2	5.5	58.3
	$k^m = 3$			Whole sample		
private	99.0	0.4	0.6	89.9	8.0	2.1
public	87.5	11.9	0.6	66.3	32.0	1.7
unempl.	92.0	0.7	7.3	75.2	10.8	14.0

Table 4: Labor market state transition probabilities

Mobility classes							
	% high ed.	% med. ed.	% low ed.	Mean exp.	% $k^y = 1$	% $k^y = 2$	
$k^m = 1$	35.24	23.79	40.97	21.39	59.09	40.91	
$k^m = 2$	29.27	14.53	56.21	14.70	28.07	71.93	
$k^m = 3$	22.48	22.20	55.32	20.21	64.28	35.72	
Income classes							
	% high ed.	% med. ed.	% low ed.	Mean exp.	% $k^m = 1$	% $k^m = 2$	% $k^m = 3$
$k^y = 1$	29.43	22.34	48.24	19.09	15.67	6.94	77.38
$k^y = 2$	19.85	19.95	60.20	20.30	15.14	24.83	60.02
Whole sample							
	% high ed.	% med. ed.	% low ed.	Mean exp.			
	25.43	21.34	53.23	19.60			

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	78.2	19.3	1.7	0.3	0.5	61.3	19.5	9.9	5.4	3.9
	14.7	64.2	19.3	1.8	0.0	25.9	35.2	19.0	13.4	6.5
	1.5	20.3	57.8	19.6	0.8	15.0	21.7	32.2	21.0	10.1
	0.1	1.7	20.2	63.8	14.2	11.5	12.7	22.6	34.3	18.8
	0.2	0.0	0.8	13.7	85.3	8.9	7.7	10.6	21.2	51.6

	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

	Predicted			Observed				
state at $t-1$	state at t			state at t				
↓	private	public	unempl.	private	public	unempl.		
private	96.6	1.5	1.7	96.9	1.5	1.5		
public	8.1	90.9	0.9	8.2	90.1	0.8	Euclidian dist.:	2.4
unempl.	39.5	3.8	56.6	41.0	4.1	54.9	Maximum dist.:	1.8
state at $t-2$								
↓								
private	95.2	2.4	2.3	96.7	1.7	1.4		
public	11.9	86.7	1.3	11.0	88.3	0.6	Euclidian dist.:	2.9
unempl.	54.1	5.8	39.9	53.1	5.4	40.7	Maximum dist.:	1.6
⋮								
state at $t-6$								
↓								
private	93.8	3.5	2.6	95.3	2.9	1.7		
public	19.4	78.9	1.6	20.7	78.8	0.3	Euclidian dist.:	3.2
Unempl.	67.8	8.7	23.4	69.2	7.6	23.0	Maximum dist.:	1.5

Out of sample prediction

	Predicted			Observed				
state at $t-10$	state at t			state at t				
↓	private	public	unempl.	private	public	unempl.		
private	94.3	3.2	2.3	94.1	4.0	1.7		
public	23.4	76.0	0.5	29.7	69.4	0.8	Euclidian dist.:	15.7
unempl.	68.1	11.5	20.2	75.3	14.4	10.1	Maximum dist.:	10.1

Table 7: Fit to job mobility data

	Predicted: income quintile at t					Observed: income quintile at t					
	63.4	20.7	7.8	4.9	2.9	68.2	19.4	6.5	3.3	2.5	Eucl. dist.:
income	21.4	48.8	20.6	5.8	3.2	19.1	52.8	19.2	6.3	2.6	10.9
quintile	7.9	21.0	45.5	20.7	4.6	6.2	20.8	48.9	20.1	4.0	
at $t-1$	4.8	6.0	20.8	50.3	17.8	3.8	5.2	21.0	54.4	15.7	Max. dist.:
	2.6	3.3	5.0	17.9	70.9	2.7	2.3	3.7	15.9	75.4	4.8
	57.7	22.4	9.2	6.2	4.1	63.5	20.2	8.9	4.1	3.1	Eucl. dist.:
income	23.2	43.9	21.8	7.1	3.7	21.7	47.1	20.3	8.1	2.5	11.3
quintile	9.5	22.5	40.1	22.0	5.7	8.6	23.3	42.3	20.3	5.3	
at $t-2$	6.1	7.2	22.5	44.8	19.3	4.8	7.2	22.4	48.4	17.0	Max. dist.:
	3.8	4.0	6.1	19.5	66.4	3.3	2.6	5.4	16.9	71.6	5.8
⋮											
	44.5	24.1	13.4	9.6	8.2	52.3	22.6	13.0	6.3	5.5	Eucl. dist.:
income	24.5	33.0	23.6	12.4	6.1	28.2	34.2	19.9	9.9	7.6	14.7
quintile	13.6	23.8	29.3	23.0	10.2	11.8	27.5	30.7	20.4	9.4	
at $t-6$	10.0	12.4	23.4	32.2	21.8	7.1	12.8	25.4	37.4	17.0	Max. dist.:
	7.8	6.7	10.1	22.3	52.9	6.6	5.7	7.6	22.5	57.4	7.9

Out of sample prediction

	Predicted: income quintile at t					Observed: income quintile at t					
	39.7	25.6	14.9	11.8	7.7	38.9	27.5	16.9	11.0	5.5	Eucl. dist.:
income	22.9	31.7	29.1	9.5	6.6	28.4	31.1	21.3	11.1	8.0	22.0
quintile	16.7	23.1	23.5	23.4	13.1	14.4	26.6	31.4	18.7	8.7	
at $t-10$	12.1	13.2	18.7	31.5	24.2	5.9	11.0	20.0	39.1	23.8	Max. dist.:
	8.3	6.3	13.7	23.5	47.9	7.8	5.8	9.7	21.4	55.1	7.9

Note: Income quintiles from the unconditional sample distribution.

Table 8: Fit to income mobility data

	Private sector			**Public sector**		
	permanent	transitory	total	permanent	transitory	total
Income	0.07 22.8%	0.22 77.2%	0.29	0.05 27.8%	0.14 72.2%	0.19
Lifetime value (log.)	0.06 35.2%	0.11 64.9%	0.17	0.04 28.2%	0.11 71.8%	0.15

Note: All numbers relate to the whole sample.

Table 9: Variance decomposition: Incomes and lifetime values

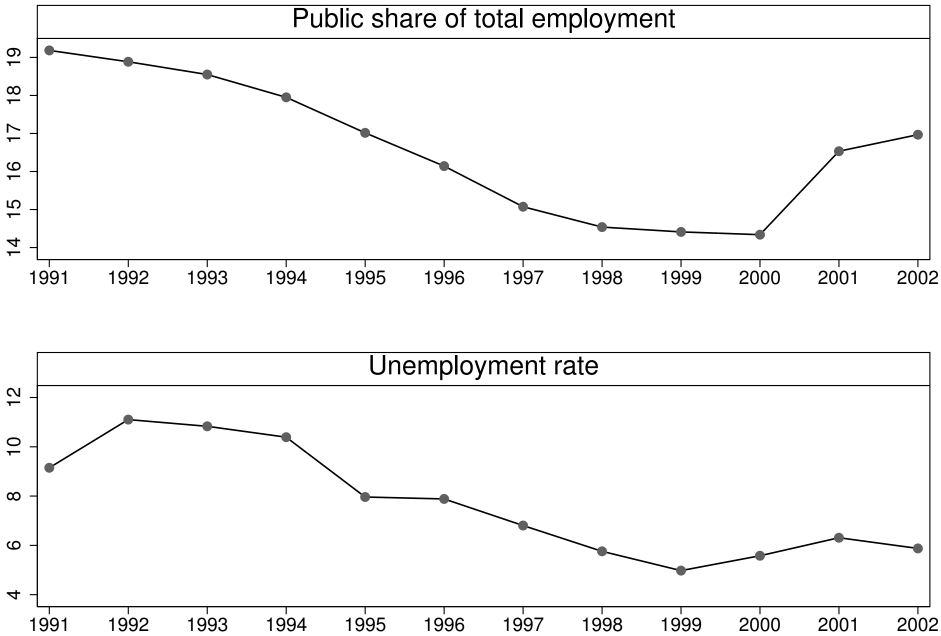

Figure 1: Sector shares and unemployment rate, males 20-55, 1991-2002

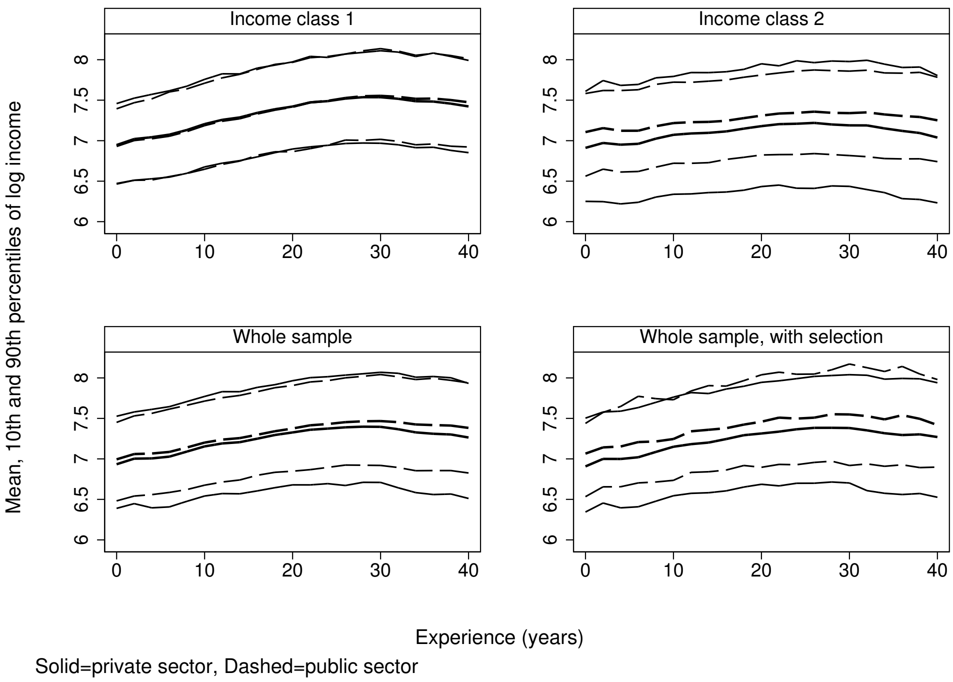

Figure 2: Income-experience profiles

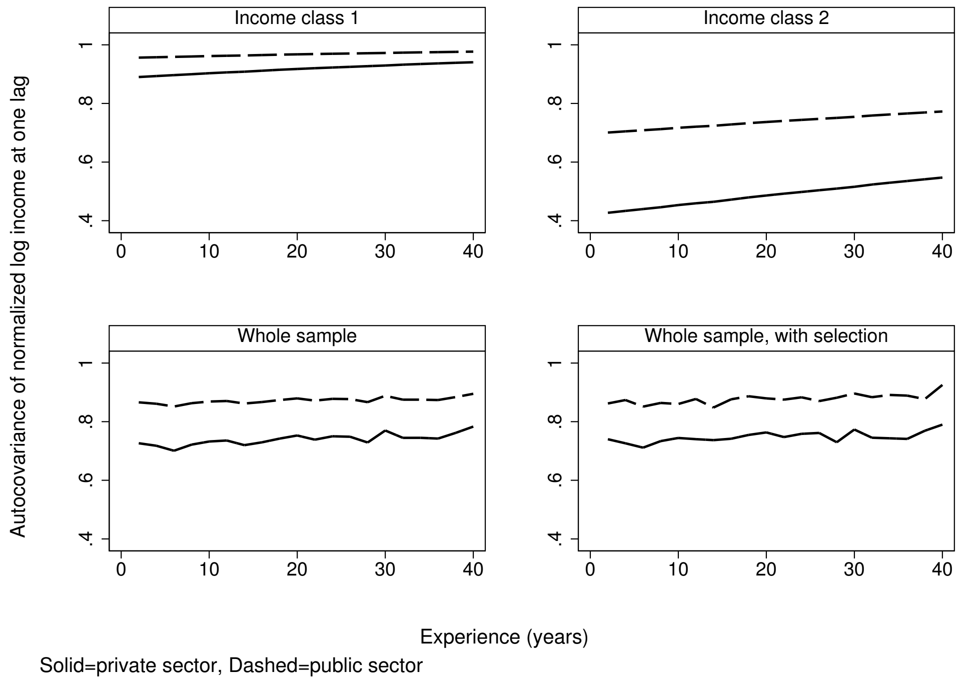

Figure 3: Autocovariance of normalized income

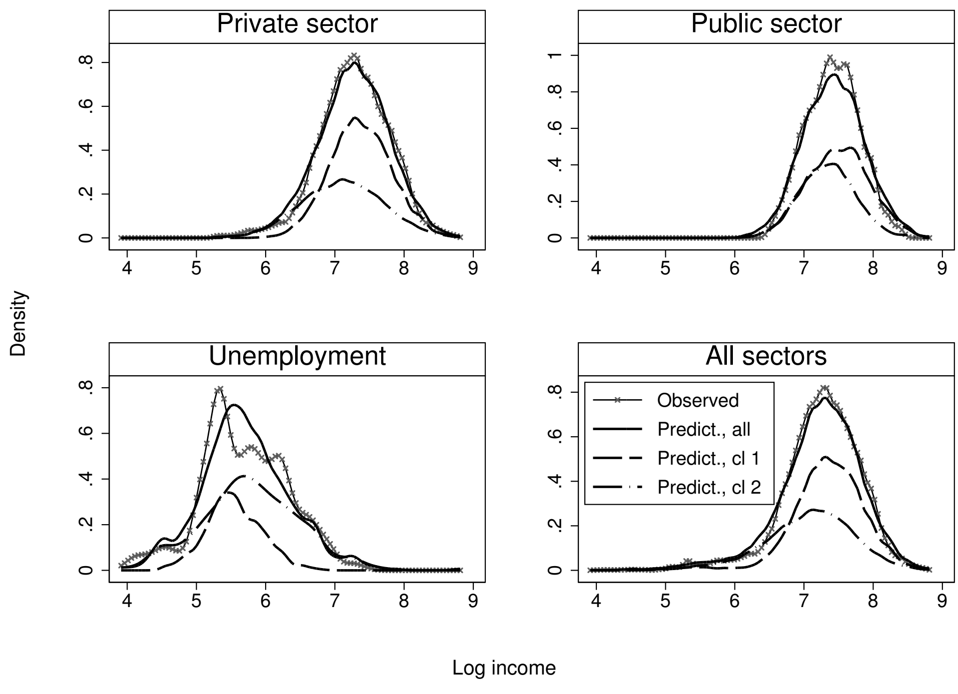

Figure 4: Income densities

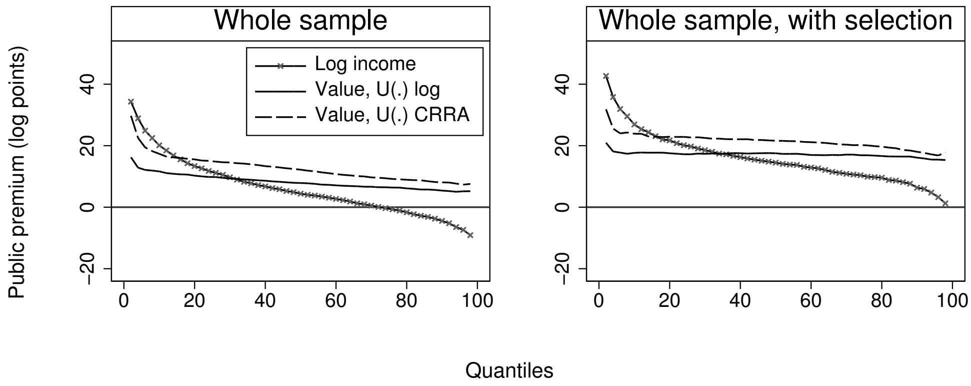

Figure 5: The public gap: jobs for life

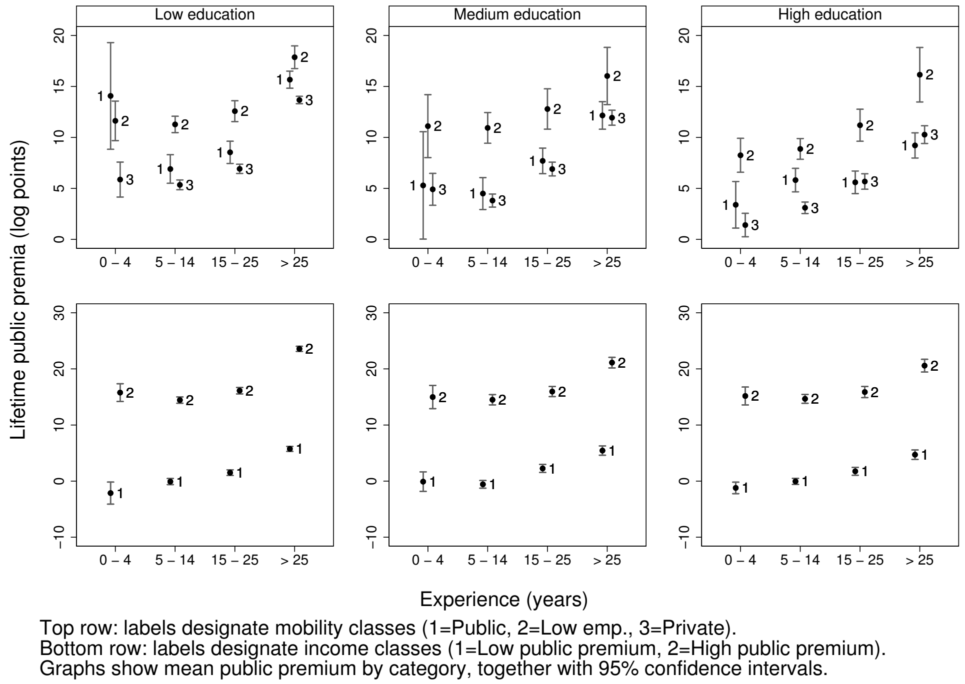

Figure 6: The conditional public gap: jobs for life

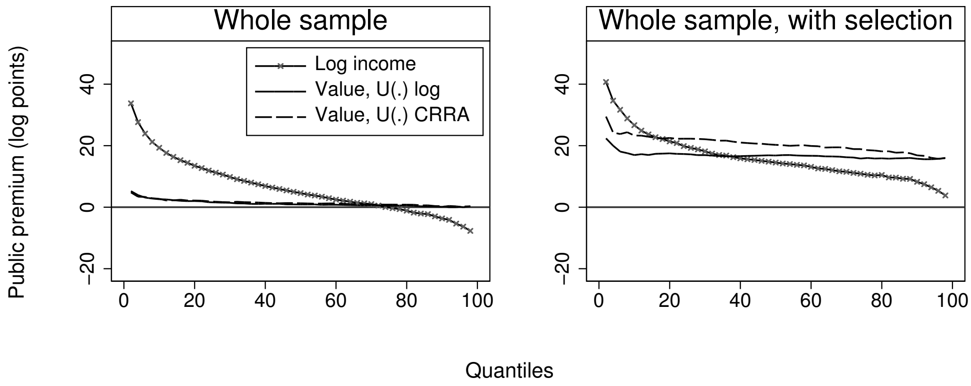

Figure 7: The public gap: income flows and lifetime values