

n° 2005-28

The Pervasive Absence of Compensating Differentials*

S. BONHOMME¹
G. JOLIVET²

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

* Repeated conversations with Jean-Marc Robin led to significant improvements in this paper. We Would also like to thank Manuel Arellano, Pierre Cahuc, Laurent Gobillon, Cédric Houdré, Francis Kramarz, Guy Laroque, Costas Meghir, Claudio Michelacci and Gerard van den Berg as well as participants from the COST Conference in Uppsala, the 2005 IZA Summer School, the 12th Panel Data conference in Copenhagen and seminar participants at CEMFI, CREST, Paris-Jourdan, Université de Paris I and Tinbergen Institute for their comments and remarks. Abba Safir had a pervasive influence on the title. The usual disclaimer applies.

¹ CEMFI, Casado del Alisal, 5, 28014 Madrid, Spain. Email : bonhomme@cemfi.es

² CREST-INSEE, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France. Email : gregory.jolivet@ensae.fr

The Pervasive Absence of Compensating Differentials*

Stéphane Bonhomme[†]
CEMFI, Madrid

Grégory Jolivet[‡]
Université de Paris I,
CREST-INSEE

November, 2005

*Repeated conversations with Jean-Marc Robin led to significant improvements in this paper. We would also like to thank Manuel Arellano, Pierre Cahuc, Laurent Gobillon, Cédric Houdré, Francis Kramarz, Guy Laroque, Costas Meghir, Claudio Michelacci and Gerard Van den Berg as well as participants from the COST Conference in Uppsala, the 2005 IZA Summer School, the 12th Panel Data conference in Copenhagen and seminar participants at CEMFI, CREST, Paris-Jourdan, Université de Paris I and Tinbergen Institute for their comments and remarks. Abla Safir had a pervasive influence on the title. The usual disclaimer applies.

[†]CEMFI, Casado del Alisal, 5, 28014 Madrid, Spain. E-mail: bonhomme@cemfi.es

[‡]CREST-INSEE, 15 bd Gabriel Péri, 92 245 Malakoff Cedex, France. E-mail: gregory.jolivet@ensae.fr

Abstract

We write and estimate a dynamic model of wages, amenities and labor mobility. Workers trade off wage and amenity offers when deciding whether to change jobs, while facing heterogeneous mobility costs. We show that these frictions can turn strong individual preferences for non wage characteristics into very small wage/amenity correlation in cross section. We use voluntary job-to-job transitions to identify workers' Marginal Willingness to Pay for amenities. The resulting selection model is solved using alternative types of constrained transitions in order to proxy the distribution of job offers. We take the model to a panel of nine European countries while controlling for unobserved heterogeneity. Our estimates show large MWP, more than 20% of the wage, for amenities such as the type of work or job security. We also find strong evidence of heterogeneity in mobility costs, and consequently weak wage/amenity in cross-section.

JEL codes: C33-35, J31-33, J63 and J81.

Keywords: Compensating differentials, hedonic wages, job mobility, amenities, unobserved heterogeneity.

Résumé

Nous estimons un modèle dynamique des salaires, aménités et transitions individuelles sur le marché du travail. Les changements d'emploi sont fondés sur un arbitrage entre salaires et aménités, conditionnellement à des coûts hétérogènes de mobilité. Nous montrons que ces frictions sont responsables d'une faible corrélation salaire/aménité en coupe et ce alors même que les individus peuvent avoir de fortes préférences pour les aménités. Les propensions marginales à payer (PMP) pour chaque aménité sont estimées sur les décisions volontaires de quitter un emploi. La sélection individuelle dans les emplois est identifiée à l'aide de plusieurs types de transitions contraintes, servant à appréhender la distribution d'offres. Le modèle est estimé sur neuf pays européens, à partir des données de l'ECHP, tout en contrôlant de l'hétérogénéité inobservée. Nos résultats témoignent d'une forte préférence pour les aménités telles que le type de travail ou la sécurité en emploi (les PMP pouvant dépasser 20%). Nous révélons également une forte hétérogénéité dans les coûts de mobilité, conduisant à de très faibles différences compensatrices de salaires en coupe.

JEL codes: C33-35, J31-33, J63 and J81.

Mots-clés: Différences compensatrices, salaires hédoniques, mobilité inter emploi, aménités, hétérogénéité inobservée.

1 Introduction

In theory, the inclusion of non wage job characteristics into labor market models potentially enriches the analysis of wage determinants, inequalities, workers' behavior or firm's technology (see Rosen, 1986). However some of these insights lack empirical support as their confrontation with the data has not led to clear conclusions. A prominent example is the mixed and sometimes conflicting estimates of workers' preferences for these *amenities* reported in the literature. The purpose of this paper is to estimate workers' Marginal Willingness to Pay (MWP hereafter) for job attributes by an original method which provides an explanation for these inconclusive findings.

In a perfectly competitive labor market, there must exist positive wage differentials for disamenities (Smith, 1776). The literature on hedonic models, initiated by Rosen (1974, 1986), has provided a relevant theoretical framework for the analysis of these *compensating* differentials. In these models, perfect competition forces the wage and amenity to belong to the worker's indifference curve, the slope of which is equal to her MWP for the amenity. Then, workers' preferences for amenities can be estimated in cross-section, and Rosen (1974) proposes to perform hedonic wage regressions to estimate the implicit–hedonic–prices of amenities.¹ However, this method has not yielded strong empirical evidence of compensating differentials. Typical estimates in this literature, starting with Thaler and Rosen (1975), are of small order of magnitude, often less than five percent of the wage. Moreover, many correlation estimates are insignificantly different from zero, possibly wrong-signed, even when controlling for individual heterogeneity (two influential studies are Brown, 1980, and Duncan and Holmlund, 1983).

An explanation for these rather inconclusive results could lie in market frictions. In an imperfect labor market, hedonic prices and workers' MWP need not coincide. Therefore, low wage/amenity correlations must not be interpreted as reflecting weak preferences for job attributes. Hwang et al. (1998) emphasize this insight in the context of an on-the-job search model with heterogeneous firms. Lang and Majumdar (2004) reach the same conclusion within a non sequential search framework where firms and workers are homogeneous. This

¹Rosen's complete proposal involves a two-step method, in which workers' preferences and firms' technology are estimated in a second step. Full identification of general equilibrium hedonic models under perfect competition faces many obstacles (Brown and Rosen, 1982). This task has been recently tackled by Ekeland et al. (2002, 2004) and Heckman et al. (2005).

concern has triggered a new empirical approach to estimate workers' MWP, which has led to strikingly different results. In an innovative study, Gronberg and Reed (1994) derive a simple relation between workers' MWP and job hazard rates. Using job duration data, they estimate high and significant MWP for two non wage attributes—measuring several aspects of working conditions—out of four.² Subsequently, Van Ommeren et al. (2000) and Dale-Olsen (2005) obtain MWP estimates ranging around one third of the wage, for commuting and safety respectively. Moreover, hedonic regression estimates based on the same data yield much smaller wage/amenity correlations. Labor market imperfections are a likely explanation for this striking contrast, but, so far, there has been no attempt at explaining these differences.

In this paper, we account for these contrasted results in a dynamic model of wages, amenities and individual labor market transitions. Job-to-job transitions and their wage and amenity outcomes play a key role in our approach. Workers value jobs according to their wage and non wage characteristics, and are constrained on their mobility as they face heterogeneous transition costs. Therefore, in the model, wage and amenity cross-sections are the result of constrained individual choices, and may imperfectly reveal workers' preferences. When taking their job change decisions, workers trade off wage and amenity offers according to their MWP for the various non wage characteristics. We show that this behavior tends to create a negative wage/amenity correlation posterior to job change. However, workers' MWP do not translate one-to-one in cross-section. The link between workers' preferences and wage/amenity correlation is made explicit in a structural relation which emphasizes the role of heterogeneity in mobility costs. This relation also involves the correlation between wage and amenity offers, generated at the firm's level. In the model, this demand-side effect is exogenous. Estimating a general equilibrium model is out of the scope of this paper.

The model rationalizes the evidence gathered by the literature. Even if workers value job characteristics significantly, there can be no “compensating” wage differences between jobs with distinct amenities. MWP can be high, so that better amenities increase the probability to accept a job offer. However, this can be insufficient for disamenities to be rewarded on the labor market if mobility costs are heterogeneous. Then, if the demand-side effect reflecting wage/amenity correlation in job offers is not sufficiently negative, regressing wages

²An early contribution by Herzog and Schlottmann (1990) makes use of workers' mobility between industries to estimate workers' MWP for safety.

on amenities for job changers will yield a small correlation— very far from the true MWP.

The identification and estimation of the model’s parameters is challenging. The two main problems we face are mirror images of each other. First, we need mobility decisions to be as less constrained as possible, so that they reveal individual preferences. This point is easily illustrated by a counterexample: if job-to-job transitions were purely exogenous from the worker’s perspective, then they would contain no information about her preferences for job characteristics. Second, as refused job offers are not observed, we have to deal with a challenging selection problem. Again, the simple example of exogenous job change is illustrative, as it is clear that there would be no selection problem in that case.

We use different types of labor market transitions to treat these two problems. Our data allow us to isolate transitions to a “better or more suitable job”, which we assume to be voluntary and to reveal individual preferences.³ Then, to identify the wage/amenity offer distribution, we use one or several types of *constrained* transitions. Doing so, we adopt a method inspired from the treatment effects literature, where the endogeneity of a “treatment” (here voluntary mobility) is corrected for by using suitable control groups. Our approach thus depends on the exogeneity of constrained mobility. As our data do not permit to precisely discriminate among constrained transitions, we address the concern that the exogeneity assumption might be violated. In the absence of a convincing instrument for job mobility, we proceed to several changes in the definition of constrained mobility and check the robustness of the results with respect to these variations. Doing so, we expect empirical regularities to emerge across these different control groups.

We complete the picture by incorporating a bivariate unobserved heterogeneity. Past literature has argued that if workers are heterogeneous with respect to their productive characteristics then hedonic regression coefficients can be strongly biased (Hwang *et al.*, 1992). We make use of the panel dimension of our data and extend Brown’s (1980) approach. We model the wage process as depending on unobserved, “productive” *job* specific characteristics. We also incorporate a job-specific effect common to all amenities and independent of the wage. This second heterogeneity is motivated by the nature of the amenity variables we

³In a recent contribution, Villanueva (2005) uses a similar definition of voluntary mobility to derive bounds on the “market price” of non wage attributes, correcting for the endogeneity of voluntary transitions. The main difference with our work lies in the considering of workers’ selection into jobs. While Villanueva treats selection as a source of bias, we claim workers’ mobility choices to be informative on their preferences.

use, which consist of self-reported measures of satisfaction with several dimensions of the job. In line with Duncan and Holmlund (1983), we think that such indicators may suffer from substantial biases, of a more “subjective” nature.

We then estimate the model on European data from the European Community Household Panel (ECHP) for the years 1994-2001. We study nine countries (Austria, Denmark, Spain, Finland, France, Ireland, Italy, the Netherlands and Portugal) and *simultaneously* allow for five amenities. The empirical results neatly illustrate our discussion of the relationships between wages, amenities and job mobility. We find positive MWP for most job attributes, significantly so in many cases. Moreover, for several amenities such as the type of work and job security, estimates range around one third of the wage. However, both the wage and amenities account for a small share of the variance of mobility decisions. We interpret this result as evidence of heterogeneity in mobility costs. When combined with the MWP, we find very small wage differentials posterior to job change. These findings are robust to changes in the model’s assumptions.

The outline of the paper is as follows: we first present our data and compute several descriptive statistics in Section 2. In Section 3, we present the model and emphasize the link between MWP and hedonic prices. Identification and estimation are discussed in Section 4 while Section 5 is devoted to estimation results and robustness checks. Lastly, Section 6 concludes.

2 Job mobility, wages and amenities: First empirical evidence

In this section, we conduct a simple descriptive analysis of a multi-country sample of individual transitions on the labor market. This allows us to emphasize a number of salient facts about workers’ mobility, wages and non wage job characteristics that will motivate our study. First, we present the data and describe the specific variables we will use for the analysis of job mobility and amenities.

2.1 The ECHP

We use the European Community Household Panel (ECHP) for the analysis of workers' mobility decisions. The ECHP is a panel of ex-ante homogenized individual data covering 15 countries from 1994 to 2001. Each household is interviewed once a year and every individual present in the initial sample is followed over the eight waves. Each observation consists of a rich set of individual characteristics, such as age and gender, together with standard information on the present job: wage, date of start... Two groups of variables are especially relevant to our analysis: the nature of job-to-job transitions, and satisfaction variables with various non wage characteristics.

Classifying job mobility: Our approach uses job-to-job transitions to identify workers' preferences. As emphasized in the introduction, constrained job change may imperfectly reveal these preferences. To discriminate between various degrees of mobility constraints, we use a variable which presents the reason why the individual has stopped working in her previous job. The twelve possible answers are the following:

| | | | |
|---|--|----|--|
| 1 | obtained better/ more suitable job | 7 | looking after old, sick, disabled persons |
| 2 | obliged to stop by employer | 8 | partner's job required move to another place |
| 3 | end of contract/ temporary job | 9 | study, national service |
| 4 | sale/ closure of own or family business | 10 | own illness or disability |
| 5 | marriage | 11 | wanted to retire or live off private means |
| 6 | child birth/ need to look after children | 12 | other |

Every answer, except 2, 3 and 4, could be thought of as a voluntary quit since the worker has not been laid off. However we consider answers 5 to 12 (when job-to-job mobility is caused *e.g.* by a marriage or the birth of a child) as a sort of constrained mobility which may not reveal the individual's preferences over jobs. In this paper, we thus define *voluntary* mobility as the transitions from one job to a "better or more suitable" one (answer 1). All the other transitions (answers 2 to 12) are *constrained*.

It is apparent from the twelve answers that a constrained transition can correspond to very different economic events. We define *displacements* to be transitions corresponding to answers 2 to 4, and cluster all the other answers (5 to 12) into the category of *partially*

constrained mobility. Although imperfect, the disaggregation of constrained transitions into these two subcategories will allow us to test the robustness of our results.

Amenities: Among the numerous job characteristics available in the ECHP is a set of job amenities. These variables give the subjective valuation of the worker with a given aspect of her job. The typical question is:

How satisfied are you with your present job in terms of (amenity)?

and individuals use a scale from 1 (*“not satisfied at all”*) to 6 (*“fully satisfied”*) to indicate their degree of satisfaction. The question remains the same for the following job characteristics:

- TY : type of work
- CD : working conditions
- WT : working times
- DI : distance to job/ commuting
- SE : job security

For the analysis to be clearer and the estimation to be more tractable, we will cluster the answers into two levels of satisfaction: an amenity equal to 1 (answers 5 and 6) will mean that individuals are actually satisfied and 0 (answers 1 to 4) that they are either unsatisfied or neutral. This clustering is consistent with the literature following Rosen (1986) where amenities take two values: zero for “bad” jobs, one for “good” jobs.⁴

Even though the ECHP is an ex-ante harmonized panel, some variables (especially amenities) may not be available in every wave and/ or country.⁵ Therefore we restrict our analysis to countries where amenities are available and rarely missing (the non-response rate is less than 1%). In this version of the paper, we focus on Austria, (AUS), Denmark (DNK), Spain (ESP), Finland (FIN), France (FRA), Ireland (IRL), Italy (ITA), the Netherlands (NLD) and Portugal (PRT). These nine countries cover the scope of the different mobility patterns one can encounter in Europe, from rather static labor markets in some Latin countries (France, Italy and Portugal) to markets with a high turnover (Denmark).

⁴It is common practice in the analysis of subjective data to estimate ordered models, such as ordered PROBIT (see Senick, 2003, and the references therein). Still these methods often involve the arbitrary clustering of some categories (typically the lowest levels of satisfaction). We also estimated our model for “good” amenities corresponding to levels 4, 5 and 6. The results remained qualitatively similar.

⁵In particular, the ECHP only lasts three years in Germany and the United Kingdom.

Individual Characteristics: We lastly present the individual characteristics we use in the subsequent analysis: “age” and “age²” are continuous variables; “male” is a gender dummy, which equals 1 for men; “married” indicates whether the individual is married (= 1) or not; and “kid” equals 1 if the individual has children under 12. Finally, “education” is a variable taking three values, from 1 (less than second stage of secondary education) to 3 (third level education).

2.2 Sample description

We merge every two consecutive waves of the ECHP and append the seven resulting tables in order to have a sample containing an *ex-ante* and an *ex-post* situation (respectively denoted as t and $t + 1$) for every individual/ year in the survey. Thus, a worker present in the eight waves is associated with seven observations, each observation containing her job status (employment, wage,⁶ amenities, etc...) and individual characteristics (age, marital status, etc...) both at date t and $t + 1$. Therefore an individual appears up to seven times in the data and for each observation she can experience one of the following transitions:

- stay employed in the same job
- stay unemployed
- make an unemployment-to-job transition
- make a job-to-unemployment transition
- make a constrained job-to-job transition
- make a voluntary job-to-job transition

Since we do not focus on labor participation, we cluster unemployment and inactivity. Moreover we define employment as paid jobs that last more than 15 hours per week.⁷ The precise construction of our samples (one per country) is presented in Appendix A.

Table 1 shows the main descriptive statistics on our samples. The first two rows present the number of individuals and the number of actual *ex-ante*/ *ex-post* observations. The next six rows give the proportions of each type of transition (except the ones from non employment to non employment). We note that individuals tend to stay in their job. The corresponding probability ranges around two thirds in all countries, which leads to an average job duration

⁶We use the logarithm of *monthly* wages detrended on year dummies.

⁷Self-employed people are likely to differ from other workers in many ways. In particular, lower risk aversion can cause much different career profiles. In this paper we assume away this issue, and drop the self-employed from our samples.

of three years. Yet, there is more dispersion across countries in the probability of making a job-to-job transition, which ranges between four and ten percent of total transitions. In particular, voluntary job-to-job mobility is significantly more frequent in Denmark (4.2%) than in Italy (1.3%). In all cases, though, these amount to a small proportion of transitions on the labor market.

Table 1: Sample description

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|---------------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|-----------------|-----------------|-----------------|
| individuals | 4 100 | 4 010 | 7 531 | 4 430 | 7 513 | 3 760 | 7 799 | 6 492 | 6 124 |
| observations | 18 455 | 20 025 | 37 683 | 16 786 | 35 571 | 16 127 | 42 527 | 31 892 | 32 877 |
| Transitions : | | | | | | | | | |
| in % of all obs. | | | | | | | | | |
| - non emp.-to-job | 959 5.2% | 1 306 6.5% | 3 798 10.1% | 1 511 9.0% | 2 464 6.9% | 1 565 9.7% | 2 713 6.4% | 2 053 6.4% | 2 211 6.7% |
| - job-to-non emp. | 1 002 5.4% | 1 133 5.7% | 2 685 7.1% | 1 083 6.5% | 2 214 6.2% | 959 5.9% | 2 289 5.4% | 1 321 4.1% | 1 560 4.7% |
| - stay in same job | 13 382 72.5% | 13 070 65.3% | 20 349 54.0% | 11 028 65.7% | 24 233 68.1% | 9 683 60.0% | 27 889 65.6% | 21 534 67.5% | 22 624 68.8% |
| - job-to-job | 933 5.1% | 2 058 10.3% | 2 677 7.1% | 962 5.7% | 1 429 4.0% | 1 140 7.1% | 1 558 3.7% | 2 296 7.2% | 1 646 5.0% |
| - voluntary j-t-j | 436 2.4% | 849 4.2% | 754 2.0% | 419 2.5% | 603 1.7% | 544 3.4% | 541 1.3% | 988 3.1% | 672 2.0% |
| - constrained j-t-j | 497 2.7% | 1209 6.0% | 1 923 5.1% | 543 3.2% | 826 2.3% | 596 3.7% | 1 017 2.4% | 1308 4.1% | 974 3.0% |
| % of wage | | | | | | | | | |
| increases among: | | | | | | | | | |
| - vol. j-t-j | 63.3% | 61.2% | 64.5% | 69.7% | 60.5% | 68.8% | 60.4% | 73.3% | 66.1% |
| - constr. j-t-j | 54.2% | 51.4% | 54.6% | 54.3% | 53.2% | 53.9% | 47.1% | 64.8% | 51.6% |
| - stay in same job | 52.1% | 47.2% | 53.5% | 52.2% | 54.5% | 52.1% | 48.2% | 59.7% | 48.1% |

The last three rows of Table 1 are important motivations for our analysis. Indeed we can see that most voluntary job changes are associated with a wage gain whereas job stayers and constrained job movers more frequently experience a wage cut. This suggests that the wage influences job change decisions. Yet, the proportion of wage increases ranges from only 60% (in France) to 73% (in the Netherlands) of *voluntary* job-to-job transitions. Up to 40% of voluntary job movers experience a wage cut even if the new job is said to be “better or more suitable” than the previous one. If at least part of these transitions with wage cuts are not spuriously generated by measurement error, these statistics suggest that the wage is not the only characteristic workers value, and we should look at other job characteristics to explain

voluntary mobility.⁸⁹

To investigate further this issue, we compute the probability matrices for the various amenities and types of transitions (voluntary, constrained or within-job).¹⁰ To save space, we show these matrices in Table 2 for Denmark only, the results being similar in all countries. As every off-diagonals show (distance to job being the only exception), the probability matrices in the first row are all non symmetric: voluntary job changers are more prone to gain than to lose amenities. For instance, 28% (resp. 9%) of voluntary job changers experience a rise (resp. a fall) in their satisfaction with the type of work. In contrast, probability matrices are strikingly symmetric when looking at job stayers. In constrained transitions, non wage characteristics are improved in terms of satisfaction, yet less so than in voluntary ones. Thus, the increase in satisfaction with non wage characteristics seems to be specific to voluntary mobility.

Table 2: Probability matrices $\begin{pmatrix} \mathbb{P}(0,0) & \mathbb{P}(0,1) \\ \mathbb{P}(1,0) & \mathbb{P}(1,1) \end{pmatrix}$ in Denmark

| Amenities | TY | CD | WT | DI | SE |
|---------------|--|--|--|--|--|
| vol j-t-j | $\begin{pmatrix} .13 & .28 \\ .09 & .50 \end{pmatrix}$ | $\begin{pmatrix} .16 & .25 \\ .16 & .43 \end{pmatrix}$ | $\begin{pmatrix} .12 & .20 \\ .16 & .52 \end{pmatrix}$ | $\begin{pmatrix} .19 & .19 \\ .20 & .42 \end{pmatrix}$ | $\begin{pmatrix} .16 & .19 \\ .16 & .49 \end{pmatrix}$ |
| constr. j-t-j | $\begin{pmatrix} .16 & .21 \\ .18 & .45 \end{pmatrix}$ | $\begin{pmatrix} .19 & .20 \\ .17 & .44 \end{pmatrix}$ | $\begin{pmatrix} .13 & .17 \\ .17 & .53 \end{pmatrix}$ | $\begin{pmatrix} .19 & .19 \\ .20 & .42 \end{pmatrix}$ | $\begin{pmatrix} .27 & .20 \\ .17 & .36 \end{pmatrix}$ |
| stay in job | $\begin{pmatrix} .17 & .10 \\ .13 & .60 \end{pmatrix}$ | $\begin{pmatrix} .20 & .13 \\ .14 & .53 \end{pmatrix}$ | $\begin{pmatrix} .14 & .10 \\ .10 & .66 \end{pmatrix}$ | $\begin{pmatrix} .21 & .07 \\ .08 & .64 \end{pmatrix}$ | $\begin{pmatrix} .15 & .12 \\ .11 & .62 \end{pmatrix}$ |

These few descriptive statistics tend to confirm the idea that the wage is not the only determinant of workers' voluntary mobility and that non wage characteristics are likely to enter job valuation. We now proceed to a formal test of this intuition.

⁸⁹Changes in hours worked provide a possible explanation for voluntary quits associated with wage cuts. However, less than 10% of these transitions correspond to changes from full-time (defined as more than 30 hours per week) to part-time work in the three countries we consider.

⁹⁰This paper focuses on non wage characteristics. Wage *growth* expectations are an alternative explanation for job transitions associated with wage cuts (*e.g.* Postel-Vinay and Robin, 2002).

¹⁰ $\mathbb{P}(a_t, a_{t+1})$ is the joint probability law of (a_t, a_{t+1}) . It is computed as the ratio of the number of voluntary (resp. constrained or within-job) $t/t + 1$ transitions from amenity $a_t = 0, 1$ to amenity $a_{t+1} = 0, 1$, divided by the number of all voluntary (resp. constrained or within-job) transitions.

3 A model of wages, amenities, and job mobility

3.1 The model

We here write a dynamic model of individual transitions on the labor market. Voluntary job mobility is the keystone of the model. Still, we also allow for transitions into and out of non employment, together with constrained job change. In the next Section, we shall argue that these transitions can be informative to deal with the endogeneity of voluntary job mobility. We here present the model with a single non wage characteristic, for clarity. The extension to multiple amenities, which we use for estimation, is reported in the Appendix.

In the model, every match between a worker and a job is described by a pair $\theta = (\theta_1, \theta_2)$, where θ_1 is a productive characteristic and θ_2 is a non productive aspect of the match, standing for the worker's subjective satisfaction with her job. Vector x represents individual characteristics.

Wages and amenities within jobs: Let a worker be employed at date $t = 0, 1, \dots$, and let θ be the characteristics of the worker/job match. We assume that the match has already lasted for at least one period. The special case of starting jobs will be addressed at the end of this subsection together with the realization of new matches' characteristics. At the beginning of period t , the wage y_t and the binary non wage characteristic $a_t \in \{0, 1\}$ of the job are drawn, given θ and x . The amenity is drawn according to:

$$a_t = \mathbf{1}\{\alpha_a x + \beta_{1a}\theta_1 + \beta_{2a}\theta_2 + u_{at} > 0\}. \quad (1)$$

Simultaneously, the (logarithm of the) wage y_t is drawn according to the *hedonic curve*:

$$y_t = \rho a_t + \alpha_y x + \beta_y \theta_1 + u_{yt}. \quad (2)$$

It is convenient to think of equation (1) as a reduced form. The “objective” characteristics of a given job (y_t, \tilde{a}_t) are supposed to depend on common characteristics: θ_1 and x , and transitory shocks u_{yt} and u_{at} . Then, amenities are evaluated by workers in terms of satisfaction. We introduce a subjective threshold s_t such that $a_t = \mathbf{1}\{\tilde{a}_t > s_t\}$. This threshold depends on x and θ_1 , and on the second (non productive) component of the quality of the worker/job match θ_2 . Equation (1) then mixes the “objective” amenity and the “subjective” interpretation of the amenity in terms of satisfaction.

Then, parameter ρ in the hedonic curve (2) is the implicit price of the amenity, which we shall refer to as the within-job *compensating differential* for the amenity. Equations (1)-(2) can follow from the negotiation of the employment contract between the worker and the firm, given the characteristics of the match. We do not model this process and thus treat the parameters in (1)-(2) as exogenous.

Once drawn, the wage and amenity are valued by the worker, according to the instantaneous indirect utility function $v(y_t, a_t, x)$. Then she can experience several types of transitions.

Adverse shocks: First, the worker can be constrained to move to another job, or forced back to non employment. Both events happen with exogenous probabilities depending on x and θ .

If the worker has experienced a constrained transition, her new wage/amenity pair (y^c, a^c) is drawn from a specific distribution given θ and x , similar to (1)-(2) but with different parameters indexed by the $(^c)$ superscript. These parameters result from the new firm's technological choices, rather than from contract negotiation with the worker. In particular, the compensating differential for the amenity posterior to constrained job change (ρ^c) is *a priori* different from the within-job compensating differential (ρ).

Job offers: If the worker has experienced no adverse shock, then she gets an offer from an outside firm that she can either accept or turn down.¹¹ Amenity offers a^* are drawn from a specific distribution:

$$a^* = \mathbf{1}\{\alpha_a^* x + \beta_{1a}^* \theta_1 + \beta_{2a}^* \theta_2 + u_a^* > 0\}, \quad (3)$$

and wages and amenities belong to another hedonic curve:

$$y^* = \rho^* a^* + \alpha_y^* x + \beta_y^* \theta_1 + u_y^*, \quad (4)$$

where ρ^* is the compensating differential for the amenity in job offers. Equation (4) reflects the trade-off between wages and (possibly costly) amenities at the firm's level. It is also taken as given. We allow the three hedonic curves— corresponding to continuing jobs, constrained job change and outside offers —not to be the same. The three compensating differentials: ρ , ρ^c and ρ^* can thus be different.

¹¹The assumption of systematic arrival of outside job offers will be discussed in the next subsection.

Note that wage and amenity offers depend on the characteristics of the *current* match, as firms are assumed to direct their search on groups of workers with given individual (x) and match (θ) characteristics .

Voluntary mobility rule: Let $V(y_t, a_t, \theta, x)$ be the present value, at the beginning of period t , of a job with characteristics θ and (instantaneous) wage/amenity values (y_t, a_t) . Future periods are discounted at a constant rate. Note that $V(y_t, a_t, \theta, x)$ takes into account possible future spells of non employment, associated with a constant instantaneous utility.

The worker bases her decision whether or not to change job on the comparison of the job offer she has just received to the expected present value of her current job. The latter is given by: $V_{stay}(\theta, x) = \mathbb{E}_{y_{t+1}, a_{t+1}} [V(y_{t+1}, a_{t+1}, \theta, x) | \theta, x]$. Note that, as shocks u_{at} and u_{yt} in (1)-(2) are transitory, this expression does not depend on past wage/amenity values $(y_t, a_t, y_{t-1}, a_{t-1}, \dots)$.

When the worker takes her decision, the characteristics of the new match are not yet realized. The decision is thus based on the expected value of the proposed job, given wage and amenity offers: $V_{mov}(y^*, a^*, \theta, x) = \mathbb{E}_{\theta^*} [V(y^*, a^*, \theta^*, x) | y^*, a^*, \theta, x]$, where θ^* stands for the characteristics of the offered worker/job match, unknown to the worker at time t .

Let z_t be the dummy variable indicating if the individual has changed job voluntarily between periods t and $t + 1$. The mobility decision reads:

$$z_t = \mathbf{1} \{V_{move}(y^*, a^*, \theta, x) > V_{stay}(\theta, x) + c(\theta, x)\},$$

where $c(\theta, x)$ are transition costs.¹² We adopt a linear specification for this relationship, and assume:

$$z_t = \mathbf{1} \{y^* + \delta^* a^* > \alpha_z x + \beta_{1z} \theta_1 + \beta_{2z} \theta_2 + u_z\}. \quad (5)$$

Since x and θ intervene in both V_{stay} and V_{move} , the interpretation of their associated parameters in (5) calls for prudence. However, the role of y^* and a^* is unambiguous and δ^* clearly reflects the trade off between wage and amenity in the mobility decision. We shall refer to δ^* as the *dynamic* Marginal Willingness to Pay for the amenity.

¹²Examples of mobility costs correlated with θ could be housing and children's education expenses. Van den Berg (1992) estimates an on-the-job search model where mobility costs depend on the current wage, constant within job.

The Right-Hand Side inside the index in equation (5):

$$\tau(x, \theta) = \alpha_z x + \beta_{1z} \theta_1 + \beta_{2z} \theta_2 + u_z, \quad (6)$$

is a combination of the value of staying in the present job and transition costs. In this paper, we do not intend to disentangle these two dimensions. Moreover, the random shock u_z renders $\tau(\theta; x)$ stochastic conditional on observed and unobserved characteristics. Although many models of transitions assume deterministic reservation wages (see *e.g.* Flinn and Heckman, 1982), we shall see in Section 5 that, in the context of job-to-job mobility, it is important to allow for variability in τ .

The voluntary job change decision has two potential outcomes. If the worker accepts the offer, she gets $(y_{t+1}, a_{t+1}) = (y^*, a^*)$ at the beginning of the next period $t + 1$. If not, she draws a new pair (y_{t+1}, a_{t+1}) in (1)-(2).

Match characteristics: If the worker has started to work at a new job at the next period $t + 1$ with starting wage and amenities (y_{t+1}, a_{t+1}) , then the productive and non productive characteristics of the match are realized. The new characteristics θ^* are drawn from a distribution depending on the starting wage/amenity values at the job. Importantly, this distribution is identical for every individual and independent of the type of the last transition. In other words, we assume that starting wage/amenity values, which do depend on the worker's previous labor market status, are sufficient statistics for the match characteristics, and we neglect state dependence beyond these starting conditions.

If instead she has remained in the same job between t and $t + 1$, then the characteristics of the match stay the same. Match characteristics are thus constant within-job. Lastly, non employed individuals are assumed to “keep” the θ of their previous match, using θ as a signal when applying for a new job.

So far, the worker has been assumed to be employed in a continuing job at time t . We now focus on the two cases when she is either non employed at time t , or starting to work in a new job.

Non employment-to-job transitions: If the worker is non employed at the beginning of period t , then we assume that she can find a job with exogenous probability depending

on the θ of her previous job.¹³ Then, at the beginning of the next period, she draws a wage/amenity pair (y^n, a^n) from a different distribution going with another compensating differential for the amenity (ρ^n) . Subsequently, the characteristics of the new match are realized.

Starting jobs: Lastly, consider a worker employed in a new job at the beginning of period t . Her starting wage and amenity are either equal to (y^*, a^*) (if she has changed job voluntarily at the end of period $t - 1$), to (y^c, a^c) (if she has been constrained to change job), or to (y^n, a^n) (if she has was previously non employed). The new match characteristics are then drawn given these starting values, and condition the arrival rates of adverse shocks and the offers she might receive from an outside firm.

Econometric specification: In the econometric model, all variables are indexed by individual i and time period t . Individual characteristics x_{it} can be time-varying (*e.g.* age), or not (sex, education). All residuals in the model are *i.i.d.*, independent of covariates, independent of one another and normally distributed with zero means.

All types of transitions— but voluntary ones—are modeled according to a common pattern. For instance, the occurrence of constrained job change is specified as:

$$z^c = \mathbf{1} \{ \alpha_z^c x + \beta_{1z}^c \theta_1 + \beta_{2z}^c \theta_2 + u_z^c > 0 \}. \quad (7)$$

The variance of all residuals in the transition equations, such as u_z^c , are normalized to one. We adopt the same normalization for the residuals in the amenity equations. In contrast, the variance of u_z in the voluntary mobility equation can be identified, as the coefficient of wage offers in (5) is set to one by assumption.

In the econometric model, characteristics θ_1 and θ_2 are *unobserved* match-specific effects. Their identification comes from within-job repetitions (see 4.1). We model θ_1 and θ_2 as *discrete* random variables. This choice allows to approximate their distributions without making explicit distributional assumptions (Heckman and Singer, 1984).

After a transition out of non employment, a constrained or a voluntary job-to-job transition, new match characteristics are drawn. They are assumed to follow an Ordered PROBIT

¹³In the estimation, we will only consider individuals who are employed initially, so that every observation will be associated to a θ .

specification, conditional on starting wage and amenities at the new job.

We adopt a similar approach to model initial conditions. Namely, we assume that all individuals are initially employed, and condition the worker’s likelihood on her starting wage and amenity. Then, the match characteristics in the initial job follow an Ordered PROBIT specification given the initial wage and amenity, with a new set of parameters.

3.2 Hedonic wage regressions and job mobility

The wage/amenity offer equations (3)-(4) and the rule of voluntary mobility (5) form the core of the model. The two parameters ρ^* and δ^* represent specific trade-offs between wages and amenities: at the firm’s level, ρ^* reflects substitution between two costs while at the worker’s level, δ^* represents the trade-off between two goods. When workers take their mobility decisions, both effects are at play. Posterior to job change, these “demand” and “supply” effects can *a priori* strengthen or weaken the correlation between wages and amenities.

One of the main insights of the theory of compensating differentials is that the two trade-offs, at the firm’s and the worker’s levels, tend to create a *negative* correlation.¹⁴ In the model, this correlation depends on parameters ρ^* and δ^* , as well as on a third parameter measuring the extent of transition costs. We now derive the structural relation linking these different parameters.

Let us define:

$$\Delta_z = \mathbb{E}(y^*|a^* = 1, z_t = 1, \theta, x) - \mathbb{E}(y^*|a^* = 0, z_t = 1, \theta, x).$$

Δ_z is the wage differential between the two amenity levels, for voluntary job changers. It could be estimated by a hedonic wage regression on the subsample of job changers, controlling for unobserved heterogeneity.

Using the specification of the hedonic curve (4), Δ_z can be rewritten as:

$$\Delta_z = \underbrace{\rho^*}_{\equiv \Delta_z^{(d)}} + \underbrace{\mathbb{E}(u_y^*|a^* = 1, z_t = 1, \theta, x) - \mathbb{E}(u_y^*|a^* = 0, z_t = 1, \theta, x)}_{\equiv \Delta_z^{(s)}}. \quad (8)$$

In (8), Δ_z has been written as the combination of a “demand” and a “supply” effect. The first effect, ρ^* , is taken as given in the model.

¹⁴Other theories of non wage characteristics have different implications concerning the wage/amenity correlation in equilibrium. See *e.g.* Hwang *et al.* (1998).

Then, the “supply” effect is generally negative, consistently with the intuition that workers can accept lower (higher) wage in exchange of better (worse) amenity when moving to another job. To see why, note that the mobility rule (5) implies that

$$\Delta_z^{(s)} = \mathbb{E}(u_y^* | u_z - u_y^* < -\mu + \rho^* + \delta^*) - \mathbb{E}(u_y^* | u_z - u_y^* < -\mu),$$

where we have defined for compactness:

$$\mu = (\alpha_z - \alpha_y^*)x - (\beta_y^* - \beta_{1z})\theta_1 + \beta_{2z}\theta_2.$$

As u_y^* and u_z are normally distributed and uncorrelated by assumption, we can write:

$$\Delta_z^{(s)} = \frac{(\sigma_y^*)^2}{\sigma} \left[\frac{\phi}{\Phi} \left(-\frac{\mu - \delta^* - \rho^*}{\sigma} \right) - \frac{\phi}{\Phi} \left(-\frac{\mu}{\sigma} \right) \right], \quad (9)$$

where ϕ (respectively Φ) denotes the standard normal pdf (resp. cdf), and $(\sigma_y^*)^2$, σ_z^2 and $\sigma^2 = (\sigma_y^*)^2 + \sigma_z^2$ are the variances of u_y^* , u_z and $u_y^* - u_z$, respectively.

Then it follows from (9) that $\Delta_z^{(s)}$ has the same sign as $-(\delta^* + \rho^*)$.¹⁵ In the benchmark case when there is no correlation in wage/amenity offers ($\rho^* = 0$), the worker’s trade-off between wage and amenity leads to a negative wage/amenity correlation posterior to job change. Even in the absence of negative correlation on the demand-side of the market, this supply effect is enough to create “compensating” wage differences between jobs with distinct amenities.

Order of magnitude: In the model, workers are constrained on their mobility, as we assume job-to-job transitions to be costly. This feature can have a strong effect on the wage differential posterior to job change Δ_z , as we now illustrate.

To proceed, let us suppose that the probability of job change is small, conditional on “good” or “bad” amenity offers, and conditional on individual and job characteristics. This assumption is consistent with the descriptive evidence in 2.2, where the aggregate probability of voluntary job change was found to be less than five per cent in the three countries we study. Precisely, for all x , θ and a^* , we suppose:

$$\Phi \left(-\frac{\mu - (\delta^* + \rho^*)a^*}{\sigma} \right) \ll 1.$$

¹⁵This is because the inverse Mills ratio $\frac{\phi}{\Phi}$ is strictly decreasing on the real line. The normality of u_z and u_y^* is not essential for this result to hold, however.

This assumption permits to approximate $\Delta_z^{(s)}$ as:¹⁶

$$\Delta_z^{(s)} \approx \frac{(\sigma_y^*)^2}{\sigma} \left[\left(\frac{\mu - \delta^* - \rho^*}{\sigma} \right) - \left(\frac{\mu}{\sigma} \right) \right] = - \left(\frac{\sigma_y^*}{\sigma} \right)^2 (\delta^* + \rho^*).$$

We then define:

$$R_{yz} \equiv \left(\frac{\sigma_y^*}{\sigma} \right)^2. \quad (10)$$

Hence $\Delta_z^{(s)}$ can be approximated as:

$$\Delta_z^{(s)} \approx -R_{yz} (\delta^* + \rho^*). \quad (11)$$

R_{yz} measures the weight of wage offers in mobility decisions. We interpret this ratio as a measure of the heterogeneity in mobility costs. If costs are homogeneous, then the variance of wage offers accounts for a large share of σ^2 and R_{yz} is close to 1. On the contrary, if mobility decisions involve many factors other than the wage, R_{yz} is small.

We do not model the probability of receiving an outside offer, since we assume that every worker who has not faced an adverse shock gets an alternative offer. However, the *stochastic* mobility costs we have introduced make this assumption immaterial. In the model, not having access to an alternative job (*i.e.* receiving no job offer) is interpreted as facing a high mobility cost, *i.e.* drawing a large u_z . It is thus clear that R_{yz} represents also the heterogeneity in opportunities to change job.

Equation (11) shows that, the larger the MWP for the amenity, the more negative the correlation posterior to job change. However, in the case where mobility costs are heterogeneous, large MWP for job attributes (*i.e.* large and positive δ^*) can translate into very weak wage/amenity correlation (if R_{yz} is low).

Now if we go back to the wage differential for job changers and still assume a low probability of job change, we can write, combining (8) and (11):

$$\Delta_z \approx (1 - R_{yz}) \rho^* + R_{yz} (-\delta^*). \quad (12)$$

According to (12), the wage/amenity correlation posterior to job change combines the two parameters ρ^* and δ^* , in a proportion determined by the heterogeneity in mobility costs

¹⁶The inverse Mills ratio has the property that $\frac{\phi}{\Phi}(x) = -x + o(x)$ when $x \rightarrow -\infty$. Hence, for probability p close to zero, $\frac{\phi}{\Phi}(\Phi^{-1}(p))$ is close to $-\Phi^{-1}(p)$. Note that, unlike in the previous paragraph, the assumption of normally distributed residuals is here critical.

R_{yz} . In the limit, when R_{yz} is close to zero, then the worker’s trade-off between wages and amenities has no impact on the correlation posterior to job change. Lastly, equation (12) involves the three key parameters of the model: wage/amenity substitution on the demand (ρ^*) and supply (δ^*) sides, together with an indicator of mobility costs (R_{yz}). Identifying and estimating these parameters is one of our main purposes.

4 Identification and estimation issues

In this section, we address the identification and estimation of the model’s parameters. We first focus on the identification problem arising from workers’ selection into jobs. We then discuss the identification of unobserved match characteristics. Lastly, we outline the estimation method.

4.1 Identification of the key parameters

We here assume that match characteristics are observed by the econometrician. Next subsection will deal with their identification.

If job offers were observed for all workers, not only actual job changers, then the model’s key parameters would be identified without further assumption. Indeed ρ^* follows directly from the hedonic curve (4). Moreover, all parameters ruling job offer equations would be identified in that case, including σ_y^* , the standard deviation of u_y^* . As for δ^* , it follows from equation (5) that:

$$\Phi^{-1} [\mathbb{P}(z_t = 1 | y^*, a^*, \theta, x)] = \frac{1}{\sigma_z} y^* + \frac{\delta^*}{\sigma_z} a^* - \frac{\alpha_z}{\sigma_z} x - \frac{\beta_{1z}}{\sigma_z} \theta_1 - \frac{\beta_{2z}}{\sigma_z} \theta_2. \quad (13)$$

Equation (13) shows that, along with data on job turnover, data on wage offers would permit to identify all parameters in (5), including the standard deviation of u_z , σ_z .

The selection problem: The absence of data on job offers in the ECHP complicates the researcher’s task. As wage and amenity offers are not observed for job stayers, we face a selection problem.

Identification of selection models is often achieved by the use of exclusion restrictions. In the case of job-to-job mobility, however, it seems difficult to find a covariate that both

significantly influences the job change probability and is uncorrelated with job offers. Potential candidates in the ECHP are the indicators of being married and having young children. However, their effect on job change propensity turns out to be small.

Therefore we here take a different route. Our approach builds on the remark that observing the realizations of job offers (y^*, a^*) is not strictly necessary for the parameters to be identified. For this purpose, knowing their *distribution* is sufficient. This is straightforward for the parameters appearing in (3)-(4). For mobility parameters in (5), the argument makes use of Bayes' rule. Namely, if the wage/amenity offer distribution is known (given characteristics), with density ℓ^* , then:

$$\mathbb{P}(z_t = 1 | y^*, a^*, \theta, x) = \frac{\ell^z(y^*, a^* | z_t = 1, \theta, x)}{\ell^*(y^*, a^* | \theta, x)} \mathbb{P}(z_t = 1 | \theta, x), \quad (14)$$

where ℓ^z denotes the density of wages and amenities accepted by voluntary job changers. As ℓ^z and $\mathbb{P}(z_t = 1 | \theta, x)$ involve observed quantities, the Left-Hand Side in (14) is identified as soon as ℓ^* is known. Then, using (13), the parameters appearing in the mobility equation can be recovered.

Finding proxies for the distribution of wage/amenity offers: For the model's parameters to be identified, a sufficient condition is thus that the distribution of job offers be itself identified. We propose to augment the model by the following identity:

$$\ell^* = \ell^c, \quad (15)$$

where ℓ^c is the density of wages and amenities drawn by *constrained* job changers.

For (15) to be satisfied, two conditions need to hold: First, constrained job change has to be exogenous from the worker's perspective. If this is the case, then ℓ^c can also be interpreted as the density of "offers" received by constrained job changers— offers which cannot be turned down by the worker. Second, constrained and voluntary job changers need to draw from the same distribution of job offers.

The first condition, exogeneity, could be violated for two reasons: workers can self-select with respect to their unobserved characteristics (ability bias) or with respect to the characteristics of the current and offered jobs (selection). We may argue that we control for the first source of endogeneity, as the probability of job change is conditional on θ in

equation (7). However, we do not control for the endogeneity possibly arising from selection with respect to job offers. As for the second condition, equality of job offers, it could be violated if, say, experiencing a constrained transition were seen by employers as a signal of low productivity. This argument could well hold in the case of lay-offs, as suggested in Gibbons and Katz (1991).

To guarantee the exogeneity of constrained job change, Gibbons and Katz (1992) consider displaced workers consecutive to plant closure. They claim that such displacements approximate the natural experiment of exogenous job loss, and use this insight to correct for ability bias in inter-industry wage differentials. A recent empirical analysis of returns to tenure by Dustmann and Meghir (2005) builds on the same idea. In our case, however, this approach is not directly applicable because the ECHP data are not precise enough to identify “true” displacements, exogenous from workers’ perspective. We are not aware of a data set providing information on amenities and voluntary mobility (as in the ECHP) together with “true” displacements (as in firm level data). Still, the ECHP data allow to distinguish between different types of constrained job transitions. This feature permits to develop informal robustness checks.

An informal test of robustness. To address the concern that the above conditions might be violated, we test the robustness of our results to variations in assumption (15). Our approach consists in substituting three other wage/amenity distributions for the original distribution of reference (constrained transitions). First, we decompose constrained transitions into displacements and partially-constrained transitions, as explained in 2.1. We take the wage/amenity distributions of either of the two types as a reference. For instance, replacing identity (15) by: $\ell^* = \ell^d$, where ℓ^d denotes the density of wages and amenities drawn by *displaced* workers, a different set of parameters ruling voluntary job mobility can be derived. We proceed similarly in the case of partially constrained transitions.

Lastly, we also try a third specification, taking wages and amenities posterior to non employment as the reference for job offers. In this case, the identity becomes: $\ell^* = \ell^n$. This restriction has been widely used in the job search literature, as an assumption permitting the job offer distribution to be identified. In these models, a subpopulation of workers (usually unemployed) draw their jobs in the offer distribution and are forced to accept the offer

because their alternative is not preferable. Hence the observed distribution of jobs drawn by these workers is the same as the actual offer distribution (see *e.g.* Christensen et al., 2005).

None of these distributions of reference is claimed to ensure the two conditions stated above but we expect empirical regularities with respect to these changes in “control” groups to be relevant. Indeed, we will see in section 5.3 that the resulting patterns, in terms of MWP and wage/amenity correlation, turn out to be remarkably similar between specifications, reinforcing the evidence.

4.2 Identification of match characteristics

In the model, match characteristics are unobserved by the econometrician. As θ_1 and θ_2 are constant within-job, however, they are identified by wage and amenity repetitions provided that jobs last more than two periods with positive probability. For θ_1 , the argument comes *e.g.* from a theorem by Kotlarski (1967). In that case, the density of θ_1 is identified non-parametrically. As for θ_2 , one also needs that the regressors vary sufficiently over time. For instance, if one of the regressors has large support in all its dimensions as in Manski (1988) then the latent distributions are identified and Kotlarski’s result applies.

Properly speaking, we do not dispose of such a regressor in amenity equation (1), but unobserved heterogeneity distributions are modeled as discrete random variables, allowing for a parsimonious number of groups. For this specification, we found no evidence of identification problems.

4.3 Estimation: EM with a Sequential M-step (ESM)

We here briefly present the estimation of the model’s parameters. The details of the procedure are given in Appendix B. We model θ_1 and θ_2 as follows. Let N denote the number of individuals in the sample. We assume that there exist two integers K_1 and K_2 , a mapping:

$$\begin{aligned} \{1\dots N\} &\rightarrow \{1\dots K_1\} \times \{1\dots K_2\} \\ i &\mapsto (k_{1i}, k_{2i}), \end{aligned}$$

and parameters $(\vartheta_{11}, \dots, \vartheta_{1K_1}), (\vartheta_{21}, \dots, \vartheta_{2K_2})$ such that $(\theta_{1i}, \theta_{2i}) = (\vartheta_{1k_{1i}}, \vartheta_{2k_{2i}})$.

We use the EM algorithm of Dempster *et al.* (1977) to estimate the model’s parameters. This amounts to treating k_{1i} and k_{2i} as random variables. Starting with initial guesses for

the parameters, one computes, in the expectation (E) step, the posterior probabilities that $(k_{1i}, k_{2i}) = (k_1, k_2)$ given the data, for all k_j in $\{1, \dots, K_j\}$, $j = 1, 2$ and for all individuals. Then in the maximization (M) step one maximizes the likelihood of the observations, weighted by the posterior probabilities.

As for the choice of K_1 and K_2 there is a trade-off between the accuracy of the description of the unobserved heterogeneity distributions and the tractability of the estimation due to the small number of voluntary job-to-job transitions. We found $K_1 = 4$ and $K_2 = 2$ to be a convenient choice for the countries we study. In the empirical analysis, we shall test the robustness of our conclusions to variations in K_1 and K_2 .

The estimation of the global model takes the form of simple steps. In the M-stage of the algorithm, all parameters— except the ones ruling voluntary job-to-job mobility, given by equations (3)-(4) and (5) —are estimated either by PROBIT, Ordered PROBIT or OLS— weighted by the posterior probabilities. Then, equations (3)-(4) and (5) form a censored regression model with endogenous threshold. Theoretically, one could estimate the model under the constraint (15). However, in the empirical analysis, we allow for five amenities and a bivariate distribution of unobserved heterogeneity, so the joint estimation turns out to be impractical.

Instead, within each M-step of the algorithm, we proceed in two steps: First, we estimate the wage/amenity distribution posterior to constrained job change (ℓ^c). Second, we estimate the parameters ruling the voluntary mobility decision setting $\ell^* = \ell^c$. Appendix B.2 details the mathematical expression of the second-step likelihood. The resulting algorithm follows the pattern introduced in Arcidiacono and Jones (2003) of EM with a sequential M-step (ESM). Therefore, our method provides consistent estimates of the parameters. However, it is not as efficient as FIML.

To compute asymptotic standard errors, we write the First-Order Conditions of the algorithm as population moment conditions. Arcidiacono and Jones (2003) propose to compute the asymptotic variance-covariance matrix by the usual GMM formula. In our case, we found that second derivatives of the type-conditional likelihoods could be very long to compute. In Appendix B.3, we propose a generalization of the information matrix identity that allows to significantly reduce computing time, by up to a 100 factor in our case.

5 Estimation results

In this Section, we first present the parameter estimates. We then contrast the estimated MWP with the wage/amenity correlations in cross-section, interpreting the results in light of section 3.2. Lastly, we check the robustness of our findings to several changes in the model’s specification.

5.1 Parameter estimates

As the model presented in 3.1 contains many parameters, we here give a partial account of the results, focusing on the parameters of interest. Additional estimates are available from the authors upon request.

Wage and amenity equations: Table 3 presents the parameter estimates of wage/amenity equations (1)-(2) for Denmark. The results for other countries are qualitatively similar. Recall that all five amenities have been simultaneously included in the estimation. Parameters ρ_1, \dots, ρ_5 , which represent the within-job compensating differentials for each of the five amenities, are reported separately in Table 4. Table 3 shows that human capital determinants and unobserved match characteristics have very different effects on the wage and the amenities. The wage is concave in age, and is higher for males and educated workers. In contrast, all amenities but distance to job are convex in age, yet no clear pattern arises from the effects of gender and education. In addition, θ_2 has a significant effect on all the amenities, whereas it is independent of the wage by construction.¹⁷ Interestingly, this effect has the same sign and roughly the same magnitude for the five different amenities. This finding could suggest that workers differ in their understanding of the 1 to 6 ranking given in the ECHP. High- θ_2 and low- θ_2 workers could attach different meanings to words such as “very satisfied” or “satisfied”, consistently with a “subjective” interpretation of θ_2 .

Moreover, observed covariates account for 33% of the wage variance,¹⁸ and unobserved heterogeneity θ_1 accounts for 44%. In contrast, both regressors have little explanatory power

¹⁷Strictly speaking, θ_2 is independent of the wage conditional on initial amenities in the job. See Appendix B for details.

¹⁸To compute this variance decomposition, we first regressed the wage on x , then on x and θ_1 . This last regression was weighted by the groups’ posterior probabilities computed at the parameter estimates (see Appendix B). We proceeded similarly for the latent variables of amenities.

in amenity equations. For instance, for the amenity “type of work”, both x and θ_1 account for 3% of the variance, while θ_2 accounts for 26%. Similar orders of magnitude are obtained for the other amenities. Therefore, while observed and unobserved characteristics account for a large part of the wage variance, the determinants of amenity variables seem essentially unobserved.

Table 3: Wage and amenities within job (DNK)

| | Wage | TY | CD | WT | DI | SE |
|---|----------------------|---------------------|---------------------|---------------------|----------------------|---------------------|
| <i>Observed heterogeneity x</i> | | | | | | |
| age | .0272 (.00081) | −.00512 (.010) | −.0244 (.0096) | −.0256 (.0098) | .0369 (.011) | −.0726 (.010) |
| age ² | −.000316 (.00001) | .000116 (.00012) | .000333 (.00011) | .000454 (.00012) | −.000188 (.00013) | .000851 (.00012) |
| male | .120 (.0028) | .0707 (.027) | .0871 (.026) | .00445 (.027) | −.0958 (.030) | −.0929 (.027) |
| edu= 2 nd level | .07160 (.0035) | .0135 (.037) | .0264 (.036) | .000929 (.036) | −.177 (.041) | .0747 (.036) |
| edu≥ 3 rd level | .146 (.0038) | .0362 (.040) | .0192 (.039) | −.0419 (.039) | −.223 (.044) | .193 (.039) |
| constant | 8.989 (.017) | .348 (.21) | .798 (.20) | 1.0379 (.21) | −1.427 (.23) | 1.820 (.21) |
| <i>Unobserved heterogeneity (θ_1, θ_2)</i> | | | | | | |
| $\theta_1 = 1$ | −.986 (.0052) | −.319 (.083) | −.522 (.079) | −.651 (.082) | .632 (.086) | −.372 (.078) |
| $\theta_1 = 2$ | −.568 (.0038) | −.603 (.053) | −.823 (.052) | −.613 (.056) | .563 (.055) | −.389 (.051) |
| $\theta_1 = 3$ | −.326 (.0036) | −.263 (.050) | −.456 (.049) | −.359 (.053) | .356 (.050) | −.0517 (.048) |
| $\theta_2 = 1$ | - | 1.0190 (.029) | .981 (.028) | .684 (.029) | 1.390 (.033) | .825 (.029) |

We now turn to parameters ρ_1, \dots, ρ_5 that give the within job compensating differentials for various amenities. As shown by Table 4, the estimates are very close to zero and, in some cases, insignificant at conventional levels. Moreover, there is no clear pattern in their sign. For instance, in the Netherlands “working conditions” and “working times” are associated to negative wage differentials, while the correlation is positive for “distance to job” and “job security”. In all four cases the estimates are significant at the five percent level, yet their order of magnitude is of less than 2%. We also computed ρ estimates without unobserved heterogeneity (the results are not featured here). Comparing these estimates with those

displayed in Table 4, we found that controlling for the ability bias, through θ_1 , tends to create or increase compensating differentials for disamenities. Even with this correction, though, the ρ point estimates remain rather low.

Table 4: Within-job compensating differentials, ρ

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|----|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|---------------------|
| TY | .0142 (.0046) | -.00328 (.0033) | .0197 (.0035) | .000724 (.0038) | .0188 (.0031) | .00561 (.0062) | .00242 (.0027) | -.00319 (.0030) | .0154 (.0044) |
| CD | -.00162 (.0046) | -.00976 (.0032) | .00672 (.0035) | -.0176 (.0037) | -.00643 (.0029) | -.0102 (.0063) | -.00422 (.0027) | -.0110 (.0027) | -.0101 (.0043) |
| WT | -.00278 (.0044) | -.0383 (.0033) | -.00544 (.0034) | -.0246 (.0036) | -.0167 (.0027) | .000235 (.0062) | -.0149 (.0026) | -.0123 (.0029) | -.0219 (.0045) |
| DI | .0456 (.0038) | .0158 (.0031) | .00931 (.0032) | .0556 (.0037) | .00896 (.0027) | -.00615 (.0059) | .00197 (.0023) | .0171 (.0029) | -.000334 (.0037) |
| SE | .0203 (.0040) | -.0102 (.0030) | .0161 (.0035) | -.00584 (.0038) | .00922 (.0027) | .0497 (.0062) | .00705 (.0026) | .00933 (.0029) | .0190 (.0044) |

The patterns found in Table 3 are qualitatively similar for the wage/amenity equations posterior to constrained job change and posterior to non employment (not featured here). Moreover, the estimates of the parameters of the Ordered PROBIT linking the characteristics of a new match to the starting wage/amenity values in a job are consistent with the patterns of Table 3: Higher wages are associated to high θ_1 . Then, a high starting satisfaction with any amenity yields a higher θ_2 .

Compensating differentials in job offers ρ^* are especially relevant to our analysis. Recall that in this section, restriction (15) is imposed, so that job offer parameters are those corresponding to constrained job change. We shall try other hypotheses at the end of this section. The ρ^* estimates displayed in Table 5 are close to zero, roughly of the same order of magnitude as the ρ estimates reported in Table 4. Moreover, as there are fewer constrained job changers than job stayers, standard errors are higher,¹⁹ resulting in mostly insignificant estimates. Interpreting these results in the light of 3.2 suggests that, for most amenities, the correlation on the demand side might not be sufficiently negative to create large “compensating” wage differences posterior to job change.

¹⁹Standard errors reported in the text are 20 to 50 percent higher than standard errors computed in the last step of the EM algorithm before convergence. In our case, therefore, accounting for group variability has a strong effect on the significance of the estimates.

Table 5: Wage/amenity correlation in job offers, ρ^*

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|----|------------------|-------------------|-------------------|------------------|--------------------|-------------------|--------------------|-------------------|------------------|
| TY | .0134 (.045) | .0261 (.018) | .0205 (.020) | .0365 (.036) | −.000838 (.027) | .0503 (.042) | .0446 (.028) | .0317 (.019) | .1056 (.033) |
| CD | −.0217 (.044) | −.00337 (.019) | −.00561 (.020) | .00150 (.035) | −.0222 (.027) | −.0106 (.051) | −.0433 (.031) | −.0332 (.018) | .0756 (.029) |
| WT | −.0786 (.043) | −.00139 (.018) | −.0110 (.019) | −.0278 (.037) | .0105 (.024) | −.123 (.0452) | .0114 (.029) | −.00351 (.019) | −.0815 (.034) |
| DI | −.0208 (.037) | −.0215 (.016) | −.0315 (.019) | −.0151 (.034) | .00314 (.022) | −.00109 (.045) | .0228 (.024) | −.0389 (.020) | −.0216 (.027) |
| SE | .0228 (.034) | −.00726 (.017) | .0556 (.022) | −.0172 (.035) | −.0142 (.023) | .0352 (.039) | −.000783 (.031) | .0268 (.018) | .00520 (.041) |

Voluntary mobility: We here focus on the determinants of voluntary mobility. In the model, job change decisions are based on the comparison of value functions and mobility costs. Therefore, if costs depend on job characteristics, the parameter estimates featured in equation (5) will be a mixture of these two elements. In this paper we make no attempt to separate the value of a job from true transition costs. Instead, we report in Table 6 the results of an OLS regression of $\mathbb{E}[\tau(\theta, x)]$, where $\tau(\theta, x)$ is defined by (6), on individual covariates and the last wage/amenity values in the current job.²⁰ We interpret these coefficients as the weights of different factors in voluntary mobility decisions.

The signs and significance of the estimates are rather intuitive. In particular, Table 6 shows that the total effect of age and/or age² on $\mathbb{E}[\tau(\theta, x)]$ is positive. Age thus reduces significantly the probability of job change. Being a woman is also associated with a lower propensity to change job. These effects have been already noted in the literature (*e.g.* Groot and Verbene, 1997, and Xenogiani, 2003). Having children and being married generally have a similar, though weaker, influence. The effect of education on voluntary mobility seems non-monotonic. Then, the higher the current wage, the lower the probability to quit voluntarily (the estimates range between .56 in Denmark and 2.2 in Finland). Lastly, the coefficients of amenities on the current job are also positive and significant. These findings, common to all countries and amenities, suggest that being satisfied either with one’s wage or non wage characteristics deters one from quitting. This result is consistent with the literature starting with Freeman (1978) which studies the effects of job satisfaction on the quit probability.

²⁰Again, OLS regressions were weighted by the groups’ posterior probabilities. Standard errors are conditional on x, y_t, \mathbf{a}_t . Yet, they do account for the variability of the parameters entering $\mathbb{E}[\tau(\theta, x)]$.

Table 6: Weight of several covariates in the decision to change job

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|-----------------------------|---------------------|--------------------|---------------------|--------------------|--------------------|---------------------|-------------------|--------------------|-----------------------|
| age | -.0125 (.018) | -.0513 (.022) | -.0205 (.023) | -.0326 (.058) | .0956 (.073) | .0389 (.039) | -.137 (.077) | -.0423 (.024) | .0986 (.050) |
| age ² | .000611 (.00026) | .00119 (.00031) | .000619 (.00032) | .00165 (.00087) | .00019 (.00089) | .000168 (.00050) | .00292 (.0013) | .00117 (.00035) | -.0000791 (.00054) |
| male | -.0260 (.041) | -.119 (.050) | -.157 (.060) | -1.211 (.31) | -1.075 (.38) | -.257 (.10) | -.784 (.27) | -.0916 (.050) | -.924 (.22) |
| edu= 2 nd level | -.258 (.054) | -.139 (.069) | .00277 (.085) | .274 (.19) | .770 (.30) | -.195 (.12) | .0893 (.14) | .550 (.086) | .224 (.17) |
| edu ≥ 3 rd level | -.151 (.089) | -.163 (.077) | -.195 (.098) | .0639 (.17) | -.388 (.22) | -.519 (.16) | .149 (.27) | .4005 (.079) | .199 (.22) |
| married | .0383 (.047) | .00702 (.051) | .0387 (.057) | .2694 (.18) | .287 (.18) | -.0974 (.13) | .475 (.22) | .0737 (.053) | -.233 (.14) |
| kid | -.0564 (.041) | .0792 (.050) | .123 (.061) | -.186 (.14) | -.188 (.17) | .197 (.11) | .154 (.17) | .0323 (.052) | .124 (.11) |
| constant | 4.843 (.31) | 6.244 (.51) | 4.218 (.52) | -8.324 (2.219) | 3.056 (2.19) | -1.323 (.83) | 2.258 (3.35) | 3.815 (.51) | -7.228 (1.41) |
| Wage | .656 (.00053) | .560 (.0066) | .725 (.0063) | 2.230 (.039) | 1.168 (.029) | 1.488 (.0019) | 1.881 (.057) | .754 (.0037) | 1.90 (.026) |
| TY | .166 (.016) | .102 (.016) | .120 (.021) | .375 (.091) | .372 (.12) | .286 (.045) | .307 (.044) | .112 (.015) | .271 (.040) |
| CD | .152 (.017) | .0905 (.013) | .120 (.020) | .200 (.077) | .410 (.15) | .217 (.046) | .128 (.054) | .0887 (.013) | .244 (.054) |
| WT | .0811 (.010) | .0842 (.0082) | .0987 (.014) | .236 (.065) | .303 (.10) | .101 (.034) | .166 (.061) | .0968 (.014) | .263 (.039) |
| DI | .00959 (.010) | .135 (.031) | -.0178 (.020) | .267 (.075) | .231 (.093) | .158 (.043) | .0852 (.047) | .154 (.030) | .238 (.063) |
| SE | .140 (.015) | .0988 (.014) | .0913 (.0092) | .374 (.075) | .486 (.15) | .172 (.030) | .376 (.046) | .0652 (.011) | .315 (.057) |

Then, we divide the coefficients of current amenities by the coefficient of the wage. We interpret the estimates reported in Table 7 as the relative weight of each amenity in the decision to change job. Unsurprisingly, we find positive and significant estimates for virtually all amenities in every country, ranging around .20. The smallest coefficients are obtained for “distance to job”, for which the estimates are insignificant at the 95% level in Austria, Finland and Italy. Note that the estimates are on average both higher and less precisely estimated in France than in the other countries.

The results in Table 7 are close in spirit to the methodology introduced by Gronberg and Reed (1994), who estimate the “Marginal Willingness to Pay” for an amenity as the ratio of the elasticities of the hazard rate of job duration with respect to the amenity and the wage, respectively. We obtain comparable results: Gronberg and Reed find that two amenities out of four—measuring several dimensions of “objective” working conditions—have a positive

and significant effect on job duration. In the case where they are significant, the relative weights are close to one third of the wage. Van Ommeren *et al.* (2000) obtain similar orders of magnitude in their analysis of commuting.

Table 7: Weight of *current* amenities in the mobility decision, relative to the *current* wage

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|----|-----------------|----------------|----------------|-----------------|----------------|-----------------|-----------------|-----------------|----------------|
| TY | .252 (.024) | .182 (.032) | .143 (.023) | .165 (.030) | .318 (.11) | .192 (.030) | .163 (.028) | .148 (.021) | .143 (.023) |
| CD | .231 (.026) | .162 (.026) | .128 (.030) | .166 (.029) | .351 (.13) | .146 (.031) | .0683 (.031) | .118 (.017) | .128 (.030) |
| WT | .124 (.015) | .150 (.016) | .138 (.023) | .136 (.021) | .260 (.093) | .0678 (.023) | .0885 (.035) | .128 (.020) | .138 (.023) |
| DI | .0146 (.016) | .241 (.058) | .125 (.035) | -.025 (.028) | .198 (.085) | .106 (.029) | .0453 (.026) | .205 (.041) | .125 (.035) |
| SE | .214 (.024) | .176 (.026) | .166 (.032) | .126 (.014) | .416 (.14) | .116 (.020) | .20 (.031) | .0865 (.015) | .166 (.032) |

Heterogeneity in mobility costs. Table 8 reports the estimates of the standard deviation of the stochastic shocks on mobility costs u_z , together with estimates of the weight of wage offers in voluntary job change R_{yz} , defined by (10).

We note that the estimates of the standard deviation σ_z range between .69 in Denmark and 3.6 in France. The second row in Table 8 illustrates the magnitude of these standard errors by reporting the estimates of R_{yz} . As explained in section 3.2, we interpret this quantity as a measure of heterogeneity in mobility costs. Estimates of R_{yz} are strikingly low in the nine countries, suggesting that many other factors than wage offers might influence the decision to quit. Moreover, voluntary mobility seems much more heterogeneous in the Latin countries (Spain, France, Italy and Portugal) with a ratio R_{yz} of less than one percent and insignificant from zero at the 95% level. This could indicate that individual mobility is highly constrained in these countries, which would be in accordance with the descriptive statistics featured in Table 1.

Table 8: Heterogeneity in mobility costs

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|------------|----------------|------------------|------------------|-----------------|-------------------|-----------------|------------------|------------------|-------------------|
| σ_z | .689 (.052) | 1.122 (.13) | 3.22 (.83) | .898 (.12) | 3.597 (1.19) | 1.81 (.32) | 3.22 (1.01) | 1.245 (.15) | 2.52 (.61) |
| R_{yz} | .122 (.018) | .0422 (.0097) | .0058 (.0030) | .0688 (.018) | .00455 (.0030) | .0296 (.011) | .0058 (.0036) | .0326 (.0078) | .00930 (.0045) |

The explanatory power of both the wage and the amenity variables in voluntary mobility decisions thus appears to be weak, suggesting that mobility costs are highly heterogeneous. Such a low explanatory power is one of our main findings, and has strong implications on the order of magnitude of wage/amenity correlation. At the end of this section we shall investigate the robustness of this result to variations in the model’s specification, as explained in section 4.1.

MWP for amenities: We lastly turn to the estimates of the key parameters, the MWP δ^* in job offers. Most estimates in Table 9 are positive, and several are significant. In particular, the type of work and job security are associated to large MWP, around .30 in Denmark and the Netherlands. In France, workers seem especially attached to the security of their job with a very high MWP for this amenity. However, although significant, this MWP is not precisely estimated (1.1 with a standard error of .4). This remark carries out to all other amenities in the Latin countries (Spain, France, Italy and Portugal), for which standard errors are much higher than in the other countries. Note that, for these countries, imprecise estimates of MWP are associated with low R_{yz} estimates (see Table 8). Hence, for these four countries only, we should consider estimation results as qualitative findings rather than focusing on the point estimates. In 5.3, we shall see that restricting the definition of constrained transitions, or splitting the sample by gender, yields more comparable results between Latin and non Latin countries.

Table 9: MWP for amenities in job offers: δ^*

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|----|-----------------|-----------------|----------------|------------------|-----------------|----------------|-----------------|------------------|----------------|
| TY | .191 (.058) | .275 (.067) | .512 (.21) | .191 (.062) | .792 (.32) | .248 (.13) | .877 (.34) | .271 (.067) | .893 (.27) |
| CD | .225 (.059) | .0798 (.053) | .203 (.18) | .0834 (.057) | .172 (.18) | .0313 (.13) | .318 (.20) | .00712 (.049) | .319 (.16) |
| WT | .162 (.052) | .0225 (.053) | -.196 (.18) | .0654 (.056) | -.0395 (.15) | -.180 (.11) | -.0371 (.19) | .00882 (.052) | .0988 (.16) |
| DI | .0553 (.041) | .0614 (.048) | .126 (.16) | -.0833 (.052) | .0886 (.15) | .134 (.11) | .317 (.17) | -.190 (.055) | .336 (.15) |
| SE | .254 (.046) | .273 (.056) | .720 (.26) | .348 (.070) | 1.133 (.40) | .426 (.14) | .994 (.35) | .265 (.057) | .868 (.26) |

The most notable exception to this general pattern is “working times” in all countries, insignificantly different from zero in most cases. This non intuitive result could be due to the

fact that wage rates and hours worked are aggregated in the model (into monthly wages), resulting in a crude modeling of hours worked in workers’ preferences. Other exceptions are “working conditions” in the Netherlands and “distance to job ” in Finland and the Netherlands, the latter amenity being associated to a negative and significant MWP.

Nevertheless, the general impression that emerges from Table 9 is one where MWP in job offers can be large. This result can be contrasted with Tables 4 and 5, showing compensating differentials for job stayers and in job offers, respectively. Thus, both the wage and non wage characteristics seem to influence voluntary mobility decisions, suggesting that non wage characteristics do enter individual preferences. In terms of variance, though, the influence appears quite weak, as both the wage and amenities have a low explanatory power in job change decisions. In the next section, we intend to quantify the impact of these two findings on the presence/absence of “compensating” wage differences in cross-section.

5.2 MWP and wage differentials

The analysis in 3.2 shows that the actual wage differential posterior to job change— between two jobs of different levels of amenities – combines the model’s three key parameters: the correlation in wage/amenity offers ρ^* , the MWP for the amenity δ^* and the heterogeneity of mobility costs R_{yz} . We here report the estimates of the various wage differentials, and of their decomposition in terms of the demand and supply effects introduced in 3.2.

For a given amenity, the demand effect $\Delta_z^{(d)}$ is equal to the compensating differential ρ^* in job offers corresponding to this amenity. Estimates of this effect can be found in Table 5. The supply term, $\Delta_z^{(s)}$, arises from workers’ trade-offs between wages and amenities when deciding to change job. In the upper half of Table 10, we report the estimates of $\Delta_z^{(s)}$ for all countries and amenities.²¹

Estimates of $\Delta_z^{(s)}$ are mostly negative, consistently with wage/amenity compensation on the supply side, as MWP δ^* are mostly positive (see Table 9). However, heterogeneity in mobility costs significantly reduces the magnitude of this effect. Comparing Table 10 with

²¹To estimate the latter, we computed the RHS in (9) for every individual in the sample, and averaged over x , weighted by the groups’ posterior probabilities. In theory, the delta-method is not sufficient to compute standard errors in this case, as one has to account also for the estimation of the expectation in (9) by a sample mean. As this latter source of variation turned out to be negligible relative to the variation in the model’s parameters, however, the delta-method was used as yielding a very good approximation of true standard errors.

Table 9 shows that MWP of .30 translate into correlations of less than .02.²² Then, the lower half of Table 10 shows the sum of the demand and supply effects $\Delta_z = \Delta_z^{(d)} + \Delta_z^{(s)}$. It is clearly apparent from the table that, when heterogeneous mobility costs and non zero correlation in job offers are taken into account, the resulting wage/amenity correlation posterior to job change does not reflect workers' underlying preferences for non wage attributes.

Table 10: Wage differentials for voluntary job changers

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|----|--|---------------------|----------------------|--------------------|-----------------------|---------------------|---------------------|---------------------|---------------------|
| | $\Delta_z^{(s)}$ | | | | | | | | |
| TY | -.0231 (.0084) | -.0112 (.0027) | -.00279 (.0012) | -.0141 (.0048) | -.00327 (.00129) | -.00781 (.0035) | -.00483 (.0017) | -.00871 (.0020) | -.00832 (.0024) |
| CD | -.0230 (.0086) | -.00286 (.0021) | -.00104 (.00093) | -.00528 (.0041) | -.000621 (.000738) | -.000544 (.0036) | -.00145 (.0010) | .000758 (.0015) | -.00331 (.0015) |
| WT | -.00943 (.0076) | -.000791 (.0021) | .00109 (.00089) | -.00233 (.0042) | .000120 (.000651) | .00801 (.0038) | .000135 (.00099) | -.000154 (.0016) | -.000146 (.0014) |
| DI | -.00392 (.0063) | -.00150 (.0019) | -.000497 (.00086) | .00614 (.0038) | -.000381 (.000626) | -.00349 (.0033) | -.00179 (.0010) | .00670 (.0017) | -.00264 (.0013) |
| SE | -.0312 (.0069) | -.00990 (.0023) | -.00405 (.0013) | -.0204 (.0049) | -.00461 (.00166) | -.0121 (.0037) | -.00520 (.0019) | -.00840 (.0019) | -.00728 (.0023) |
| | $\Delta_z = \Delta_z^{(d)} + \Delta_z^{(s)}$ | | | | | | | | |
| TY | -.00970 (.040) | .0149 (.018) | .0177 (.020) | .0224 (.034) | -.00411 (.027) | .0425 (.041) | .0398 (.028) | .0230 (.018) | .0973 (.033) |
| CD | -.0447 (.039) | -.00624 (.019) | -.00665 (.020) | -.00377 (.033) | -.0228 (.027) | -.0112 (.049) | -.0447 (.031) | -.0325 (.018) | .0722 (.029) |
| WT | -.0880 (.039) | -.00218 (.017) | -.00987 (.019) | -.0302 (.035) | .0106 (.024) | -.115 (.044) | .0115 (.029) | -.00366 (.019) | -.0817 (.033) |
| DI | -.0247 (.033) | -.0230 (.016) | -.0320 (.019) | -.00893 (.032) | .00276 (.022) | -.00458 (.044) | .0210 (.024) | -.0322 (.020) | -.0242 (.027) |
| SE | -.00845 (.031) | -.0171 (.017) | .0516 (.022) | -.0377 (.033) | -.0188 (.023) | .0231 (.038) | -.00598 (.031) | .0184 (.017) | -.00208 (.041) |

Thus, evidence of “compensating” wage differences is rather weak in cross-section, although workers seem to value non wage characteristics. Workers’ trade-offs between wages and amenities translate into a very small, possibly still negative, correlation. This section has emphasized two key elements in this mechanism: the low explanatory power of the wage and amenities in job mobility decisions (low R_{yz}), and the often insignificant correlation

²²Notice that the approximation of the wage differential $\Delta_z^{(s)}$ given in (12) works here very well. Indeed, combining the results in Tables 8, 9 and 10 it is easy to check that compensating differentials are roughly the product of MWP δ_z^* (net of ρ^* , which is close to zero in most cases) and heterogeneity in mobility costs R_{yz} .

in job offers (low $|\rho^*|$). Our results thus shed light on the difficulty of finding compensating differentials in cross-section, even conditional on unobserved heterogeneity, and even if individuals value non wage characteristics significantly relatively to the wage.

5.3 Robustness checks

We here check the robustness of the parameter estimates reported in 5.1 to changes in the model’s specification. First, we disaggregate constrained transitions into *partially constrained transitions* and *displacements*, and model each process separately. We do so to address the concern that constrained transitions, as defined in this paper, may recover different phenomena. For instance, family-related job mobility is an example of partially constrained transition. However, employer-related job changes (firm closure, layoff...), that we call displacements, could be more exogenous from the worker’s perspective.

We model partially constrained transitions and displacements as in equation (7), with different sets of parameters. Likewise, wage and amenity equations follow the pattern of (1)-(2), again with different parameters. The sequence of events is the following: between t and $t+1$ employed individuals can experience a job to non employment transition. If they do not, they can still lose their job and get a new one before the next interview. Then, if they are not displaced, they can quit their present job for personal reasons and make a partially constrained transition to a new job. All the probabilities and corresponding wage/amenity distributions are conditional on x and θ . Lastly, if the worker has experienced none of these shocks, she receives a job offer that she can accept or turn down, as in 3.1.

Within this framework, we try three different specifications for job offers. Subsequently, job offer parameters are assumed equal to the ones in displacements, partially constrained transitions and transitions out of non employment.²³ Table 11 presents the estimation results for MWP in job offers δ^* and heterogeneity in mobility costs R_{yz} .

Table 11 reinforces the two main qualitative findings of 5.1. First, MWP for amenities are mostly positive, and can be large for some amenities. Thus, the type of work and job security are associated with positive and significant MWP for almost all countries and every choice of transition of reference. Moreover, many MWP amount to a large share of the wage,

²³In several cases (5 out of 27), there was not enough information in the data to identify $K_1 = 4$ groups of heterogeneity θ_1 . We have thus retained $(K_1, K_2) = (3, 2)$ for the following crossings: displacements and out-of-non employment in Ireland and partially constrained transitions in Spain, Ireland and Portugal.

up to 40% in non Latin countries. Second, every specification shows large heterogeneity in mobility costs. For instance, in the Netherlands wage offers account for less than 8% of the variation in voluntary mobility, irrespective of the type of transitions chosen as a reference.

Table 11: MWP (δ^*) and R_{yz} estimates for various transitions of reference

| | AUS | DNK | ESP | FIN | FRA | IRL | ITA | NLD | PRT |
|--|------------------|------------------|-------------------|------------------|-------------------|-------------------|-------------------|-------------------|------------------|
| <i>Displacements</i> | | | | | | | | | |
| TY | .263 (.051) | .193 (.055) | .659 (.31) | .177 (.065) | .725 (.29) | .143 (.083) | 1.569 (1.11) | .208 (.038) | .437 (.10) |
| CD | .162 (.050) | .0366 (.045) | .223 (.24) | .0676 (.059) | .0951 (.16) | .0779 (.088) | .645 (.55) | -.0169 (.030) | .130 (.081) |
| WT | .178 (.045) | -.0570 (.047) | -.348 (.25) | .0114 (.059) | -.111 (.15) | -.206 (.082) | -.470 (.50) | -.00579 (.032) | .0526 (.089) |
| DI | .0482 (.036) | .116 (.042) | .0735 (.21) | -.102 (.055) | .0245 (.14) | .0146 (.080) | 1.012 (.68) | -.122 (.032) | .221 (.079) |
| SE | .378 (.046) | .418 (.057) | 1.033 (.43) | .370 (.076) | 1.323 (.43) | .401 (.095) | 3.020 (1.96) | .3296 (.036) | .758 (.13) |
| R_{yz} | .117 (.023) | .0489 (.010) | .00338 (.0023) | .0446 (.019) | .00555 (.0035) | .0517 (.014) | .00150 (.0019) | .0806 (.012) | .0289 (.0080) |
| <i>Partially constrained transitions</i> | | | | | | | | | |
| TY | .157 (.077) | .423 (.092) | .171 (.053) | .165 (.041) | .411 (.077) | .227 (.099) | .434 (.11) | .301 (.075) | .725 (.20) |
| CD | .428 (.082) | .156 (.067) | .146 (.054) | -.0374 (.041) | .117 (.064) | -.0490 (.10) | .308 (.093) | .0451 (.052) | .183 (.14) |
| WT | .134 (.069) | .130 (.065) | .0181 (.055) | .10 (.040) | .185 (.060) | -.0185 (.086) | .0418 (.086) | .0402 (.056) | -.00932 (.15) |
| DI | .0899 (.053) | .00902 (.058) | .152 (.060) | -.0243 (.035) | .143 (.056) | .261 (.092) | .0624 (.067) | -.206 (.060) | .0953 (.12) |
| SE | .222 (.059) | .131 (.063) | -.0409 (.050) | .305 (.047) | .0895 (.056) | .237 (.095) | .235 (.083) | .156 (.056) | .467 (.16) |
| R_{yz} | .0772 (.017) | .031 (.0088) | .0591 (.015) | .174 (.046) | .0242 (.0068) | .046 (.014) | .0217 (.0068) | .0241 (.0063) | .0108 (.0049) |
| <i>Out-of-non employment</i> | | | | | | | | | |
| TY | .902 (.25) | .154 (.070) | -.0155 (.059) | .209 (.10) | .454 (.11) | .405 (.14) | .292 (.088) | .401 (.065) | .125 (.048) |
| CD | .759 (.22) | .127 (.061) | .0445 (.060) | .175 (.099) | .518 (.11) | -.179 (.15) | -.0536 (.080) | .0241 (.045) | .202 (.053) |
| WT | -.046 (.17) | .173 (.062) | -.0469 (.058) | -.0525 (.094) | .119 (.073) | -.369 (.14) | -.00877 (.082) | .109 (.049) | -.0088 (.053) |
| DI | .228 (.13) | -.0918 (.057) | .0723 (.055) | -.310 (.11) | .204 (.074) | .332 (.13) | -.0537 (.066) | -.0598 (.045) | .0944 (.047) |
| SE | .597 (.17) | .317 (.066) | .0931 (.058) | .580 (.14) | .354 (.088) | .0402 (.12) | .363 (.086) | .227 (.048) | .0845 (.052) |
| R_{yz} | .0104 (.0044) | .0281 (.0071) | .0516 (.0082) | .0104 (.0044) | .0106 (.0032) | .00949 (.0035) | .00581 (.0017) | .0635 (.012) | .0531 (.0092) |

The main results of 5.1 thus appear robust to changes in the reference distribution. Still,

a closer look at Table 11 reveals several interesting features. First, we note that the choice of the distribution of reference rarely influences the sign or significance of the MWP estimates. However, it can alter their ranking in a given country. For instance, the type of work and job security are associated with the highest MWP in any country when using either constrained transitions or displacements as the reference. In contrast, when partially constrained job change is used, the MWP for job security strongly decreases. For instance, in Denmark the MWP estimate for this amenity is .42 when using displacements as the reference, and .13 (yet still significant at the 95% level) when using partially constrained transitions. An intuitive explanation could be that such transitions are experienced by workers who change jobs for personal reasons (marriage, geographic mobility...) but whose alternative is *not* non employment. In this interpretation, MWP estimates for job security based on the use of displacements are higher, because they incorporate the risk of non employment— possibly correlated with the aversion to job insecurity.

A second interesting feature of Table 11 concerns the grouping of countries which emerged from Tables 8 and 9. In the Latin countries, indeed, in the two cases where displacements are not part of the transitions of reference, heterogeneity in mobility costs is reduced. In Spain, France, Italy and Portugal, the R_{yz} estimates for these two specifications are higher, and significantly different from zero. Simultaneously, MWP estimates are more in line with the results for non Latin countries, ranging around one third of the wage for the type of work and job security. Moreover, when restricting constrained transitions to partially constrained or out-of non employment transitions, standard errors of MWP estimates are lower, indicating that MWP are better estimated. These findings suggest that displacements (as defined in this paper) are not a satisfying control group in Latin countries.

Additional checks: Testing for the robustness of the estimates with respect to changes in the distribution of reference was essential to our approach and we view these empirical regularities as supportive of our results and interpretation. Still, one can question other features of the model than the identification assumption discussed in 4.1. We thus proceeded to a series of alternative robustness checks.²⁴ First, we varied the number of groups for θ_1 or θ_2 and tried the following (K_1, K_2) pairs: (1, 1), (2, 2), (3, 2), (5, 2) and (4, 3). The results,

²⁴In the rest of this section, constrained transitions are taken as reference and restriction (15) is imposed, as in 5.1.

as far as MWP and R_{yz} estimates are concerned, are qualitatively similar when allowing for more than $K_1 = 2$ and $K_2 = 2$ groups. In some countries such as Denmark and Ireland, this is also the case when no heterogeneity is allowed for. In the Netherlands or Portugal, however, results differ greatly in the homogeneous case, with σ_z estimates of 3.9 and 11.0, respectively, and large and badly estimated MWP. We modified the wage observations by adding an *i.i.d.* perturbation, normally distributed with standard error equal to 10% of that of the observed wage. We found the estimates strongly robust to this kind of measurement error.

Lastly, we estimated the model on separate groups of similar age or gender. In the model, voluntary mobility depends on individual covariates in a parametric way, as shown by (5). However, it could be that younger and older workers, or women and men, have very different mobility behaviors, which could not be well captured by a parametric specification. Moreover, this problem could affect some countries more than others. For instance, in Table 6, we see that, compared to other covariates, being a man has a large positive effect on the probability of changing job voluntarily in France, Italy or Portugal. To address this issue, we proceeded to the estimation of the model on the subsamples of men. We found that the heterogeneity in mobility costs remains practically similar in non Latin countries, except in the Netherlands where σ_z increases to 1.6. In Latin countries, σ_z decreases but stays at rather high levels. For instance, the σ_z estimate goes down to 2.8 in Spain, to 2.4 in France and to 2.6 in Italy (the σ_z estimates are 3.2, 3.6 and 3.2, respectively, when pooling men and women, see Table 8). Moreover, if the MWP for amenities are not qualitatively affected, their order of magnitude is somewhat closer to the one obtained for non Latin countries. For instance, in Italy the MWP for “type of work” is reduced from .90 (men and women) to .58 (men only), the MWP for “job security” from .99 to .63. The highest estimate reported in Table 9, the MWP for “job security” in France, goes down from 1.13 to a still high– yet more reasonable– .76. To save space, we do not report the corresponding results in the present version but they are available upon request.

To summarize the main results of this section, we find large and significant MWP in many cases, especially for two amenities: the type of work and job security. We also find, under every specification, high heterogeneity in mobility costs. An intriguing finding is the contrast between Latin and non Latin countries, the former being associated with even higher mea-

tures of heterogeneity in mobility costs, together with imprecise MWP estimates. In the end of this section, we have emphasized two possible explanations: the greater heterogeneity in mobility behavior between men and women in Latin countries, and the greater heterogeneity in types of constrained transitions (displacements *versus* partially constrained).

6 Conclusion

The theory of compensating differentials builds on Adam Smith’s original statement:²⁵

“The whole of the advantages and disadvantages of the different employments of labour and stock must, in the same neighborhood, be either perfectly equal or continually tending to equality.”

On the labor market, this implies that “bad” non-monetary characteristics of one’s job must be compensated by higher wages. However, hedonic wage regressions lead to weak or even wrong-signed wage/amenity correlations. In this paper, we show that these results must not be interpreted as reflecting individual preferences for non wage amenities. Smith had indeed pointed out the conditions under which the “*equality of advantages and disadvantages*” was to be expected:

“This at least would be the case in a society where things were left to follow their natural course, where there was perfect liberty, and where every man was perfectly free both to choose what occupation he thought proper, and to change it as often as he thought proper.”

In modern European economies, very low rates of voluntary mobility suggest that workers are far from being “*perfectly free*” to change jobs. Consequently, the predictions of the theory of compensating differentials are unlikely to hold.

Our estimation results show significant valuation of several non wage characteristics, mostly the type of work and job security, in spite of low wage/amenity correlations. However, the low explanatory power of both the wage and amenities in job mobility, and the small correlation in job offers, imply that workers’ preferences do not translate into significant negative correlation.

The method advocated in this paper makes use of the difference in the degree of constraints in mobility decisions to reveal individual preferences: *constrained* transitions allow

²⁵*An Inquiry Into the Nature and Causes of the Wealth of Nations*, Book 1, Chapter 10, Introduction.

to estimate the available alternatives, then *voluntary* ones permit to measure the true effects of individual choice. This approach could be applied to other fields where hedonic methods are widely used. An example is the estimation of MWP for air quality in environmental economics. In a recent paper, Bayer, Keohane and Timmins (2005) estimate a model of residential sorting allowing for mobility costs. Their MWP estimates are larger than usual hedonic regression estimates. The method and results of the present paper suggest that, in such a field, distinguishing between the reasons to migrate could prove fruitful to deal with endogeneity problems.

Lastly, on the labor market, our results shed light on the empirical content of non wage job characteristics, as they are part of workers' preferences. This evidence could be seen as a motivation for labor economists to incorporate other job attributes than the wage into their models. We note that there have recently been several attempts at broadening the analysis of earnings inequality to that of monetary compensation inequality (*e.g.* Pierce, 2001) or, more generally, inequality in the returns to work (Hamermesh, 1999). In a different field, dynamic structural models of the labor market are just starting to take amenities into account. For instance, Dey and Flinn (2005) write and estimate an equilibrium job search model where firms can also provide health insurance to their employees. We view our results and their implications as supportive of these multi dimensional analyses of the labor market, for which the availability of more informative data sets now seems to be the main obstacle.

References

- [1] Arcidiacono, P., and J.B. Jones (2003): "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71(3), 933-946.
- [2] Bayer, P., N. Keohane, and C. Timmins (2005): "Migration and Hedonic Valuation: The Case of Air Quality," *mimeo*.
- [3] Brown, C. (1980): "Equalizing Differences in the Labor Market," *Quarterly Journal of Economics*, 94, 113-134.
- [4] Christensen, B. J., R. Lentz, D. T. Mortensen, G. R. Neumann and A. Werwatz (2005) "On the Job Search and the Wage Distribution," Forthcoming Journal of Labor Economics.
- [5] Clark, A. E., and A. J. Oswald (1994): "Unhappiness and Unemployment," *Economic Journal*, 104, 648-659.
- [6] Cox, D. R. (1975): "Partial Likelihood," *Biometrika*, 62(2), 269-276.

- [7] Dale-Olsen, H. (2005): “Estimating Workers’ Marginal Willingness to Pay for Safety Using Linked Employer-Employee Data,” *Economica*, forthcoming.
- [8] Dempster, A. P., N. M. Laird and D. B. Rubin (1977): “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *Journal of the Royal Statistical Society*, B 39(1), 1-38.
- [9] Dey, M., and C.J. Flinn (2005): “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica*, 73(2), 571-628.
- [10] Dolton, P.J., G. Makepeace, and W. Van der Klaauw (1989): “Occupational Choice and Earnings Determination: The Role of Sample Selection and Non-pecuniary Factors,” *Oxford Economic Papers*, 41(3), 573-594.
- [11] Duncan, G., and B. Holmlund (1983): “Was Adam Smith Right After All? Another Test of the Theory of Compensating Differentials,” *Journal of Labor Economics*, 1, 366-379.
- [12] Dustmann C., and C. Meghir (2005): “Wages, Experience and Seniority,” *Review of Economic Studies*, 72(1), 77-108.
- [13] Ekeland I., J. Heckman, and L. Nesheim (2002): “Identifying hedonic models,” *American Economic Review*, 92(2), 304-309.
- [14] Ekeland I., J. Heckman, and L. Nesheim (2004): “Identification and estimation of hedonic models,” *Journal of Political Economy*, 112, S60-S109.
- [15] Flinn C. J., and J. Heckman (1982): “New Methods for Analyzing Structural Models of Labor Force Dynamics,” *Journal of Econometrics*, 18, 115-168.
- [16] Freeman R. B. (1978): “Job Satisfaction as an Economic Variable,” *American Economic Review*, Papers and Proceedings, 68, 115-168.
- [17] Gibbons R. and L. Katz (1991): “Layoffs and Lemons,” *Journal of Labor Economics*, 9(4), 351-380.
- [18] Gibbons R. and L. Katz (1992): “Does Unmeasured Ability Explain Inter-Industry Wage Differentials,” *Review of Economic Studies*, 59(3), 515-535.
- [19] Gronberg, T., and R. Reed (1994): “Estimating Workers’ Marginal Willingness to Pay for Job Attributes Using Duration Data,” *Journal of Human Resources*, 29, 911-931.
- [20] Groot, W., and M. Verbene (1997): “Aging, Job Mobility, and Compensation,” *Oxford Economic Papers*, New Series, 49, 380-403.
- [21] Hammermesh, D. S. (1999): “The Changing Inequality for Workplace Amenities,” *Quarterly Journal of Economics*, 114, 1085-1123.
- [22] Heckman, J., R. Matzkin, and L. Nesheim (2005): “Simulation and Estimation of Hedonic Models,” in T.J. Keohe, T.N. Srinivasan, and J. Whalley, ed., *Frontiers in Applied General Equilibrium Modeling, In Honor of Herbert Scarf*. Cambridge.

- [23] Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271-320.
- [24] Herzog, H.W., and A.M. Schlottmann (1990): "Valuing Risk in the Workplace: Market Price, Willingness to Pay, and the Optimal Provision of Safety," *Review of Economics and Statistics*, 72(3), 463-470.
- [25] Hwang, H.S., D. Mortensen, and W.C. Reed. (1998): "Hedonic Wages and Labor Market Search," *Journal of Labor Economics*, 16, 815-847.
- [26] Hwang, H. S., W. C. Reed, and C. Hubbard (1992): "Compensating Differentials and Unobserved Productivity," *Journal of Political Economy*, 100, 835-858.
- [27] Kotlarski, I. (1967): "On Characterizing the Gamma and Normal Distribution," *Pacific Journal of Mathematics*, 20, 69-76.
- [28] Lang, K., and S. Majumdar (2004): "The Pricing of Job Characteristics When Markets do not Clear: Theory and Policy Implications," *International Economic Review*, 45, 1111-1128.
- [29] Manski, C. F. (1988): "Identification of Binary Response Models," *Journal of the American Statistical Association*, 83, 729-738.
- [30] Newey, W.K., and D.L. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics* vol 4: 2111-2245. Amsterdam: Elsevier Science.
- [31] Pierce, B. (2001): "Compensation Inequality," *Quarterly Journal of Economics*, 116(4), 1493-1525.
- [32] Postel-Vinay, F. and J.-M. Robin (2002) "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica* 70(6), 2295-350.
- [33] Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34-55.
- [34] Rosen, S. (1986): "The Theory of Equalizing Differences," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics* vol 1: 2 641-692. Amsterdam: Elsevier Science.
- [35] Senick, C. (2005): "What Can We Learn from Subjective Data: the Case of Income and Well-Being," *Journal of Economic Surveys*, 19(1), 43-63.
- [36] Smith, A.: *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776.
- [37] Thaler, R., and S. Rosen (1975): "The Value of Saving a Life: Evidence from the Labor Market", in N. Terleckyj, ed., *Household Production and Consumption*. New York: NBER.
- [38] Van den Berg, G. (1992): "A Structural Dynamic Analysis of Job Turnover and the Costs Associated with Moving to Another Job," *Economic Journal*, 102, 1116-1133.

- [39] Van Ommeren, J., Van den Berg, G., and C. Gorter (2000): “Estimating the Marginal Willingness to Pay for Commuting,” *Journal of Regional Science*, 40, 541-563.
- [40] Villanueva E. (2005): “Compensating Wage Differentials and Voluntary Job Changes: Evidence from West Germany,” *mimeo* UPF.
- [41] Xenogiani, T. (2003): “Job Satisfaction, Mobility Decisions and Wage Gains by Gender,” LSE Working Paper.

APPENDIX

A Data

The definition of jobs: We let individuals be in either one of the two following labor market states: employed or unemployed. Unemployment comprises self declared unemployment, inactivity, employment during less than 15 hours per week or with wages lower than the first percentile (which, for example, is around 235 Euros per month in France, that is 25% of the median wage). We drop every individual who experiences a self employment spell since we assume her trajectory (and especially her job mobility decisions) not to be governed by the same processes as those of workers in paid employment.

Attrition: Some of the observation periods are right censored, *i.e.* individuals do not always stay in the ECHP during the eight waves. We assume this right censoring to be exogenous to the wage, amenity and job mobility process.

Missing data: The problem of missing data is twofold: there can be non reported variables for a given wave where the individual is present or the individual can “disappear” from the survey during a year within his observation period and come back one year later. When it is possible, we impute missing data on wages and/ or amenity using the previous or following wave if the individual is still in the same job: we substitute the missing wage for the mean of the previous and following wage and draw the amenity from a binomial distribution weighting both the previous and following amenity with probability 0.5 (the amenity can change within a job). These substitutions affect less than a thousand observations (over *e.g.* more than 30 000 in France). For the few observations that still show missing data, we create two individuals out of one. This rather arbitrary treatment of less than 1% of our sample does not affect the consistency of the estimates and the loss of efficiency is likely to be small.

B The estimation procedure

In this section of the Appendix, we detail the estimation procedure of the model presented in 3.1. We start by setting the notations. We let $i \in \{1 \dots N\}$ denote individuals, and $t \in \{1 \dots T\}$ denote time periods. Let e_{it} be the dummy variable indicating if individual i is employed at time t . The model allows for multivariate amenities $\mathbf{a}_{it} \in \{0, 1\}^J$. In the empirical analysis, we take $J = 5$. As outlined in 3.1, the dynamics of employment, wages and amenities follow:

$$\begin{aligned} (e_{it+1}, y_{it+1}, \mathbf{a}_{it+1} | e_{it} = 1) &= (0, 0, 0) \quad \text{if } z_{it}^e = 1, \\ &= (1, y_{it}^c, \mathbf{a}_{it}^c) \quad \text{if } z_{it}^e = 0 \text{ and } z_{it}^c = 1, \\ &= (1, y_{it}^*, \mathbf{a}_{it}^*) \quad \text{if } z_{it}^e = 0, z_{it}^c = 0 \text{ and } z_{it} = 1, \\ &= (1, y_{it}^r, \mathbf{a}_{it}^r) \quad \text{if } z_{it}^e = 0, z_{it}^c = 0 \text{ and } z_{it} = 0, \\ (e_{it+1}, y_{it+1}, \mathbf{a}_{it+1} | e_{it} = 0) &= (1, y_{it}^n, \mathbf{a}_{it}^n) \quad \text{if } z_{it}^n = 1, \\ &= (0, 0, 0) \quad \text{if } z_{it}^n = 0, \end{aligned}$$

where z_{it} indicates voluntary mobility, and $(^c)$, $(^n)$ and $(^e)$ superscripts refer to constrained job-to-job mobility, non employment-to-job and job-to-non employment transitions, respectively. Moreover, to avoid confusions we have used superscript $(^r)$ to index within-job wage and amenity distributions, which appear in equations (1)-(2) in 3.1.

All equations in section 3.1 are easily adapted to the case of multivariate amenities. For instance, wage and amenity offers (3)-(4) become:

$$\begin{aligned} a_{j,it+1}^* &= \mathbf{1} \{ \alpha_{j,a}^* x_{it} + \beta_{j,1a}^* \theta_{1it} + \beta_{j,2a}^* \theta_{2it} + u_{j,ait+1}^* > 0 \} \quad \forall j \in \{1..J\}, \\ y_{it+1}^* &= \boldsymbol{\rho}^* \mathbf{a}_{it+1}^* + \alpha_y^* x_{it} + \beta_{j,y}^* \theta_{1it} + u_{yit+1}^*. \end{aligned} \quad (\text{B1})$$

In (B1), $\boldsymbol{\rho}^* = (\rho_1^*, \dots, \rho_J^*)$ is the vector of compensating differentials in job offers associated to the J amenities.

B.1 The EM algorithm

Consider an individual i , and a given job which lasts from t_{i0} to $t_{i1} - 1$. Unobserved match characteristics are assumed constant on $[t_{i0} + 1, t_{i1}]$. Moreover, they are realized after the individual has started to work in the new job. It is thus convenient to estimate the incomplete likelihood (Dempster *et al.*, 1977) of the individual observation between $t_{i0} + 1$ and t_{i1} , *conditional* on wage/amenity realizations at t_{i0} :

$$\sum_{\theta_1, \theta_2} \pi_{\theta_1, \theta_2} ((y, \mathbf{a})_{it_{i0}}; \Theta_1) \prod_{t=t_{i0}}^{t_{i1}-1} f((y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it} | e_{it}, \theta_1, \theta_2, x_{it}, x_{it+1}; \Theta_2).$$

In this expression, Θ_1 and Θ_2 are sets of parameters. In the rest of this section we shall denote as $\Theta = (\Theta_1, \Theta_2)$ the set of parameters with respect to which the incomplete likelihood is to be maximized. Then, π_{k_1, k_2} are the prior probabilities $\mathbb{P}(\theta_{1it} = k_1, \theta_{2it} = k_2 | (y, \mathbf{a})_{it_{i0}})$, conditional on the wage and amenities at t_{i0} . It is implicitly assumed that both θ_1 and θ_2 are independent of $x_{t_{i0}}$, conditional on the wage and amenities at t_{i0} .

We model θ_1 and θ_2 as two independent random variables (conditional on $(y, \mathbf{a})_{it_{i0}}$) following the Ordered PROBIT specification. Precisely, we assume that there exist two latent variables:

$$\tilde{\theta}_{1i} = \alpha_{1y}^{init} y_{it_{i0}} + \beta_{1y}^{init} \mathbf{a}_{it_{i0}} + u_{1it_{i0}}^{init}, \quad (\text{B2})$$

$$\tilde{\theta}_{2i} = \beta_{2y}^{init} \mathbf{a}_{it_{i0}} + u_{2it_{i0}}^{init}, \quad (\text{B3})$$

and $K_1 + K_2 - 2$ thresholds $s_{1,1}, \dots, s_{1,K_1-1}, s_{2,1}, \dots, s_{2,K_2-1}$, such that:

$$\theta_{1i} = 1 \text{ if } \tilde{\theta}_{1i} < s_{1,1}, \quad \theta_{1i} = 2 \text{ if } s_{1,1} \leq \tilde{\theta}_{1i} < s_{1,2}, \quad \theta_{1i} = K_1 \text{ if } s_{1,K_1-1} \leq \tilde{\theta}_{1i},$$

and:

$$\theta_{2i} = 1 \text{ if } \tilde{\theta}_{2i} < s_{2,1}, \quad \theta_{2i} = 2 \text{ if } s_{2,1} \leq \tilde{\theta}_{2i} < s_{2,2}, \quad \theta_{2i} = K_2 \text{ if } s_{2,K_2-1} \leq \tilde{\theta}_{2i}.$$

Residuals $u_{1it_{i0}}^{init}$ and $u_{2it_{i0}}^{init}$ are independent of each other and covariates, and follow standard normal distributions. We allow the parameters in (B2)-(B3) to be different, in the case where t_{i0} corresponds to the first date of observation for individual i (initial conditions).

Let us partition Θ_2 into subsets corresponding to different transitions. For instance, Θ_2^{yr} corresponds to the parameters in the hedonic equation (2) of 3.1, including the standard deviation of u .

Similarly, Θ_2^z is defined as the set of parameters ruling voluntary mobility decisions, see equation (5) in the same section.

Then we can factorize $f((y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it}^n | e_{it}, \theta_1, \theta_2, x_{it}, x_{it+1}; \Theta_2)$ into:

$$\begin{aligned}
& f(e_{it+1} | e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^e)^{e_{it}=1} \cdot f(e_{it+1} | e_{it} = 0, \theta_1, \theta_2, x_{it}; \Theta_2^n)^{e_{it}=0} \\
& \cdot \left[f(y_{it+1} | \mathbf{a}_{it+1}, z_{it}^n = 1, \theta_1, \theta_2, x_{it}; \Theta_2^{yn}) \cdot \prod_{j=1}^J f(a_{j,it+1} | z_{it}^n = 1, \theta_1, \theta_2, x_{it}; \Theta_2^{jan}) \right]^{z_{it}^n=1} \\
& \cdot f(z_{it}^c | z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^c)^{z_{it}^n=0} \\
& \cdot \left[f(y_{it+1} | \mathbf{a}_{it+1}, z_{it}^c = 1, z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^{yc}) \right. \\
& \cdot \left. \prod_{j=1}^J f(a_{j,it+1} | z_{it}^c = 1, z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^{jac}) \right]^{z_{it}^n=0, z_{it}^c=1} \\
& \cdot f\left(z_{it} = 1, (y, \mathbf{a})_{it+1} | z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^z, \Theta_2^{y*}, \{\Theta_2^{ja*}\}_j\right)^{z_{it}^n=z_{it}^c=0, z_{it}=1} \\
& \cdot f\left(z_{it} = 0 | z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^z, \Theta_2^{y*}, \{\Theta_2^{ja*}\}_j\right)^{z_{it}^n=z_{it}^c=z_{it}=0} \\
& \cdot \left[f(y_{it+1} | \mathbf{a}_{it+1}, z_{it} = z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it+1}; \Theta_2^{yr}) \right. \\
& \cdot \left. \prod_{j=1}^J f(a_{j,it+1} | z_{it} = z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it+1}; \Theta_2^{jar}) \right]^{z_{it}^n=z_{it}^c=z_{it}=0}.
\end{aligned}$$

We can thus rewrite:

$$\begin{aligned}
& f((y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it}^n | e_{it}, \theta_1, \theta_2, x_{it}, x_{it+1}; \Theta_2) \\
= & g((y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it}^n | e_{it}, \theta_1, \theta_2, x_{it}, x_{it+1}; \Theta_2^{-z}) \\
& \cdot \left[f(z_{it} = 0 | z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^z, \Theta_2^{-z})^{z_{it}^n=z_{it}^c=z_{it}=0} \right. \\
& \cdot \left. f(z_{it} = 1, (y, \mathbf{a})_{it+1} | z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^z, \Theta_2^{-z})^{z_{it}^n=z_{it}^c=0, z_{it}=1} \right] \\
= & g((y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it}^n | e_{it}, \theta_1, \theta_2, x_{it}, x_{it+1}; \Theta_2^{-z}) \\
& \cdot h(z_{it}, (y, \mathbf{a})_{it+1} | z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^z, \Theta_2^{-z})
\end{aligned}$$

where Θ_2^{-z} is the subset of Θ_2 containing all the parameters in Θ_2 but those in Θ_2^z , and g is a product of conditional likelihoods.

The joint maximization of the likelihood being cumbersome, we take advantage of this factorization to estimate the model's parameters sequentially. We do so in each M-step of the EM algorithm, as we now explain. Given initial values for the parameters, $\Theta_1^{(s)}$ and $\Theta_2^{(s)}$, the two steps of the sequential EM write as follows.

E-Step: Compute the posterior probabilities of (θ_1, θ_2) given the data

$$X_i = \{(y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it}\}_{t_{i0}+1 \leq t \leq t_{i1}},$$

and $X_{i0} = (y, \mathbf{a})_{it_{i0}}$, and conditional on $x_i = \{x_{it}\}_{t_{i0}+1 \leq t \leq t_{i1}}$, as:

$$p_{\theta_1, \theta_2}(X_i | X_{i0}; \Theta^{(s)}) = \frac{\pi_{\theta_1, \theta_2}(X_{i0}; \Theta_1^{(s)}) f(X_i | \theta_1, \theta_2, x_i; \Theta_2^{(s)})}{\sum_{k_1, k_2} \pi_{k_1, k_2}(X_{i0}; \Theta_1^{(s)}) f(X_i | k_1, k_2, x_i; \Theta_2^{(s)})}.$$

M-step: Update the parameters as follows:

$$\Theta_1^{(s+1)} = \underset{\Theta_1}{\text{Argmax}} \sum_i \sum_{\theta_1, \theta_2} p_{\theta_1, \theta_2}(X_i | X_{i0}; \Theta^{(s)}) \ln \pi_{\theta_1, \theta_2}(X_{i0}; \Theta_1), \quad (\text{B4})$$

$$\begin{aligned} \Theta_2^{(s+1), -z} &= \underset{\Theta_2^{-z}}{\text{Argmax}} \sum_{i, J(i)} \sum_{\theta_1, \theta_2} \left[p_{\theta_1, \theta_2}(X_i | X_{i0}; \Theta^{(s)}) \dots \right. \\ &\quad \left. \dots \sum_{t=t_{i0}}^{t_{i1}-1} \ln g((y, \mathbf{a})_{it+1}, e_{it+1}, z_{it}^c, z_{it} | e_{it}, \theta_1, \theta_2, x_{it}, x_{it+1}; \Theta_2^{-z}) \right], \end{aligned} \quad (\text{B5})$$

$$\begin{aligned} \Theta_2^{(s+1), z} &= \underset{\Theta_2^z}{\text{Argmax}} \sum_{i, J(i)} \sum_{\theta_1, \theta_2} \left[p_{\theta_1, \theta_2}(X_i | X_{i0}; \Theta^{(s)}) \dots \right. \\ &\quad \left. \dots \sum_{t=t_{i0}}^{t_{i1}-1} \ln h(z_{it}, (y, \mathbf{a})_{it+1} | z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_1, \theta_2, x_{it}; \Theta_2^z, \Theta_2^{(s+1), -z}) \right], \end{aligned} \quad (\text{B6})$$

where $J(i)$ indexes the jobs held by individual i .

The M-step of this algorithm differs from the maximization of the complete likelihood as maximization with respect to Θ_2 is achieved sequentially, maximizing the part of the complete likelihood corresponding to voluntary mobility given previous estimates of the offer parameters. This sequential EM algorithm is considered by Arcidiacono and Jones (2003), who prove that it yields consistent estimates of the parameters. They also show that, as this estimator differs from ML, it is not asymptotically efficient. It is to be noted that Arcidiacono and Jones (2003) focus on *unconditional* EM. The extension of their analysis to the case where prior probabilities are conditional on some covariates is however straightforward.

The calculations in the M-step are as follows. In (B4), we estimate the two components of Θ_1 by ordered PROBIT. The dependent variable is the indicator of the sample number, where the modified dataset is the original one duplicated $K_1 K_2$ times. The maximization is weighted by

the groups' posterior probabilities. In (B5), we estimate all sets of parameters corresponding to wage equations, like Θ_2^{yc} for instance, by OLS. We estimate parameters corresponding to amenity equations, like Θ_2^{1ac} , by PROBIT. Lastly, we estimate the set of parameters corresponding to shocks or decisions, like Θ_2^e , also by PROBIT. All regressions are weighted by the posterior probabilities.

To update the parameters ruling voluntary mobility (Θ_2^z), we maximize a likelihood that we present in the next section. The maximization of this likelihood being somewhat longer than the rest of the E and M steps, we proceeded in two stages. In a first stage, we estimated the global model assuming no selection effects in the voluntary mobility process, *i.e.* we replaced (5) with:

$$z_t = \mathbf{1} \left\{ \tilde{\alpha}_z x + \tilde{\beta}_{1z} \theta_1 + \tilde{\beta}_{2z} \theta_2 + \varepsilon_z > 0 \right\},$$

where ε_z is normally distributed, *i.i.d.*, and independent of covariates. We then computed the posterior probabilities for every individual in the sample, and maximized the likelihood corresponding to the voluntary mobility rule, weighted by these probabilities. We finally runned a second sequential EM, allowing for selection in voluntary mobility, and taking the latter estimates as initial conditions for the maximization of the corresponding likelihood. Our experiments showed that the number of iterations necessary for EM to converge numerically was much reduced by proceeding this way.

B.2 Voluntary mobility

The likelihood corresponding to the voluntary mobility rule, for one transition $t/t + 1$, conditional on $(x_{it}, z_{it}^c = z_{it}^n = 0, e_{it} = 1, \theta_{1it} = \theta_1, \theta_{2it} = \theta_2)$ writes:

$$\begin{aligned} \mathcal{L} &= \prod_{i, z_{it}=1} f(z_{it} = 1, (y, \mathbf{a})_{it+1}; \Theta_2^z, \Theta_2^{-z}) \times \prod_{i, z_{it}=0} f(z_{it} = 0; \Theta_2^z, \Theta_2^{-z}), \\ &= \prod_{i, z_{it}=1} f((y, \mathbf{a})_{it+1}; \Theta_2^{-z}) \times \prod_{i, z_{it}=1} f(z_{it} = 1 | (y, \mathbf{a})_{it+1}; \Theta_2^z, \Theta_2^{-z}) \times \prod_{i, z_{it}=0} f(z_{it} = 0; \Theta_2^z, \Theta_2^{-z}). \end{aligned} \tag{B7}$$

Note that the first term in (B7) depends on Θ_2^{-z} only, namely on parameters $\alpha_y^*, \rho^* \dots$. As the maximization in (B6) is with respect to Θ_2^z only, it is equivalent to maximize only the two last terms in (B7).

From equation (5) the latter read:

$$\begin{aligned} \tilde{\mathcal{L}} &= \prod_{i, z_{it}=1} f(z_{it} = 1 | (y, \mathbf{a})_{it+1}; \Theta_2^z, \Theta_2^{-z}) \times \prod_{i, z_{it}=0} f(z_{it} = 0; \Theta_2^z, \Theta_2^{-z}), \\ &= \prod_{i, z_{it}=1} \Phi \left(\frac{y_{it+1} + \boldsymbol{\delta}^* \mathbf{a}_{it+1} - \alpha_z x_{it} - \beta_{1z} \theta_1 - \beta_{2z} \theta_2}{\sigma} \right) \\ &\quad \cdot \prod_{i, z_{it}=0} \left[\sum_{\mathbf{b} \in \{0,1\}^J} \left(\prod_{j=1}^J \Phi \left((-1)^{b_j} (\alpha_{j,a}^* x_{it} + \beta_{j,1a}^* \theta_1 + \beta_{j,2a}^* \theta_2) \right) \right) \right. \\ &\quad \left. \cdot \Phi \left(\frac{\alpha_y^* x_{it} + \beta_y^* \theta_1 + (\boldsymbol{\rho}^* + \boldsymbol{\delta}^*) \mathbf{b} - \alpha_z x_{it} - \beta_{1z} \theta_1 - \beta_{2z} \theta_2}{\sigma_z} \right) \right]. \end{aligned}$$

where we dropped the conditioning variables for simplicity. Recall that the parameters of wage and amenity offers are supposedly known at this stage of the EM algorithm. For instance, they are equal

to the parameters of the wage and amenity equations posterior to constrained job change. Lastly, to find initial conditions in the first M-step of the algorithm we performed a PROBIT regression of z_t on x , weighted by the posterior probabilities.

B.3 Inference

Lastly, we turn to the estimation of asymptotic standard errors. Slightly modifying the moment conditions derived by Arcidiacono and Jones (2003) to account for the conditioning, we obtain, from (B4)-(B6):

$$\begin{aligned} \mathbb{E} \left(\sum_{\theta_1, \theta_2} p_{\theta_1, \theta_2}(X|(y, \mathbf{a})_{t_0}; \Theta) \frac{\partial}{\partial \Theta_1} \ln \pi_{\theta_1, \theta_2}((y, \mathbf{a})_{t_0}; \Theta_1) \right) &= 0, \\ \mathbb{E} \left(\sum_{\theta_1, \theta_2} p_{\theta_1, \theta_2}(X|(y, \mathbf{a})_{t_0}; \Theta) \frac{\partial}{\partial \Theta_2^{-z}} \ln g((y, \mathbf{a}), e, z^c, z|e, \theta_1, \theta_2, x; \Theta_2^{-z}) \right) &= 0, \\ \mathbb{E} \left(\sum_{\theta_1, \theta_2} p_{\theta_1, \theta_2}(X|(y, \mathbf{a})_{t_0}; \Theta) \frac{\partial}{\partial \Theta_2^z} \ln h(z, (y, \mathbf{a})|z^c = z = 0, e = 1, \theta_1, \theta_2, x; \Theta_2^z, \Theta_2^{-z}) \right) &= 0. \end{aligned}$$

Arcidiacono and Jones (2003) propose to estimate the asymptotic variance-covariance matrix by using the usual GMM formula. However, this formula involves first-order derivatives of the moment conditions; that is: second derivatives of the (sequential) likelihood. In models allowing for many parameters, the calculation of these quantities can be cumbersome. In our case, indeed, we found second derivatives of the likelihoods to be long to compute. With more than 200 parameters in the model, the computation of the matrix of second derivatives requires 20000 evaluations (numerical evaluations of the likelihood, or analytical computations of second derivatives).

To circumvent this problem, we use the *generalized information matrix equality* derived by Newey and McFadden (1994), p. 2163. To see how this works in the case of a mixture of sequential likelihoods, let us rewrite the above system of moment equations in compact form:

$$\mathbb{E} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) = 0, \tag{B8}$$

$$\mathbb{E} \left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) = 0, \tag{B9}$$

$$\mathbb{E} \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) = 0, \tag{B10}$$

where:

$$f(y|x; \pi, \alpha, \beta) = \sum_k \pi_k(x; \pi) g_k(y|x; \alpha) h_k(y|x; \alpha, \beta),$$

is the incomplete likelihood, and

$$p_k = p(k|y, x; \pi, \alpha, \beta) = \frac{\pi_k(x; \pi) g_k(y|x; \alpha) h_k(y|x; \alpha, \beta)}{f(y|x; \pi, \alpha, \beta)},$$

is the posterior probability of group k given the data. Vector parameters α, β, π are mutually exclusive.

As f is a mixture of sequential, or *partial* likelihoods in the sense of Cox (1975), the following equality holds:

$$\int g_k(y|x; \alpha_1) h_k(y|x; \alpha_2, \beta) dy = 1, \quad (\text{B11})$$

for all k , all x and for *all* $(\alpha_1, \alpha_2, \beta)$. This point is crucial, as it is used in the consistency proof of the partial likelihood estimator. Moreover, it implies that the moment conditions (B8)-(B10) are satisfied over the whole parameter space, as noticed by Newey and McFadden (1994) in the general context of GMM estimators.

To see why, let us consider (B9). Differentiating (B11) with respect to α_1 yields, for *all* parameters, and under suitable regularity conditions:

$$\begin{aligned} \int \frac{\partial \ln g_k(y|x; \alpha_1)}{\partial \alpha_1} g_k(y|x; \alpha_1) h_k(y|x; \alpha_2, \beta) dy &= \int \frac{\partial g_k(y|x; \alpha_1)}{\partial \alpha_1} h_k(y|x; \alpha_2, \beta) dy, \\ &= \frac{\partial}{\partial \alpha_1} \int g_k(y|x; \alpha_1) h_k(y|x; \alpha_2, \beta) dy, \\ &= 0. \end{aligned}$$

It follows that:

$$\begin{aligned} \int \sum_k p(k|y, x; \pi, \alpha_1, \alpha_2, \beta) \frac{\partial \ln g_k(y|x; \alpha_1)}{\partial \alpha_1} f(y|x; \pi, \alpha_1, \alpha_2, \beta) dy &= \\ \sum_k \pi_k(x; \pi) \int \frac{\partial \ln g_k(y|x; \alpha_1)}{\partial \alpha_1} g_k(y|x; \alpha_1) h_k(y|x; \alpha_2, \beta) dy &= 0, \end{aligned}$$

for *all* parameters, at not only at true values.

The same argument applies to (B10). Moreover, for all parameters:

$$\begin{aligned} &\int \sum_k p(k|y, x; \pi, \alpha_1, \alpha_2, \beta) \frac{\partial \ln \pi_k(x; \pi)}{\partial \pi} f(y|x; \pi, \alpha_1, \alpha_2, \beta) dy \\ &= \sum_k \frac{\partial \pi_k(x; \pi)}{\partial \pi} \int g_k(y|x; \alpha_1) h_k(y|x; \alpha_2, \beta) dy \\ &= \frac{\partial \sum_k \pi_k(x; \pi)}{\partial \pi} = 0. \end{aligned}$$

Let us now consider the first term in the matrix of second derivatives. From the previous argument we obtain:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \pi} \left(\int \sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} f(y) dy \right), \\ &= \mathbb{E} \left(\frac{\partial}{\partial \pi} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \right) + \int \sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \frac{\partial f(y)'}{\partial \pi} dy, \\ &= \mathbb{E} \left(\frac{\partial}{\partial \pi} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \right) + \mathbb{E} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \frac{\partial \ln f(y)'}{\partial \pi} \right), \end{aligned}$$

where from now on all functions and derivatives are evaluated at true values. Now, notice that:

$$\begin{aligned}\frac{\partial f(y)}{\partial \pi} &= \frac{\partial}{\partial \pi} \left(\sum_k \pi_k g_k h_k \right), \\ &= \sum_k \frac{\partial \ln \pi_k}{\partial \pi} \pi_k g_k h_k, \\ &= \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) f(y).\end{aligned}$$

It thus follows that:

$$\mathbb{E} \left(\frac{\partial}{\partial \pi} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right)' \right).$$

We can derive similar expressions for the eight other terms by using the same methodology. Namely, we find after some calculation:

$$\mathbb{E} \left(\frac{\partial}{\partial \alpha} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \left(\sum_k p_k \frac{\partial \ln(g_k h_k)}{\partial \alpha} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \beta} \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right) \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \pi} \left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \alpha} \left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) \left(\sum_k p_k \frac{\partial \ln(g_k h_k)}{\partial \alpha} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \beta} \left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln g_k}{\partial \alpha} \right) \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \pi} \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) \left(\sum_k p_k \frac{\partial \ln \pi_k}{\partial \pi} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \alpha} \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) \left(\sum_k p_k \frac{\partial \ln(g_k h_k)}{\partial \alpha} \right)' \right),$$

$$\mathbb{E} \left(\frac{\partial}{\partial \beta} \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) \right) = -\mathbb{E} \left(\left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right) \left(\sum_k p_k \frac{\partial \ln h_k}{\partial \beta} \right)' \right).$$

It is clear from these expressions that the information matrix equality is not satisfied if h_k depends in a non-trivial manner on α , *i.e.* if the sequential EM is not equivalent to ML. However, the computation of the matrix of cross-products of first derivatives involves roughly the same number of calculations as the computation of the matrix composed of the above terms.

In problems where the number of parameters is large and the (sequential) likelihood is not straightforward to compute, this idea provides a fast alternative way to compute standard errors.