DISCUSSION PAPER SERIES

No. 8841

JOB CHARACTERISTICS AND LABOR TURNOVER: ASSESSING THE ROLE OF PREFERENCES AND OPPORTUNITIES IN TEACHER MOBILITY

Stéphane Bonhomme, Grégory Jolivet and Edwin Leuven

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8841.asp**

ISSN 0265-8003

JOB CHARACTERISTICS AND LABOR TURNOVER: ASSESSING THE ROLE OF PREFERENCES AND OPPORTUNITIES IN TEACHER MOBILITY

Stéphane Bonhomme, Centre for Monetary and Financial Studies (CEMFI)
Grégory Jolivet, University of Bristol
Edwin Leuven, University of Oslo and CEPR

Discussion Paper No. 8841
February 2012

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8841

February 2012

ABSTRACT

Job Characteristics and Labor Turnover: Assessing the Role of Preferences and Opportunities in Teacher Mobility*

Abstract Job characteristics can affect worker turnover through their effect on utility and through their effect on outside job opportunities. We separately identify and estimate the roles of these two channels. Our method exploits information on job changes and relies on an augmented sample selection correction. Taking our approach to an exhaustive register of Dutch primary school teachers, and using arguably plausible exclusion restrictions, we show a detailed picture of preferences for school characteristics. We also study how preference estimates may be biased when ignoring information on job opportunities and discuss the implications for the analysis of teacher turnover.

JEL Classification: C34, C36, J40, J62 and J63
Keywords: compensating differentials, labour turnover, sample selection and teaching labour markets

Stéphane Bonhomme
CEMFI
Casado del Alisal,5
28014 MADRID
SPAIN

Email: bonhomme@cemfi.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=165703

Grégory Jolivet
University of Bristol
Department of Economics
Room 3B11
8 Woodland Road
Bristol BS8 1TN

Email: Gregory.Jolivet@bristol.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159849

Edwin Leuven
Department of Economics
University of Oslo
Postboks 1095 Blindern
0317 Oslo
NORWAY

Email: edwin.leuven@econ.uio.no

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=146953

* We thank Marty West, Frank Windmeijer, and numerous seminar audiences for excellent feedback and discussion.

Submitted 30 January 2012

1 Introduction

The study of labor turnover plays a central role in labor market analysis.[1] More specifically, a large literature has studied how the determinants of job quit decisions relate to wages and wage dynamics (Topel and Ward, 1992), or more generally to job satisfaction (Freeman, 1978; Akerlof et al., 1988). A standard approach in the literature consists in modeling the job quit probability as a function of the characteristics of the current job. These characteristics can affect a worker's decision to leave her job through two channels: preferences (their effect on the worker's utility) and job offers (their effects on the worker's outside job opportunities). The main contribution of this paper is to separately identify and estimate the role of these two channels in worker turnover.

Job quit probabilities contain information on workers' preferences for job characteristics. A commonly used preference measure is the marginal rate of substitution between two job attributes. If one of these attributes is the wage, this rate becomes the Marginal Willingness to Pay (MWP hereafter) for a given amenity. Recognizing the difficulties to estimate MWP from hedonic wage regressions,[2] Gronberg and Reed (1994) argued that, within the context of a standard job search model, MWP may be directly estimated as the (relative) coefficients of amenities in a job turnover regression. Their insight is that the job hazard rate depends on wages and amenities only through the worker's utility function, so the more a worker is satisfied with her job, the less likely she is to leave it.

In this paper, we show that estimates from job turnover regressions may not solely reflect workers' preferences for job attributes. The intuitive argument is that a worker may be staying in her job not because she likes it, but because this job reduces her access to attractive job opportunities. In the context of a job search model this will happen if, unlike in Gronberg and Reed (1994), the arrival rate or the distribution of job offers depends on the characteristics of the current job. In this case, job characteristics will not only affect turnover through workers'

[1] Although this paper focuses on the determinants of job turnover, this topic has also received considerable attention in microeconomic theory, see Burdett (1978) or Jovanovic (1979), and in macroeconomics, see Hall (1972), to cite only a few references.

[2] Hedonic wage regressions consist in regressing wages on amenities and a variety of controls (f.e. Thaler and Rosen, 1975). It was recognized early on that failing to account for unobserved individual heterogeneity in wage regressions could result in preference estimates that are severely biased (f.e. Brown, 1980, Hwang et al., 1992). More recently Hwang et al. (1998) and Gronberg and Reed (1994) showed that even in the absence of confounding unobserved heterogeneity, search frictions prevent one from identifying workers MWP directly from cross-sectional wage/amenity correlations.

preferences for jobs, but also because these jobs affect outside opportunities.[3]

The method that we propose separates these two determinants of job turnover and recovers workers' preferences as well as the correlations between current and outside job characteristics. This not only corrects MWP estimates for a potential bias, but also produces a new set of results on the relationship between characteristics of the current job and the outside job opportunities. Disentangling preferences from opportunities is important from a policy perspective: while preferences are relevant for workers' welfare, the relationship between job characteristics and job opportunities is related to sorting and allocation issues that may be influenced by the policy maker.[4] Overall, our approach thus leads to a better understanding of the channels through which job characteristics affect turnover.

Our approach exploits information on job-to-job transitions. In the simple theoretical framework that we use as a motivation, we let worker's decision to move from one job to another depend on current and outside job characteristics. However, unlike Gronberg and Reed (1994), we let outside job offers and current job characteristics be correlated. To recover the model's parameters, we show that this problem is formally equivalent to the standard sample selection problem in econometrics (Heckman, 1976). This is because job characteristics posterior to a job change are selected within the set of available job opportunities. Identification of this type of models can then be achieved if one or several determinants of the mobility decision (cost shifters) can be excluded from the job offer equations. We show that one such exclusion restriction is sufficient to identify the distribution of outside job's characteristics, and that workers' preferences may then be recovered in a second step by "differencing out" the effect of job characteristics on job opportunities.

Our proof of identification does not rely on parametric assumptions. However, in the presence of a large set of job attributes (ten, in our application), nonparametric estimation is likely to be subject to a severe curse of dimensionality. To deal with this, we impose a linear index structure

[3] We shall discuss potential sources of correlation between current and outside job characteristics when presenting our application on teacher turnover. In the context of a job search model where the wage is the only job characteristic, a structural source of correlation can be found in Postel-Vinay and Robin (2002). If an outside firm competes with a worker's current employer over, say, wages, the most productive of these two firms will outbid the other. If the worker changes job, the wage at the new firm will thus be equal to the highest wage that her previous employer could afford. This means that the worker's outside job offer depends on her current firm's productivity.

[4] Sorting on the teacher labor market is an important question (Rivkin et al., 2005). However, due to data limitations (the absence of data on teacher quality) we will not be able to document sorting patterns in this paper.

on workers' utility and job characteristics equations. When parametric assumptions such as normality are made, estimation simply consists in applying a multi-outcome version of Heckman (1979)'s two-step selection correction method, augmented with a third step to recover workers' preferences. This simple estimation method may easily be extended in various directions. In particular, we show how to allow for individual-specific unobserved heterogeneity in the mobility and outcomes equations. Building on Wooldridge (1995), this can be done by applying our method period by period. In addition, we also propose an easy-to-implement semiparametric version of our estimator that is based on Newey (2009).

We apply our approach to an exhaustive administrative data set of primary school teachers in the Netherlands, in a setting where wages are rigid, and therefore where other characteristics are likely to influence teacher mobility. We estimate teachers' preferences for a large number of attributes, including the percentage of disadvantaged students and school quality. There is a long standing interest in the estimation of teachers' preferences for school characteristics, as knowing the determinants of school change decisions may help policy makers set the right incentives to work in disadvantaged schools (Antos and Rosen, 1975; Hanushek et al., 2004).

In the context of teacher labor markets, there are reasons to expect that current job characteristics affect school turnover through job opportunities, in addition to their effect via teachers' preferences. For example, some school characteristics such as school quality may directly affect student intakes, and thus the school's demand for teachers. In addition, teachers' current job attributes may also constrain outside opportunities: teachers working in schools in poorer neighborhoods may find it more difficult to get a job offer from a school in a more affluent neighborhood, or private schools may prefer to hire teachers who already worked in private schools before.

The validity of our approach relies on the presence of convincing exclusion restrictions. We use two excluded covariates in our empirical work. The first one is based on an interaction between demographic shocks and funding rules which lead to shocks to schools' staff budget. The second excluded covariate is based on the fertility of teachers' colleagues, which affects the school's demand for teachers. In both cases we argue that a teacher's outside job opportunities are unlikely to be influenced by these variables. Moreover, having two excluded covariates yields overidentification conditions that provide evidence on the joint validity of our exclusion restrictions.

Our estimates of teachers' preferences show that the main school characteristics driving teacher mobility between schools are the proportion of disadvantaged pupils, the pupils/teacher ratio, the support/teaching staff ratio, and teaching hours. According to our estimates, Dutch teachers also value school quality, measured as an average test score rank. In addition, we find qualitatively similar estimates when allowing for teacher unobserved heterogeneity, and in a semiparametric specification that is robust to non-normality. Comparing our results with an approach that ignores information on outside job characteristics, we find that the latter does a relatively good job in highlighting which amenities matter for teachers but that estimates based on job turnover alone can be biased in terms of magnitude. The differences between our method and the job duration approach are driven by significant correlations between the characteristics of current and outside jobs which differ markedly across amenities.

Since our approach also delivers estimates of the correlations between current and outside characteristics, it provides a new set of results relevant for the analysis of worker turnover which, as far as we know, has not yet been reported. For example, we find that teachers working in schools with larger proportions of students with low-educated parents are less likely to draw an offer from a school where this proportion is small. To put these results in perspective, we then decompose the effect of job attributes on turnover into two parts: a "preference" effect and a "job opportunities" effect. As outside and current job characteristics are positively correlated, these two effects go in opposite directions. In particular our results suggest that, if turnover was driven only by teachers' preferences, the effect of school quality or the proportion of disadvantaged minority students on job turnover would be substantially *stronger*.

We are not the first to argue that duration-based methods may provide biased estimates of workers' preferences. For example, Boyd et al. (2005) note that job transition probabilities reflect not only a teacher's choice to transfer, but also her opportunities to do so. Boyd et al. (2010) analyze teacher and school preferences separately thanks to a rich data set on the centralized transfer request system in New York City (see also Boyd et al., 2003). An important advantage of the approach we propose in this paper is that it is widely applicable and can be used to analyze labor markets where there is no centralized application system, such as teacher labor markets in many European countries or – more generally – non-teacher labor markets. All one needs to implement our approach is a standard labor force survey with information on amenities and a reliable exclusion restriction.[5]

The paper is organized as follows. In section 2 we introduce the framework and derive a nonparametric identification result for teachers' preferences. In section 3 we describe our three-step estimation strategy. The application to teacher turnover starts in section 4 where we discuss the Dutch teacher market and describe the data. In section 5 we present our estimates of teachers' preferences and study the role of individual preferences and job opportunities in teacher turnover. In section 6 we perform several robustness checks. Finally, section 7 concludes.

2 Identifying workers' preferences using job-to-job mobility

In this section, we derive a nonparametric identification result for workers' preferences, using data on job turnover and job characteristics before and *after* job change.[6] Our approach is general, but in light of our application we shall refer to the teacher labor market in order to illustrate the model and assumptions.

We are interested in identifying workers' preferences over a set of J job attributes which we denote as $A = (a_1, ..., a_J)$. Specifically, let X denote teachers' characteristics, and let $V(A,X)$ be the value of a job. The marginal willingness to trade amenity a_j for amenity a_k is defined as:[7]

$$\text{MWP}_{jk}(A,X) = \frac{\partial V(A,X)/\partial a_j}{\partial V(A,X)/\partial a_k}, \quad \text{for } j \neq k.$$

The marginal willingness to trade is the variation in a_k needed to keep the value of the job constant when a_j increases marginally. Hence, it measures the worker's relative preferences for two job characteristics. When a_k is the wage, MWP_{jk} is the marginal willingness to pay for a_j. We use the notation MWP throughout the paper, although the "numeraire" a_k is not the wage

[5]Plausible exclusion restrictions may also be available in other labor markets. For example, Gibbons and Katz (1992) and Dustmann and Meghir (2005) use plant closure as an exogenous shock on workers' mobility.

[6]Note that standard job turnover regressions do not require data on job attributes posterior to a job move. This is thus an additional data requirement of our approach.

[7]Job attributes are assumed continuous in this discussion. In the empirical analysis, only one out of the ten job characteristics is discretely distributed (the public school dummy).

in our empirical application. The aim of this section is to provide conditions under which the preference parameters $\text{MWP}_{jk}(\cdot)$ are identified, for each (j, k) pair of job attributes.

Our approach relies on job-to-job transitions as a source of identification for individual preferences. Suppose that, at a given point in time, an alternative job with characteristics A^*, and value $V(A^*, X)$, is available to the worker. In the following we refer to alternative jobs as "outside jobs" or "job offers", indistinctly. Suppose also that the worker decides to move if:

$$V(A^*, X) > V(A, X) + c(X, Z), \tag{1}$$

where $c(\cdot)$ is a mobility cost. For example, $c(\cdot)$ could reflect the current school's demand for teachers, or monetary/psychic costs associated with changing job. Note that this representation is quite general. If workers receive multiple job offers in a period, $V(A^*, X)$ may be interpreted as the value of the best alternative.

According to the mobility rule (1), workers weigh the various attributes of their job in proportion to their preferences when deciding whether to change job or to stay in their current job.[8] However, exploiting (1) for identifying preferences is not direct, as the characteristics A^* of an alternative job are not observed in the data if the worker chooses to remain in her job. To address this selection problem, we rely on the availability of exogenous cost shifters Z which are not related to outside job offers. Our main identifying assumption is as follows.

Assumption 1. *The characteristics A^* of outside jobs are statistically independent of the cost shifters Z, conditionally on the current job's amenities A and teacher characteristics X.*

Assumption 1 allows the characteristics of current and outside jobs to be correlated. The cost shifters Z play the role of standard excluded regressors. The validity of the resulting exclusion restriction must be carefully studied. We shall provide an extensive discussion of our choice of excluded covariates in the empirical section. In addition, in order to strengthen the exclusion restriction, we shall allow for additional controls and for unobserved teacher heterogeneity.

In addition, we shall make the following assumption on mobility costs.

[8] Note that the function $V(A,X)$ is the expected value, not the instantaneous utility $u(A,X)$, attached to a job. As a result, $\text{MW}P_{jk}(A,X)$ may not be equal to the ratio $\frac{\partial u(A,X)/\partial a_j}{\partial u(A,X)/\partial a_k}$. Under the assumptions of a standard search model, these two quantities are equal. However they may differ if a worker's search environment depends on her current job characteristics. Hence $\text{MW}P_{jk}(A,X)$ may not be interpreted as a primitive preference parameter.

Assumption 2. *$c(X,Z)$ has an absolutely continuous distribution given (A^*, A, X, Z). In addition, $c(X,Z)$ is independent of A given $(A^*, V(A,X), X, Z)$.*

Assumption 2 states that mobility costs are independent of current job characteristics, conditionally on amenity offers, the current job value, individual heterogeneity, and cost determinants. In particular, the latter include the cost shifters Z. In the application, Z will capture the fact that a school's demand for teachers may correlate with the school's characteristics. Regarding individual heterogeneity, we will control for a range of teacher covariates, and in a robustness check for teacher unobserved heterogeneity.

The following identification result is shown in the appendix.

Theorem 1. *Let Assumptions 1 and 2 hold. In addition, suppose that the technical assumption A1 in the appendix is satisfied. Then, the marginal willingness to trade $MWP_{jk}(A,X)$ is nonparametrically identified for all $j \neq k$.*

The idea of the proof is that the following ratio of job change probabilities, for two different values of the cost shifters z_1 and z_2, is nonparametrically identified from the data:

$$\frac{\Pr\left(Q=1|A^*,A,X,Z=z_2\right)}{\Pr\left(Q=1|A^*,A,X,Z=z_1\right)},$$

where the distribution of job offers A^* has been "differenced out" due to Assumption 1. As, by Assumption 2 this ratio only depend on A through $V(A,X)$, one can then directly recover the marginal willingness to trade parameters.[9]

Before ending this discussion on identification, note that so far we have ignored that individuals can also decide to exit the labor market. As our data contain no information on labor market outcomes for individuals who leave the teacher labor market, we focus the empirical analysis on individuals who do not leave. For our approach to be consistent, the conditions of Theorem 1 need to hold conditionally on the individual not exiting the teacher labor market. In particular, we are thus assuming that there is no dependence between teaching and non-teaching job opportunities, conditional on the current job's characteristics, teacher characteristics, and

[9]Theorem 1 deals with the identification of workers' preferences. Identification of the conditional distribution of job opportunities (A^* given A and X) may be shown if Z has large support by an identification at infinity argument, as in Heckman (1990). The parametric method that we describe in the next section estimates both preferences and job opportunities.

cost shifters. Note that, as independence is conditional on the current job characteristics, this assumption may still hold if a shock affects the teaching and the non-teaching labor markets in similar ways. In addition, to make our empirical analysis robust to local demand shocks that could induce some correlation between teaching and non-teaching job opportunities (even conditional on the current job's characteristics), we also control for local labor market conditions.

3 Estimation

In this section, we present the econometric model and its estimation. We first focus on the benchmark model with observed covariates, and then extend the model to allow for additive unobserved individual heterogeneity.

3.1 The selection model

In our application, jobs are characterized by ten attributes, which describe the type and composition of the school. Because of data limitations, fully nonparametric estimation of preferences following Theorem 1 faces a severe curse of dimensionality. To deal with this issue, we impose an index structure on the model. Specifically, we assume that the job value V is linear in A and X and that the mobility cost $c(X,Z)$ is the sum of an index linear in X and Z and a stochastic component ν independent of A, X and Z.[10] For individual i at date t, the mobility rule (1) thus takes the form:

$$Q_{it} = 1\left\{\theta(A^*_{it} - A_{it}) + \theta_X X_{it} + \theta_Z Z_{it} + \nu_{it} > 0\right\}. \tag{2}$$

The marginal willingness to pay for the various job attributes can directly be recovered from the vector $\theta = (\theta_1, ..., \theta_J)$, as $\text{MWP}_{jk} = \theta_j/\theta_k$.

Similarly, we model amenity offers as follows:

$$A^*_{it} = \alpha A_{it} + \alpha_X X_{it} + \varepsilon_{it}. \tag{3}$$

Note that (3) is a system of J equations, where J is the number of job attributes. In particular, α is a $J \times J$ matrix of coefficients which plays an important role here, as it measures to which

[10]Note that mobility costs thus satisfy Assumption 2.

extent amenity offers and current amenities are correlated.

We assume that ε_{it} is independent of A_{it}, X_{it}, *and* Z_{it}. Hence, (3) satisfies the exclusion restriction of Assumption 1. In our empirical application we use two excluded regressors. This allows us to obtain overidentifying restrictions implied by the exclusion. In addition, note that the unobservables ν_{it} and ε_{it} are allowed to be correlated.

Combining (2) and (3) we obtain the reduced-form equation:

$$Q_{it} = 1\left\{\psi A_{it} + \psi_X X_{it} + \theta_Z Z_{it} + \eta_{it} > 0\right\}, \tag{4}$$

where $\psi_X = \theta_X + \theta\alpha_X$, and where $\eta_{it} = \nu_{it} + \theta\varepsilon_{it}$ is independent of A_{it}, X_{it}, and Z_{it}.

We thus obtain the following mapping between the reduced-form parameter ψ and the preference parameter θ:

$$\psi = \theta\left(\alpha - I_J\right), \tag{5}$$

where I_J denotes the $J \times J$ identity matrix. Equation (5) shows that ψ is a composite of teachers' preferences (θ) and characteristics of the job offer process (α). This provides a clear separation of the effect of job characteristics on job turnover into a preference effect and a job opportunities effect. Our estimation strategy aims at separately estimating θ and α. We proceed in three simple steps:

Step 1. We estimate the reduced-form parameters ψ in (4). Parametric estimators (such as probit or logit) or semiparametric estimators (e.g., Ichimura, 1993, Klein and Spady, 1993) may be used for this purpose.

In this first step we normalize the variance of η_{it} to one. This means that we recover the vector ψ up to scale. The output of the first step consists of the parameter estimates $\widehat{\psi}$ and the predicted probabilities $\widehat{\Pr}\left(Q_{it} = 1 | A_{it}, X_{it}, Z_{it}\right)$.

Step 2. To estimate α we start by noting that, conditional on job change, we have:

$$\mathbb{E}\left(A_{it}^{*} | A_{it}, X_{it}, Z_{it}, Q_{it} = 1\right) = \alpha A_{it} + \alpha_X X_{it} + \mathbb{E}\left(\varepsilon_{it} | \psi A_{it} + \psi_X X_{it} + \theta_Z Z_{it} + \eta_{it} > 0\right). \tag{6}$$

Therefore, α can be consistently estimated by regressing the job attributes A_{it}^* of teachers who have just moved (that is, for $Q_{it}=1$) on A_{it}, X_{it}, and a flexible function of the estimated job change probability $\widehat{\Pr}(Q_{it}=1|A_{it},X_{it},Z_{it})$. For example, one may use the methods described in Das et al. (2003) or Newey (2009). This step delivers the estimate $\widehat{\alpha}$.

As an example, under normality (6) becomes, for $j=1,...,J$:

$$\mathbb{E}\left(A_{jit}^*|A_{it},X_{it},Z_{it},Q_{it}=1\right)=\alpha A_{it}+\alpha_X X_{it}+\rho_j\sigma_j\lambda\left(\psi A_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right),$$

where σ_j is the standard error of the jth element of ε_{it}, and where ρ_j is its correlation with η_{it}. The function $\lambda(\cdot)$ is the familiar inverse Mills' ratio: $\lambda(\cdot)=\frac{\phi(\cdot)}{\Phi(\cdot)}$, where $\phi(\cdot)$ (resp. $\Phi(\cdot)$) denotes the probability (resp. cumulative) distribution function of the standard normal. In this case, the estimate $\widehat{\alpha}$ thus coincides with the Heckman (1979) two-step estimator.

Step 3. Finally, we estimate the $1\times J$ vector θ as:

$$\widehat{\theta}=\widehat{\psi}\left(\widehat{\alpha}-I_J\right)^{-1}.$$

Note that $\widehat{\theta}$ depends on a scale normalization which does not affect the MWP estimates:

$$\widehat{\text{MWP}}_{jk}=\frac{\widehat{\theta}_j}{\widehat{\theta}_k}.$$

Our three-step estimation method thus consists of a simple selection correction estimator, augmented with a final step where teachers' preferences are recovered. In our empirical application, we specify η and ε as normally distributed variables, so that the first two steps of our estimation follow the standard Heckman (1979) procedure except that we have a multidimensional outcome. In addition, we also present results using a semiparametric specification following Newey (2009). For inference, we use the nonparametric bootstrap, since this conveniently takes into account the multi-step nature of the estimation algorithm and the clustering of the errors at the school level.

Note that the first step of the estimation procedure does not deliver consistent estimates of teachers' MWP in general. This is because, when amenity offers are correlated with current amenities, it is generally the case that:

$$\frac{\widehat{\psi}_j}{\widehat{\psi}_k} \neq \frac{\widehat{\theta}_j}{\widehat{\theta}_k} = \widehat{\text{MWP}}_{jk}.$$

When amenity offers are correlated with current amenities, the first-step coefficients ($\widehat{\psi}$) are composite estimates of preferences ($\widehat{\theta}$) and determinants of job opportunities ($\widehat{\alpha}$). Our selection correction approach allows to separately estimate these two parameters.

3.2 Individual unobserved heterogeneity

The benchmark specification in our empirical analysis accounts for school and teacher characteristics which are observed in the data. To extend our framework and allow for unobserved individual heterogeneity, we introduce heterogeneous intercepts in both the offer equation (3) and the mobility equation (4). To provide a motivation for this specification, note that workers may differ with respect to the mobility costs they incur when changing jobs (in particular, our data do not have information on marital status or children). Also, search costs can be heterogeneous across workers.

An attractive feature of the estimation method presented in the previous subsection is that it can easily be extended to allow for the presence of unobserved heterogeneity. To do so, we build on the approach suggested by Wooldridge (1995) and treat the unobserved intercepts as correlated random effects.[11] We model each individual effect as a linear function of X_{i1}, A_{i1} plus a residual.[12] Formally the individual effects in the offer equation and in the selection equation are modeled as $\beta_A A_{i1} + \beta_X X_{i1} + \beta_i$ and $\mu_A A_{i1} + \mu_X X_{i1} + \mu_i$, respectively, where β_i and μ_i are independent of A_{it}, X_{it}, Z_{it} for $t \geq 1$.

[11] See also Semykina and Wooldridge (2007). Another approach, suggested by Kyriazidou (1997), considers individual fixed effects. A theoretical and empirical comparative analysis of these two methods is conducted in Dustmann and Rochina-Barrachina (2007). In Kyriazidou's approach, the main source of information is provided by individuals who move at least twice while the determinants of mobility vary little or none. We found that our data are not suitable for this type of approach and require us to impose slightly more structure via a correlated random-effects specification.

[12] Wooldridge (1995) suggests a conditioning of individual effects on the whole sequence of regressors, X_{it}, A_{it} for all t. Because the A_{it}'s are not strictly exogenous in our case, we only condition individual effects on the initial values.

The selection model with unobserved individual heterogeneity is thus:

$$A_{it}^{*} = \alpha A_{it} + \alpha_X X_{it} + \beta_A A_{i1} + \beta_X X_{i1} + \widetilde{\varepsilon}_{it}, \tag{7}$$

$$Q_{it} = 1\left\{\psi A_{it} + \psi_X X_{it} + \theta_Z Z_{it} + \mu_A A_{i1} + \mu_X X_{i1} + \widetilde{\eta}_{it} > 0\right\}, \tag{8}$$

where $\widetilde{\varepsilon}_{it} = \beta_i + \varepsilon_{it}$ and $\widetilde{\eta}_{it} = \mu_i + \eta_{it}$. This selection model is almost identical to model (3)-(4). As in Wooldridge (1995), we can thus use the same three-step estimation technique, period by period, to recover the preference parameters.

4 Primary school teachers in the Netherlands

4.1 The Dutch market for primary school teachers

We use our approach to estimate the preferences of primary school teachers in the Netherlands. In this subsection, we present some features of the Dutch education system that are relevant for our analysis.[13] First, there is financial and statutory equality between public and private schools. The latter, which are not governed by a public legal entity, are subject to private law, have discretion in their teaching content and practice (within rules and end goals set by the ministry of education) and can refuse admission to pupils. Otherwise, private schools do not differ from public schools. In particular, both types of schools are publicly funded and cannot charge student fees. Schools are governed by a school board which, for public schools, is the municipal authority. Some school boards administer more than one school.

Primary school teachers must have obtained a teaching certificate. They are qualified to teach all subjects with the exception of sports, arts, and foreign languages which are taught by special teachers. Teachers are employed by the school board which has full discretion in the management of its labor force (within rules set by the ministry of education). However, wage scales are set centrally by the government. Basically, teachers are on a wage ladder and move up one rung every year until they reach the top of the ladder and then move on to the next one (there are three wage scales overall). A teacher's wage is thus an almost deterministic function of her experience, the only source of randomness being that some teachers skip the first rung when they

[13] For a detailed description of the Dutch education system, see Eurydice (2008).

move from one wage ladder to the next. There is no wage compensation for working in a given type of school. This is an important feature as teacher selection between schools is therefore only based on non-wage job characteristics.

The school year runs from August 1^{st} to July 31^{st} of the following calendar year. There is a 6-week holiday during the summer and other shorter holidays throughout the year. Primary schools receive government funding under three budget headings: running costs, accommodation and staff. The latter budget is a function of the number and types of pupils registered in the school on October 1^{st}. Schools funding is driven by a compensatory policy aimed at giving more resources to schools with a larger number of disadvantaged pupils. The scheme is based on weights as follows: a weight of 1.9 is assigned to pupils with a non-Dutch cultural background and whose parents have a low level of education, 1.7 to pupils from traveler families, 1.4 to those living in a children's home or a foster family, 1.25 to children whose parents are Dutch and have a low level of education and 1 to everyone else.[14] Therefore, a school's demand for teachers depends both on the number and on the types of children who register. Below we use changes in this budget (which reflect changes in the pupil population) as an important source of variation in teachers' mobility.

Schools that have more disadvantaged students are allotted more funding for staff. However, they cannot offer a teacher a wage higher than what her experience grants her. Schools can thus spend this additional funding on support staff (increasing the support/teaching staff ratio), on teaching material (e.g. on computers) or on hiring more teachers. There is no class size rule in the Netherlands so schools with large numbers of disadvantaged pupils can hire more teachers and make smaller classes. Schools can thus use their budget to compete for teachers on non-wage job attributes, which motivates our empirical analysis of teacher preferences for these characteristics.

4.2 Data

We use administrative data that contain every contract between a teacher and a primary school in the Netherlands. Merging this register with other data sets on schools, we construct a matched

[14] Our data span over the period 1999-2002. Since then, a new scheme has been introduced in August 2006.

teacher-school panel with one observation per teacher i and year t.[15] We restrict our sample to female teachers since the overwhelming majority (over 80 percent) of primary school teachers in the Netherlands are women. We further consider only teachers whose age is between 20 and 60. There are essentially no teachers younger than 20, and to avoid potentially confounding effects of retirement we cut our sample at age 60.

We have access to data for three school years, from 1999 to 2002. For every teacher one observation per school year is kept, corresponding to her main employment. We assume that this is the observation for the school $s = s(i, t)$ where the teacher has the highest teaching load on December 31st of year t.[16] The choice of December 31st is motivated by an empirical regularity in teacher school-to-school transitions (the vast majority of school changes take place between July and November). The school change indicator Q_{it} equals 1 if $s(i, t) \neq s(i, t + 1)$ and 0 otherwise. The total attrition rate is 14%, which includes retirements.

We have no information on teachers' outcomes if they stop working or take a non-teaching job. We thus abstract from individual decisions to leave the Dutch teacher labor market. We discussed at the end of Section 2 the assumption that allows us to conduct our analysis only for teachers who stay in the market. We assume that non-teaching job offers are not correlated with alternative teaching jobs conditionally on the current teaching job. To control for local labor market conditions and thus reinforce the validity of this assumption, we include four region dummies as well as the regional unemployment rate in levels and changes. We also control for the unemployment insurance rate and vacancy rate at the provincial level (12 provinces).

Figure 1 shows mobility rates by age, which serves as our main proxy for experience. Mobility rates are particularly high and nonlinear at the beginning of teachers' career. For this reason, our controls X_{it} include age in a flexible manner with single year age dummies up to age 25, after which we have dummies for 26-30, 30-39, 40-49 and 50+. We also include the teacher's current wage as an individual characteristic. As we pointed out above, selection across jobs cannot operate through wages because these follow a rigid scheme set by the government.[17]

[15] Since our data set covers the whole country, we do not have the attrition problem faced by studies based on state or district-level data (e.g. Hanushek et al., 2004 or Boyd et al., 2005).

[16] Most teachers work in one school but some arts, sports or foreign language teachers may be employed in several schools. We cannot identify these teachers but we expect them to have smaller teaching loads in each school they work at. Also, we drop observations posterior to an exit from and a re-entry in the teacher labor market.

[17] The wage may thus be interpreted as an additional proxy for teaching experience.

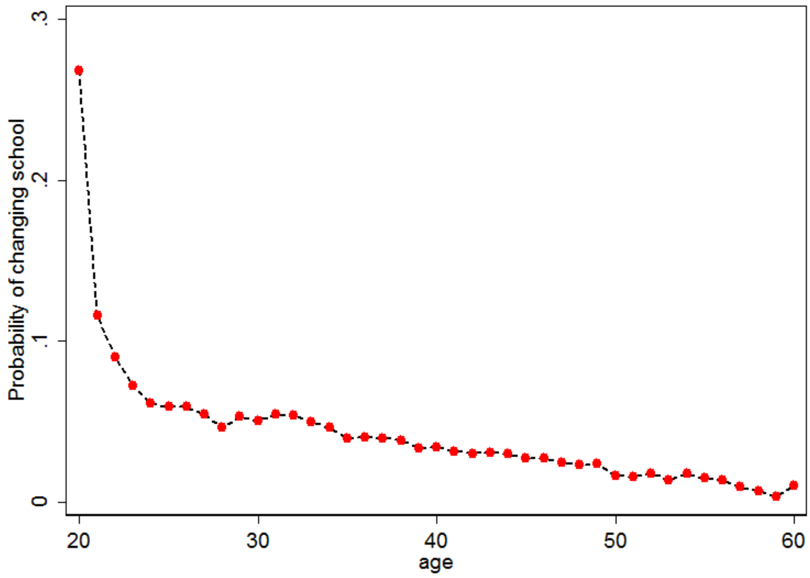


Figure 1. Mobility rates by age

In addition, our data allow us to compute a dummy that equals one if the teacher is on a parental leave during the second semester of year t (i.e. between July and December of year t), as we observe the starting and ending dates as well as the reason of all individual absence spells. We expect this variable to affect a teacher's decision to change job on that year. Finally, we add relative seniority within the school and an extensive set of controls to strengthen our main identifying assumption. We discuss these controls in the next subsection when motivating our exclusion restrictions.

Table 1 shows a set of basic descriptive statistics for the teachers in our sample. On average, teachers are 41 years-old, and in any given school-year in our observation period about 3.4 percent are on parental leave.

Since most job transitions take place between school years it seems natural that teachers care about the pupil population and school attributes of the school-year that is about to start and not of the school-year that just ended. We therefore assume that teachers base their job change decision on the upcoming school-year's attributes of their current school. We also include the current teaching load (year t) to the vector of job attributes A_{it} since we do not observe i's counterfactual teaching load in her old school in case she moves.

Our data contain information on ten job attributes that may enter teachers' value function. These variables are presented in Table 2, where means and standard deviations are computed

Table 1. Descriptive statistics on teachers

Average age (years)	40.5
% $<$ 30 years old	19.3
% 30 – 39 years old	22.6
% 40 – 49 years old	37.8
% $>$ 50 years old	20.3
% Parental leave	3.4
% Movers	3.5
Number of observations	167,550
Number of individuals	70,159

Note: "Parental leave": at least one parental leave. "Movers": at least one school change.

among the population of teachers.

We measure the socio-economic composition of the school through the proportions of disadvantaged children within the school. The proportion of children coming from a disadvantaged ethnic minority is around 16 percent on average. In comparison, the proportion of children coming from disadvantaged native Dutch families is 14 percent.[18]

The pupil-teacher ratio – a proxy for class size – is 20 on average. Population density is defined as the logarithm of the number of inhabitants per square kilometer in the school's municipality. About one third of teachers work in public schools. Notice however that, as discussed above, private schools in the Netherlands are publicly financed. The teaching load is a variable taking values between zero and one, and giving the full-time equivalent number of teaching hours.

The test score variable is computed using a national exit test taken at the end of primary education (in February). We control for the average percentile score within the school. Some schools (14%) do not implement this exam. We drop these schools from our sample for our benchmark estimation results. We have run robustness checks in which we included these schools and dropped the test variable from the list of job attributes. We also account for the average age and gender among teachers within the school. Finally, since our data set contains the employment contracts of non-teaching staff, we compute a variable that gives the number of support staff per full-time teacher within the school.

[18]Disadvantaged minority pupils include all pupils in weighting categories 4 and 5 (see section 3.1). Disadvantaged Dutch pupils are all children in categories 2 and 3. Since there are very few children in categories 3 and 4 we chose to merge them with category 2 and 5, respectively.

Table 2. Descriptive statistics on schools in 2000

	Mean	Std.Dev.
Amenities		
Disadv. minority pupils (fraction)	0.163	0.250
Disadv. Dutch pupils (fraction)	0.138	0.116
Pupil-teacher ratio	20.1	3.8
Teacher hours (in full-time equivalents (FTE))	0.734	0.250
Population density - log(population/km^2)	6.7	1.2
Public school	0.322	0.467
School quality (average percentile test score)	0.493	0.131
Age teachers (average)	41.4	3.8
Female teachers (fraction)	0.824	0.095
Support staff (in FTE as fraction of total staff)	0.098	0.113
Excluded covariates	Z^{bud}	Z^{pl}
Mean	-0.97	0.28
Quantile 25% / 50% / 75%	-11.6/-1.0/9.3	0/0/0.46
$\Pr(Z \leq 0)$	0.53	0.65
$\Pr(Z > 0)$	0.47	0.35
Number of schools	5,758	
Number of teachers per school (FTE)	10.4 (7.6)	
Number of pupils per school	223	

Note: Z^{bud} is the change in a school's budget. Z^{pl} is the total teaching loads of a teacher's colleagues who are on parental leave.

4.3 *Exclusion restrictions*

We rely on two covariates as determinants of job mobility that are excluded from amenity offer equations. The first one is a shock to the pupil population of the school which, given the Dutch institutional context, triggers a shock to the staff budget. The second variable is based on the fertility of a teacher's colleagues. We shall use these variables both separately and jointly as exclusion restrictions. We discuss each of them in turn.

Shocks to the school's budget. A school's staff budget B_t for a given year is computed as the weighted sum of the five groups of pupils registered at the school (the student numbers are taken on October 1st). We define our first excluded regressor as $Z_t^{bud} = B_{t+1} - B_t$, that is the change in the school budget. This variable exploits demographic shocks to the school's student body, both in terms of the number of pupils and of the distribution of types (such as the share of disadvantaged pupils).

The variable Z_t^{bud} captures how a school's demand for teachers changes from one year to the next. Ideally, we would like to know whether the school is closing (or opening) a class but this information is not available in our data. Schools probably smooth the impact of budget shocks to some extent. We expect however that a teacher is more likely to leave (resp. to stay in) a school if Z^{bud} is negative (resp. positive). This is confirmed by our estimation results which show that these shocks are a significant predictor of mobility (see Table B1 in the appendix). A set of descriptive statistics on schools, together with the distribution of Z^{bud} across schools in 2000 is reported in Table 2.

For our exclusion restriction to be valid, the shock Z^{bud} on individual mobility needs to be independent of the characteristics of alternative jobs available to a teacher. To strengthen the exclusion restriction, we control for some potential confounders in addition to the controls presented above. A first potential concern is that a given region is hit by an aggregate demographic shock. In that case, a teacher who has to leave her school may have access to fewer outside jobs because the pupil population in other schools also decreases. We address this concern by controlling for two aggregate demand proxies: the sum of Z^{bud} among all the other schools that are in the same town as school s, and the sum of Z^{bud} among all the schools that are in the same district but not in the same town as school s.[19] A second concern is that, since there are no catchment areas in the Netherlands, a school's pupil population may decrease as a result of it being perceived as a "bad" school. In this case, other schools may be less inclined to hire its former teachers. We account for this possibility by controlling for the rank of a given school in the distribution of school average test scores within the town and the district.

Finally, note that this type of exclusion restriction is not new in the education and labor economics literature. For example, Hoxby (2000) uses similar variation in student populations to study the effect of class size on test scores. Gibbons and Katz (1992) and Dustmann and Meghir (2005) use workers displaced as a result of firm closure to produce an arguably exogenous sample of job changers.

Fertility of colleagues. We also construct a second variable based on fertility. At all dates we observe all the teaching loads in the school, and whether teachers are on a parental leave. For each teacher i we compute Z_i^{pl} as the sum of the teaching loads of her colleagues who are on a parental leave (between July and December).

For the exclusion restriction to be valid we need to assume that the parental leaves of a teacher's colleagues affects her probability to leave the school but not her outside job opportunities. This assumption seems likely to hold but one may think that colleagues' parental leave is too small a phenomenon to have an impact on teacher turnover. It turns out that Z^{pl} is positive for 35% of our observations. Moreover, our estimation results show that colleagues' fertility plays a significant role in a teacher's mobility decision (see Table B1 in the appendix).

5 Results

This section presents our main estimation results. We start with the estimates of teacher preferences for job characteristics. Then, we study the role of job opportunities in the analysis of individual preferences and of labor turnover.

[19]Districts are administrative areas, larger than cities, defined by the Dutch ministry of education.

Table 3. Estimates of preference parameters (θ) and MWP (θ/θ_{PT})

	θ_j		θ_j/θ_{PT}	
Disadv. minority pupils	-0.410***	(0.079)	-43.2***	(15.6)
Disadv. Dutch pupils	-0.146	(0.092)	-15.4	(10.7)
Pupil-teacher ratio (PT)	-0.009**	(0.004)	ref.	
Teacher hours	0.254***	(0.070)	26.8**	(13.0)
Population density	-0.056***	(0.020)	-5.9*	(3.3)
Public school	-0.137**	(0.063)	-14.5	(8.9)
School quality	0.958**	(0.378)	101.0*	(55.8)
Age teachers	0.010***	(0.003)	1.0*	(0.6)
Female teachers	0.253**	(0.110)	26.7*	(16.2)
Support staff	-0.513***	(0.100)	-54.1**	(26.9)

Note: */**/*** statistically significant at the 10/5/1 percent level.

5.1 Teacher preferences

We start by reporting our estimates of the weights of each job characteristic in the value of a job. The first two columns in Table 3 present the estimates of the preference parameters θ. All reported standard errors are bootstrapped using 499 replications and take clustering at the school level into account. The last two columns in the table show the MWP parameter estimates together with their standard errors, using as reference characteristic the pupil-teacher ratio.

The sign and significance of the parameters θ convey information on teacher preferences. The results show the following general picture: teachers are less willing to work in schools with a high proportion of disadvantaged pupils, large classes or a small teaching/support staff ratio. They prefer to work in schools with higher average test scores, a more experienced staff (that is, a higher average age) and a higher proportion of female teachers. They would also rather work more hours and in less densely populated areas.

The proportion of disadvantaged minority pupils is perceived as a disamenity by teachers as its θ coefficient is significantly negative. This is consistent with previous findings in the literature (e.g., Hanushek et al. 2004; Scafidi et al. 2007). Depending on the institutional context and/or data availability these previous studies typically use the proportion of minority pupils and of pupils eligible for subsidized lunch to control for students' socio-economic background. In our data, in contrast, we observe the proportions of pupils with low-educated parents from a Dutch or a non-Dutch background.

Not surprisingly, teachers prefer schools with a smaller pupil-teacher ratio. As we mentioned above, schools with a larger budget cannot post higher wages since wages are set at the national level and are tied to experience. However, schools can hire more teachers and reduce class size in order to attract teachers. Our results in the second column of Table 3 show the changes in pupil-teacher ratio required to compensate for a one unit change in each amenity (MWP). For example, to compensate for a 10 percentage point increase in the proportion of minority students one would need to reduce the pupil-teacher ratio by more than 4 (0.1×43.2).

To put these results in perspective, remember that the weighting in the Dutch budget scheme is such that a school's budget almost doubles when the proportion of disadvantaged minority pupils goes from 0 to 100% (disadvantaged minority pupils have a weight of 1.9 in the funding scheme). In practice we observe that schools where this proportion is 0 have a pupil-teacher ratio of 23 on average whereas schools where this proportion is 100% have a pupil-teacher ratio of 12. It seems that these latter schools use most of their extra budget to reduce class size. This is consistent with our results in the sense that schools try to provide what teachers value. Yet this not enough to fully compensate teachers. A decrease of $23 - 12 = 11$ in the pupil-teacher ratio only compensates for a $100 * 11/43.2 \approx 25$ percentage point increase in the proportion of disadvantaged minority pupils. This simple calculation may explain why schools in disadvantaged areas can have problems retaining their teachers.

The average age of teachers within the school plays a positive and significant role in teachers' utility. This effect is almost equivalent to the effect of reducing the pupil-teacher ratio by one unit. It is difficult to interpret this effect without more detailed data. Since a teacher's age is a good indicator of her experience, one interpretation would be that teachers prefer more experienced colleagues. Another interpretation could be that teaching positions in schools with a more experienced staff are more secure than in other schools. Below we present preference estimates for different age groups that shed more light on this. Also difficult to interpret is the teachers' preferences to work in schools with larger proportions of female teachers (remember that our estimations are run on a sample of female teachers).

Teachers prefer schools with a lower support-to-teaching staff ratio. Support staff can be seen as one of the many indicators of working conditions and we may expect teachers to prefer schools where the support staff is large. Indeed, the survey by Guarino et al. (2006) shows that schools with more administrative support for teachers tend to show a lower teacher attrition rate. Note that Table 3 shows that teachers in the Netherlands prefer the *relative* size of the support staff to be low. In other words, they would rather work in schools that spend their budget on hiring more teachers than on hiring support staff. This result is therefore not inconsistent with previous findings.

We find significant preferences for more teaching hours. We suspect that there may be heterogeneity by age in the preferences for this variable at the extensive margin (two-thirds of the teachers in our sample do not have a full-time contract), something we will investigate below. Population density seems to have a negative effect on teachers' utility. Since wages are set by a fixed national scheme, teachers may prefer less densely populated areas where they would enjoy a higher real wage.

The preference parameter estimate for public schools is negative and borderline significant at the 10% level (p-value = 0.102). As we mentioned in section 4.1, public and private schools in the Netherlands mainly differ with respect to religion and to discretion in the way teaching is organized. Funding, wages and curriculum are the same. It thus seems that the limited differences between the two types of schools stills affect the teachers' utility. However, like teaching hours, preferences for public schools may differ across individuals.

Lastly, we find that school quality as captured by student achievement plays a major role in teachers' preferences, especially when compared with the proportion of disadvantaged minority students (who score on average 1 standard deviation lower than non-minority students). Hanushek et al. (2004) also find that student achievement is one of the drivers of teacher turnover. Scafidi et al. (2007) show that the effect of test scores on turnover may be due to the correlation between this variable and other school characteristics, especially ethnic composition. Our results indicate that in the Netherlands, even when one controls for the education and nationality of students' parents, test scores still play an important role in teachers' preferences for schools.

5.2 *Job opportunities: dependence between current and offered job characteristics*

While teacher preferences for school characteristics are the main targets of the estimation, the analysis of teacher turnover between schools should also account for the search environment that teachers face when making their mobility decisions. Our approach produces new results on the dependence between the current job characteristics of a teacher and her outside job opportunities. The results reported in Table 4 show that many α_{jk} estimates are significantly positive. However, the extent to which current job characteristics affect outside opportunities shows substantial variation across amenities.

If we look at the elements on the diagonal of Table 4, we note that a_j^* significantly and positively depends on a_j for all job attributes. For example, a teacher working in a school with a larger proportion of disadvantaged minority pupils is more likely to have access to an alternative school with a large proportion of similar students. We saw in Table 3 that working in a school with a large proportion of disadvantaged students has a negative effect on teachers' utility. Here we see that this also decreases her chances of moving to a school where this proportion is low.

Table 4 also shows strong dependence between the current teaching load of a teacher and the job offers she receives, between the status (public or private) of the current and the outside schools, and between the population density of the area of the current and the outside schools. The same goes for the average test score in the current school and that in the outside schools. In contrast, for the last three amenities in the table, the dependence between the current job and the job offers is weaker.

The fact that the degree of dependence varies across amenities is relevant for measuring teachers' preferences. Table 5 compares our estimates of preference parameters and MWP with the estimates from a simple turnover regression. In the bottom panel, the MWP are computed taking the pupil teacher ratio as the reference amenity. The MWP estimates are qualitatively similar, so it is fair to say that the Gronberg and Reed (1994) approach paints a relatively accurate picture of teacher preferences in our data. Still, we note differences in MWP estimates between the two methods which, in the case of school quality, can be large (101 vs. 66). We find that this difference is significant at the 5% level for population density, and at the 10% level for public school and school quality.

Table 4. Estimates of job offers parameters (α), equation (7)

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
a_1^*	0.204***	0.024**	-1.809***	0.022	0.194**	0.116***	-0.050***	-0.089	-0.013	0.036***
a_2^*	-0.057*	0.161***	1.617***	-0.008	0.089	0.033	-0.011	-0.489	-0.010	-0.015
a_3^*	0.001	0.001***	0.052***	0.000	0.008*	-0.003	-0.001*	-0.070***	0.001*	0.001
a_4^*	0.059***	-0.005	-0.728***	0.519***	0.211***	0.011	-0.014*	-0.047	-0.002	0.004
a_5^*	0.027***	-0.007***	-0.245***	0.008**	0.477***	0.002	-0.004*	-0.109*	-0.002	0.012***
a_6^*	0.064***	0.002	-1.408***	0.031***	0.083***	0.629***	-0.023***	0.554***	-0.011***	0.003
a_7^*	-0.279***	-0.172***	0.838	0.097	0.000	-0.173	0.317***	5.589***	-0.040	0.008
a_8^*	-0.000	-0.001***	0.008	0.002***	-0.001	0.002	0.001**	0.089***	0.000	0.000
a_9^*	-0.115***	-0.037**	1.433**	0.038	-0.129	0.024	0.046**	0.244	0.050***	-0.021
a_{10}^*	0.112***	-0.010	-1.831***	0.031	0.352***	0.159***	-0.025	-0.033	-0.011	0.172***

Note: */**/*** statistically significant at the 10/5/1 percent level (standard errors available on request). Amenities are abbreviated as follows, a_1: Disadv. minority pupils, a_2: Disadv. Dutch pupils, a_3: Pupil-teacher ratio, a_4: Teacher hours,a_5: Pop. density, a_6: Pub. school, a_7: School quality, a_8: Age teachers, a_9: Female teachers, a_{10}: Support staff.

Table 5. Estimates of preference parameters and MWP: θ vs. ψ

	θ		$-\psi$	
Disadv. minority pupils	-0.410***	(0.079)	-0.247***	(0.054)
Disadv. Dutch pupils	-0.146	(0.092)	-0.109	(0.066)
Pupil-teacher ratio (PT)	-0.009**	(0.004)	-0.007**	(0.003)
Teacher hours	0.254***	(0.070)	0.169***	(0.028)
Population density	-0.056***	(0.020)	-0.012	(0.009)
Public school	-0.137**	(0.063)	-0.021	(0.016)
School quality	0.958**	(0.378)	0.434**	(0.213)
Age teachers	0.010***	(0.003)	0.007***	(0.002)
Female teachers	0.253**	(0.110)	0.130	(0.091)
Support staff	-0.513***	(0.100)	-0.337***	(0.061)
	θ/θ_{PT}		ψ/ψ_{PT}	
Disadv. minority pupils	-43.2***	(15.6)	-37.4**	(14.9)
Disadv. Dutch pupils	-15.4	(10.7)	-16.5	(12.5)
Pupil-teacher ratio (PT)	ref.		ref.	
Teacher hours	26.8**	(13.0)	25.5*	(13.3)
Population density	-5.9*	(3.3)	-1.9	(1.6)
Public school	-14.5	(8.9)	-3.1	(2.7)
School quality	101.0*	(55.8)	65.6	(46.6)
Age teachers	1.0*	(0.6)	1.1	(0.7)
Female teachers	26.7*	(16.2)	19.7	(16.6)
Support staff	-54.1**	(26.9)	-50.8*	(28.6)

Note: */**/*** statistically significant at the 10/5/1 percent level.

Interestingly, the magnitude of the coefficient estimates $\widehat{\theta}$ and $\widehat{\psi}$ in Table 5 is substantially different. While, as we have just seen, this does not necessarily imply large differences in preference estimates (which are defined up to scale), this difference in magnitude has implications for teacher turnover.

5.3 Teacher turnover: a decomposition exercise

In our model, school characteristics affect turnover in two different ways: through preferences and job opportunities. In the rest of this section, we perform a simple decomposition exercise to assess the influence of these two effects on the job turnover probability.

We consider the effect of a change in the vector of amenities, from A_{it} to $\widetilde{A}_{it}$. The difference in the turnover probability is then given by:

$$\begin{aligned}\Delta\left(A_{it},\widetilde{A}_{it}\right) &= \Pr\left(Q_{it}=1|\widetilde{A}_{it},X_{it},Z_{it}\right)-\Pr\left(Q_{it}=1|A_{it},X_{it},Z_{it}\right)\\ &= F\left(\theta\alpha\widetilde{A}_{it}-\theta\widetilde{A}_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right)-F\left(\theta\alpha A_{it}-\theta A_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right),\end{aligned}$$

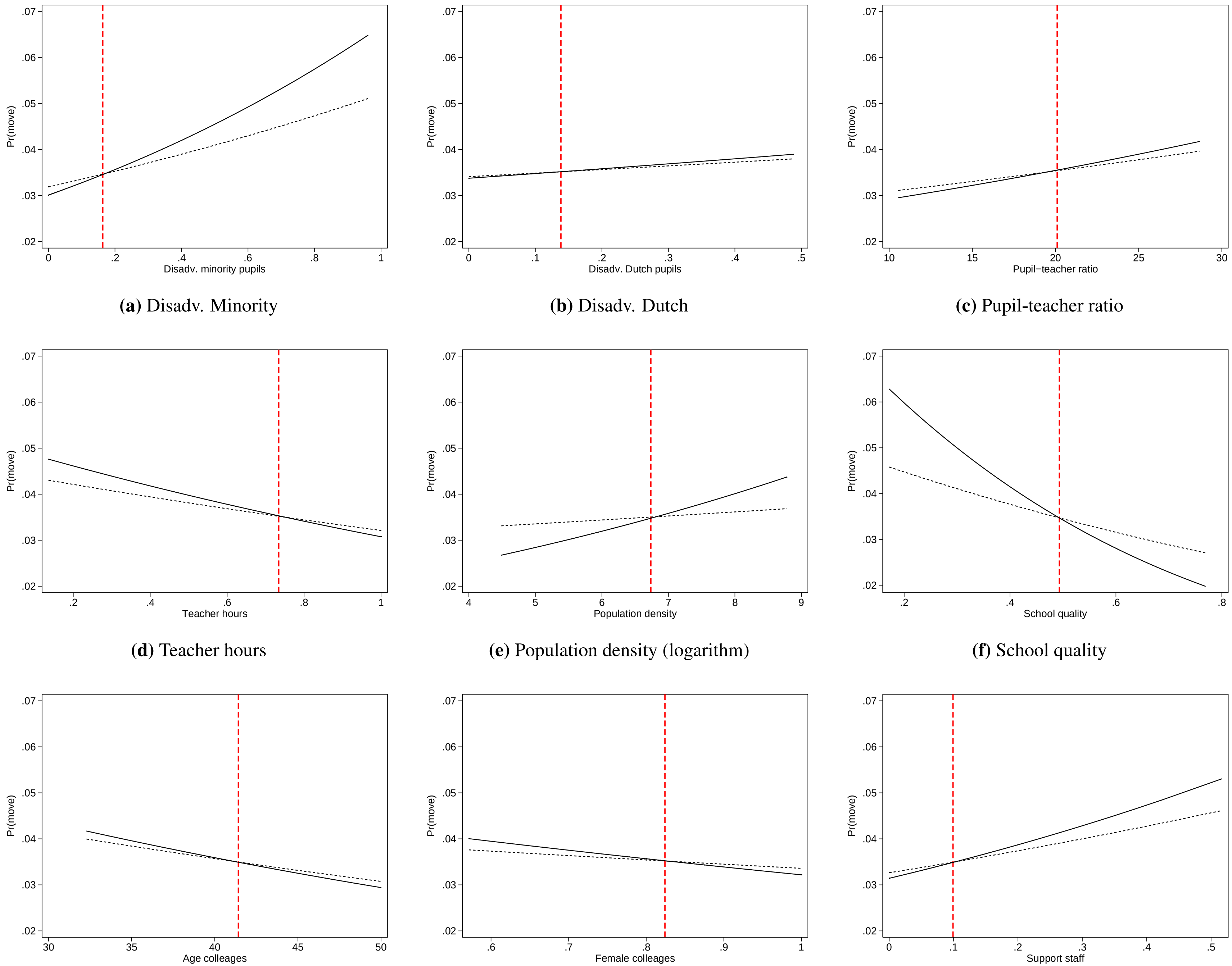
where we have used (4) and (5), and where we have denoted as $F(\cdot)$ the cdf of $-\eta_{it}$ (standard normal in the main specification).

We decompose this difference into a preference effect and a job opportunity effect as follows:

$$\begin{aligned}\Delta\left(A_{it},\widetilde{A}_{it}\right) = &\left[F\left(\theta\alpha A_{it}-\theta\widetilde{A}_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right)-F\left(\theta\alpha A_{it}-\theta A_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right)\right]\\ &+\left[F\left(\theta\alpha\widetilde{A}_{it}-\theta\widetilde{A}_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right)-F\left(\theta\alpha A_{it}-\theta\widetilde{A}_{it}+\psi_X X_{it}+\theta_Z Z_{it}\right)\right]. \quad (9)\end{aligned}$$

We interpret the first term on the right-hand side of (9) as the effect of a change in amenities on teachers' turnover that goes through preferences, holding outside opportunities constant. The second term is interpreted as the effect of a change in current amenities on turnover that goes through outside opportunities, holding teachers' utility constant.

To illustrate these two effects, we consider counterfactual changes in one single amenity at a time. For example, we compute the probability that a worker leaves her school assuming that the proportion of disadvantaged minority pupils takes another value, while keeping the value



(a) Disadv. Minority **(b)** Disadv. Dutch **(c)** Pupil-teacher ratio

(d) Teacher hours **(e)** Population density (logarithm) **(f)** School quality

(g) Age teachers **(h)** Female teachers **(i)** Support staff

Note: Dashed curve = preferences and job opportunities - Solid curve = preferences only. Vertical line = sample mean of amenity.

Figure 2. Probability of leaving school: preferences and opportunities

of other amenities (pupil/teacher ratio, school quality...) equal to the value that we observe in the data. This counterfactual probability, averaged across teachers, is shown in dashed on the top-left graph of Figure 2, as a function of the proportion of disadvantaged pupils. It reflects both the preference effect and the job opportunities effect. On the same graph, the solid curve shows another counterfactual probability, where the job opportunities effect is shut down and thus turnover is only driven by preferences. The same exercise is repeated for each of the other job characteristics (except public/private which is binary) in the other graphs of Figure 2.[20]

The graphs show that the preference and job opportunities effects go in opposite direction. This may be seen by noting that the slope of the dashed curve (total effect) is consistently lower than the slope of the solid curve (preference effect). The reason for this is that offered job characteristics and current job characteristics are positively associated, as shown by the estimates in Table 4.

As an example, the top left graph on the figure shows that teachers are more likely to leave their school if the proportion of disadvantaged minority pupils increases (dashed curve). However, this effect would be much stronger if turnover only depended on preferences, as illustrated by the steeper solid curve. This difference is due to the fact that teachers working in a school where the proportion of disadvantaged pupils is high do not have many opportunities to work in schools where this proportion is low.

The relationship between job opportunities and current amenities varies substantially with the amenity considered. For the proportion of disadvantaged minority pupils the combined effect is roughly half of the preference effect. The difference is also substantial for school quality and to a lesser extent for population density.

6 Robustness checks

In this last section of the paper we start by studying the robustness of our results to the presence of additive unobserved teacher heterogeneity. We then document preference heterogeneity by

[20]Formally, let $A_{it}^{(j,a)}$ be the vector equal to A_{it} except for its jth attribute, equal to a. For each j, the dashed curve in Figure 2 is the average (over i,t) of $\Pr\left(Q_{it}=1|A_{it}^{(j,a)},X_{it},Z_{it}\right)=F\left(\theta\alpha A_{it}^{(j,a)}-\theta A_{it}^{(j,a)}+\psi_X X_{it}+\theta_Z Z_{it}\right)$. The solid curve, net of the job opportunity effect, shows the sample average of $F\left(\theta\alpha A_{it}-\theta A_{it}^{(j,a)}+\psi_X X_{it}+\theta_Z Z_{it}\right)$.

age groups. Finally, we end the section by performing several robustness checks pertaining to our exclusion restrictions and our parametric assumptions.

6.1 Unobserved teacher heterogeneity

The results we showed in the previous section are based on a model specification that does not account for unobserved teacher heterogeneity. As we discussed in section 3.2, it is easy to augment the selection and outcome equations of model (2)-(3) with an individual effect modeled, following Wooldridge (1995), as a linear function of A_{i1} and X_{i1} plus a normally distributed residual. Table 6 reports the estimates of preference parameters θ when the model allows for different specifications of individual unobserved heterogeneity: when the individual intercept is assumed to depend on X_{i1}, on A_{i1}, or on both X_{i1} and A_{i1}. For comparison, we also show in the first column the benchmark estimation results, without unobserved heterogeneity.

The first two columns show that there are essentially no differences between the preference parameters estimated in a model without unobserved heterogeneity and those estimated in a model that allows for individual effects correlated only with individual characteristics' initial values X_{i1}. If we look at the next column, we see that differences do arise when the individual effects are correlated with the job characteristics' first observed values A_{i1}. First, teachers now show more significant and more negative preferences for the proportion of disadvantaged Dutch pupils in the school. We also see a sign reversal for population density. Meanwhile, the preferences are now less precisely estimated. Once we allow for the individual intercept to be correlated with both A_{i1} and X_{i1}, the parameter estimates for teaching hours, population density, public schools and school quality are no longer significant. We note that the point estimates remain quite large and the loss of significance seems to come from a loss of precision due to the large number of parameters we need to introduce to account for unobserved heterogeneity in that specification.

Overall, the general qualitative picture of teacher preferences remains similar when allowing for unobserved heterogeneity. The main difference arise from the parameter estimate associated with the proportion of disadvantaged Dutch pupils, which increases in magnitude as the specification of heterogeneity gets more flexible, and the parameter estimates associated with

Table 6. Estimates of preference parameters (θ) - Individual heterogeneity

	No unobserved heterogeneity		Individual effect correlated with					
			X_1		A_1		X_1, A_1	
Disadv. minority	-0.410***	(0.079)	-0.389***	(0.073)	-0.458**	(0.180)	-0.513***	(0.158)
Disadv. Dutch	-0.146	(0.092)	-0.125	(0.088)	-0.696**	(0.304)	-0.736***	(0.283)
Pup.-teach. ratio	-0.009**	(0.004)	-0.009**	(0.004)	-0.031***	(0.007)	-0.032***	(0.006)
Teacher hours	0.254***	(0.070)	0.261***	(0.068)	0.253*	(0.131)	-0.007	(0.129)
Pop. density	-0.056***	(0.020)	-0.054***	(0.020)	0.103*	(0.059)	0.053	(0.055)
Public school	-0.137**	(0.063)	-0.134**	(0.061)	-0.218	(0.198)	-0.198	(0.178)
School quality	0.958**	(0.378)	0.893**	(0.370)	0.949**	(0.393)	0.698	(0.440)
Age teachers	0.010***	(0.003)	0.009***	(0.003)	0.036***	(0.006)	0.036***	(0.006)
Female teachers	0.253**	(0.110)	0.243**	(0.098)	1.336***	(0.228)	1.374***	(0.234)
Support staff	-0.513***	(0.100)	-0.505***	(0.094)	-0.538***	(0.186)	-0.510***	(0.187)

Note: */**/*** statistically significant at the 10/5/1 percent level.

teaching hours, population density and school quality, which are less precise when heterogeneity is accounted for.

6.2 Preference heterogeneity by age

Teachers may value different amenities at different points in their life. Table 7 presents preference estimates for three age groups that roughly correspond to young school teachers who are starting their career (aged 20 to 26), mid-career teachers who often need to combine work with raising young children (aged 27 to 39), and teachers aged 40 to 60.

As can be seen from Table 7, these estimates stratified by age group are less precise than the ones based on the whole population and shown in Table 3. We nevertheless see some interesting patterns arise. First, teachers in all age groups prefer schools with fewer disadvantaged minority pupils, with young teachers also preferring fewer disadvantaged Dutch pupils. Although not always precisely estimated, teachers' taste for smaller classes and also for school quality is remarkably similar across age groups.

We also observe some differences. With age teachers value working hours differently. In particular younger teachers seem to value more hours, whereas the point estimate suggests that mid-career teachers want to decrease the amount of time they work. We also see that older teachers prefer to work in less densely populated areas, and older teachers also appear to prefer to work in private schools. As to the composition of the school staff, in particular colleagues' age seems to be valued differently by younger and older teachers: the younger teachers prefer younger colleagues whereas the 40 to 60-year-old teachers prefer older colleagues.

The overall picture that emerges is that, although there is some heterogeneity across age groups, teachers' preferences for amenities such as student composition, class size, and school quality seems to vary little with age.

6.3 The exclusion restriction and parametric assumptions

Our estimation strategy relies on two determinants of teachers' mobility to achieve identification. The benchmark estimation results reported above assume that the school budget shock, Z^{bud}, and colleagues' parental leaves, Z^{pl}, enter the job change decision (2) but are excluded from the job

Table 7. Preference estimates (θ) by age group

	20-26		27-39		40-60	
Disadv. minority pupils	-0.308**	(0.155)	-0.329***	(0.111)	-0.527***	(0.166)
Disadv. Dutch pupils	-0.373*	(0.196)	0.005	(0.139)	-0.158	(0.185)
Pupil-teacher ratio	-0.009	(0.009)	-0.011*	(0.006)	-0.010	(0.007)
Teacher hours	0.884***	(0.181)	-0.146	(0.136)	-0.040	(0.200)
Population density	-0.023	(0.039)	-0.032	(0.026)	-0.120**	(0.051)
Public school	0.159	(0.107)	-0.150*	(0.087)	-0.412**	(0.204)
School quality	1.209	(0.783)	1.217*	(0.717)	1.464*	(0.781)
Age teachers	-0.021**	(0.008)	-0.005	(0.005)	0.042***	(0.008)
Female teachers	0.113	(0.238)	0.360**	(0.182)	0.312	(0.195)
Support staff	-0.069	(0.212)	-0.672***	(0.174)	-0.631***	(0.225)

Note: */**/*** statistically significant at the 10/5/1 percent level.

Table 8. Estimates of preference parameters (θ) - Exclusion restriction and Semiparametric model

	Excluded covariates				Semiparametric	
	Budget		Mat. Leave		model	
Disadv. minority pupils	-0.413***	(0.077)	-0.365***	(0.088)	-0.470***	(0.103)
Disadv. Dutch pupils	-0.146	(0.095)	-0.129	(0.084)	-0.173	(0.115)
Pupil-teacher ratio	-0.010**	(0.004)	-0.008*	(0.004)	-0.011**	(0.005)
Teacher hours	0.256***	(0.070)	0.227***	(0.068)	0.327**	(0.144)
Population density	-0.057***	(0.020)	-0.051**	(0.021)	-0.066***	(0.024)
Public school	-0.139**	(0.065)	-0.123**	(0.058)	-0.158**	(0.071)
School quality	0.967**	(0.377)	0.855**	(0.358)	1.109**	(0.451)
Age teachers	0.010***	(0.003)	0.009***	(0.003)	0.011***	(0.004)
Female teachers	0.255**	(0.107)	0.225**	(0.102)	0.295**	(0.123)
Support staff	-0.518***	(0.104)	-0.459***	(0.115)	-0.599***	(0.126)

Note: */**/*** statistically significant at the 10/5/1 percent level.

offer equations (3). Since we only need one exclusion restriction to identify the model we can assess the sensitivity of our results to changes in the set of excluded covariates. The first two columns of Table 8 show the estimates of the preference parameters θ where the set of excluded covariates consists only of Z^{bud} or only of Z^{pl}. The point estimates and their precision remain almost constant across the two specifications, providing support for the validity of our exclusion restriction.

Finally, to further study the impact of the parametric assumptions on the results, we estimate a more flexible (non-normal) specification of the selection model. The mobility decision is now specified as $Q_{it} = 1\{\Lambda(\psi A_{it} + \psi_X X_{it} + \theta_Z Z_{it}) + \nu_{it} > 0\}$, where $\Lambda(\cdot)$ is a second-order polynomial, and where v follows a logistic distribution. In the second step, we replaced the inverse Mills' ratio of Heckman (1979)'s method by a polynomial in the first-stage index.[21] Results, shown in the last column of Table 8, are similar to the benchmark estimates from Table 3. This provides evidence that our estimates are not driven by our parametric assumptions.

7 Conclusion

In this paper we have argued that job characteristics affect worker turnover not only through their preferences, but also through their effect on job opportunities. We have proposed a simple three-step method to estimate these two effects, and shown how it can easily be modified to allow for unobserved worker heterogeneity.

Taking our model to an administrative data set of primary school teachers in the Netherlands, we have obtained estimates of preferences that are qualitatively consistent with earlier results in the literature (e.g., Hanushek et al., 2004, Scafidi et al., 2007), the main school characteristics driving teachers' mobility being the proportion of disadvantaged pupils, the pupil-teacher ratio, the support to teaching staff ratio and teaching hours (although our estimates suggest that preferences are age-dependent for that latter amenity). Our estimates also indicate that Dutch school teachers value school quality. Although existing funding rules double the budget of schools that teach disadvantaged minority children, we have argued that using this budget to reduce pupil-teacher ratios does not fully compensate teachers. One important open question is whether equivalent wage compensation would keep teachers indifferent between teaching disadvantaged and non-disadvantaged pupils.

In addition, we find that the two effects of job characteristics on turnover go in opposite directions, the preference effect being partly (though not completely) offset by the job opportunities effect. For example, we find that the effect of the proportion of disadvantaged minority pupils or school quality on turnover would be substantially stronger if turnover was only driven by teachers' preferences.

The two channels that we identify are both relevant for the policy maker. Knowing teachers' preferences is important for welfare analysis, and to study workers' decisions to enter or leave the

[21] See Newey (2009) for more details.

teacher labor market. Constraints on job opportunities are also worth investigating as they may be influenced by specific policies aimed at changing the allocation of teachers across schools.

In order to assess the effect of a given policy using our framework, one needs to put more economic structure on the correlations between current and outside job characteristics that we have estimated in this paper. One source for these correlations could arise from inter-firm competition, as in Postel-Vinay and Robin (2002). Another reason could be that, through directed search, unobserved heterogeneity drives a worker's preferences as well as her search for outside jobs. On the other side of the market, it could be that employers use the applicant's former job characteristics as a signal of their preferences or of their ability. The recent paper by Boyd et al. (2010) offers a first insight into these issues. It would be interesting to embed our approach into an equilibrium model of the labor market to study how the job offer distribution, which we considered to be exogenous, could result from the interactions of worker and firm preferences.

Another possible extension is to better account for teacher heterogeneity. Allowing for heterogeneity in teacher quality is particularly relevant when studying sorting patterns in the market (Rivkin et al., 2005). Extending our framework to allow for these additional features and applying it to matched teacher-class-school data, or more generally matched worker-firm-productivity data, is an interesting avenue for future work.

References

Akerlof, R., A. Rose, and J. Yellen (1988). Job switching and job satisfaction in the us labor market. *Brooking Papers on Economic Activity 1988*(2), 495–594.

Antos, J. and S. Rosen (1975). Discrimination in the Market for Public School Teachers. *Journal of Econometrics 3*(2), 123–150.

Boyd, D., H. Lankford, S. Loeb, M. Ronfeldt, and J. Wyckoff (2010). The role of teacher quality in retention and hiring: Using applications-to-transfer to uncover preferences of teachers and schools. Unpublished manuscript.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2003). Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Estimating Compensating Differentials in

Imperfect Labor Markets. Working Paper 9878, National Bureau of Economic Research, Cambridge, Mass., USA.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2005). Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students. *American economic review 95*(2), 166–171.

Brown, C. (1980). Equalizing Differences in the Labor Market. *Quarterly Journal of Economics 94*(1), 113–134.

Burdett, K. (1978). A theory of employee job search and quit rates. *American Economic Review 68*(1), 212–220.

Das, M., W. Newey, and F. Vella (2003). Nonparametric Estimation of Sample Selection Models. *Review of Economic Studies 70*(1), 33–58.

Dustmann, C. and C. Meghir (2005). Wages, Experience and Seniority. *Review of Economic Studies 72*(1), 77–108.

Dustmann, C. and M. Rochina-Barrachina (2007). Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equations. *Econometrics Journal 10*(2), 263–293.

Eurydice (2008). Organisation of the Education System in the Netherlands.

Freeman, R. (1978). Job satisfaction as an economic variable. *American Economic Review 68*(2), 135–141.

Gibbons, R. and L. Katz (1992). Does Unmeasured Ability Explain Inter-Industry Wage Differentials. *Review of Economic Studies 59*(3), 515–535.

Gronberg, T. and W. Reed (1994). Estimating workers' marginal willingness to pay for job attributes using duration data. *Journal of Human Resources 29*(3), 911–931.

Guarino, C., L. Santibanez, and G. Daley (2006). A Review of the Research Literature on Teacher Recruitment and Retention. *Review of Educational Research 76*, 173–208.

Hall, R. (1972). Turnover in the labor force. *Brooking Papers on Economic Activity 1972*(3), 709–764.

Hanushek, E., J. Kain, and S. Rivkin (2004). Why public schools lose teachers. *Journal of Human Resources 39*(2), 623–654.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement 5*(4), 475–492.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*, 153–161.

Heckman, J. (1990). Varieties of selection bias. *American Economic Review 80*(2), 313–318.

Hoxby, C. (2000). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *Quarterly Journal of economics 115*(4), 1239–1285.

Hwang, H., D. Mortensen, and W. Reed (1998). Hedonic Wages and Labor Market Search. *Journal of Labor Economics 16*(4), 815–847.

Hwang, H., W. Reed, and C. Hubbard (1992). Compensating Wage Differentials and Unobserved Productivity. *Journal of Political Economy 100*(4), 835–858.

Ichimura, I. (1993). Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-IndexModels. *Journal of Econometrics 58*(1-2), 71–120.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy 87*(5), 972–990.

Klein, R. and R. Spady (1993). An Efficient Semiparametric Estimator for Binary Response Models. *Econometrica 61*(2), 387–421.

Kyriazidou, E. (1997). Estimation of a Panel Data Sample Selection Model. *Econometrica 65*, 1335–1364.

Newey, W. (2009). Two-step series estimation of sample selection models. *Econometrics Journal 12*, S217–S229.

Postel-Vinay, F. and J. Robin (2002). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica 70*, 2295–2350.

Rivkin, S., E. Hanushek, and J. Kain (2005). Teachers, schools, and academic achievement. *Econometrica 73*(2), 417–458.

Scafidi, B., D. Sjoquist, and T. Stinebrickner (2007). Race, Poverty, and Teacher Mobility. *Economics of Education Review 26*(2), 145–159.

Semykina, A. and J. Wooldridge (2007). Estimation of Dynamic Panel Data Models with Sample Selection. Technical report, Mimeo.

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*, pp. 265–300. New York: Columbia University Press.

Topel, R. and M. Ward (1992). Job mobility and the careers of young men. *Quarterly Journal of Economics 107*(2), 439–479.

Wooldridge, J. (1995). Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions. *Journal of Econometrics 68*(1), 115–132.

A Proof of Theorem 1

In addition to Assumptions 1 and 2 we need the following technical assumptions.

Assumption A1. *i)* $f(A^*|A,X) > 0$ *almost surely. ii)* $\Pr(Q=1|A^*,A,X,Z) > 0$ *almost surely. iii) For some* z_1 *and* z_2 *in the support of* Z*:*

$$\frac{\partial}{\partial a_j}\left(\frac{\Pr(Q=1|A^*,A_j=a_j,A_{-j},X,Z=z_2)}{\Pr(Q=1|A^*,A_j=a_j,A_{-j},X,Z=z_1)}\right) \neq 0 \quad a.s., \quad \text{for all } j.$$

Assumption A1*i)* rules out situations where the support of amenity offers depends on current amenities and worker characteristics. A1*ii)* states that for any values of the vector (A^*,A,X,Z) workers have a strictly positive probability of changing job. This assumption has a technical purpose, as our identification proof relies on ratios of job change probabilities to exploit the variation in the excluded variable. Lastly, A1*iii*) is a technical assumption that states that the ratio of job change probabilities for two values of the cost shifter is a non-trivial function of amenities of the current job. As the proof below shows, this ratio can be recovered from the data, so A1*iii*) is testable.

Proof. Let us start by defining:

$$B(A^*, A, X, Z) = \Pr(Q = 1|A, X, Z)\, f(A^*|A, X, Z, Q = 1),$$

where f is a generic notation for a distribution function. Note that $B(A^*, A, X, Z)$ is readily available from the data. Using Bayes' rule we have:

$$\begin{aligned} B(A^*, A, X, Z) &= \Pr(Q = 1|A^*, A, X, Z)\, f(A^*|A, X, Z) \\ &= \Pr(Q = 1|A^*, A, X, Z)\, f(A^*|A, X), \end{aligned}$$

where we have used Assumption 1. Using Assumptions A1*i*) and A1*ii*) we obtain, for z_1 and z_2 in the support of Z:

$$\frac{B(A^*, A, X, z_2)}{B(A^*, A, X, z_1)} = \frac{\Pr(Q = 1|A^*, A, X, Z = z_2)}{\Pr(Q = 1|A^*, A, X, Z = z_1)},$$

where the offer distribution has been "differenced out". Lastly, using (1) we also have:

$$\begin{aligned} \Pr(Q = 1|A^*, A, X, Z) &= \Pr\Big(c(X, Z) < V(A^*, X) - V(A, X) \mid A^*, A, X, Z\Big) \\ &= \Pr\Big(c(X, Z) < V(A^*, X) - V(A, X) \mid A^*, X, Z\Big) \end{aligned}$$

by Assumption 2, which depends on A only through $V(A, X)$. Taking (z_1, z_2) as in Assumptions A1*iii*) thus implies that, almost surely:

$$\begin{aligned} \mathrm{MWP}_{jk}(a, x) &= \frac{\partial V(a, x)/\partial a_j}{\partial V(a, x)/\partial a_k} \\ &= \frac{\frac{\partial}{\partial a_j}\left(\frac{\Pr(Q=1|A^*,A=a,X=x,Z=z_2)}{\Pr(Q=1|A^*,A=a,X=x,Z=z_1)}\right)}{\frac{\partial}{\partial a_k}\left(\frac{\Pr(Q=1|A^*,A=a,X=x,Z=z_2)}{\Pr(Q=1|A^*,A=a,X=x,Z=z_1)}\right)} = \frac{\frac{\partial}{\partial a_j}\left(\frac{B(A^*,a,x,z_2)}{B(A^*,a,x,z_1)}\right)}{\frac{\partial}{\partial a_k}\left(\frac{B(A^*,a,x,z_2)}{B(A^*,a,x,z_1)}\right)}, \end{aligned}$$

so $\mathrm{MWP}_{jk}(a,x)$ is non-parametrically identified for all $j \neq k$.

□

B Additional results

Table B1. Estimated reduced form turnover equation (4)

Amenities:		
Disadv. minority pupils	0.247***	(0.054)
Disadv. Dutch pupils	0.109	(0.066)
Pupil-teacher ratio	0.007**	(0.003)
Teacher hours	-0.169***	(0.028)
Population density	0.012	(0.009)
Public school	0.021	(0.016)
School quality	-0.434**	(0.213)
Age teachers	-0.007***	(0.002)
Female teachers	-0.130	(0.091)
Support staff	0.337***	(0.061)
Individual characteristics:		
Age 21	0.111*	(0.060)
Age 22	0.052	(0.046)
Age 23	-0.013	(0.039)
Age 24	-0.043	(0.038)
Age 25	0.005	(0.036)
Age 20-29	0.620***	(0.038)
Age 30-39	0.497***	(0.027)
Age 40-49	0.270***	(0.023)
ln(wage)	0.026	(0.109)
On maternity leave	-0.674***	(0.054)
Tenure (rank in school)	0.060**	(0.025)
Temporary contract	0.897***	(0.023)
School rank at municipality level	0.010	(0.051)
School rank at district level	0.057	(0.100)
Local labor market controls:		
Sum Z^{bud} at municipality level	0.063*	(0.036)
Sum Z^{bud} at district level	0.022	(0.020)
Region = North	-0.045	(0.071)
Region = South	-0.065*	(0.038)
Region = East	-0.075***	(0.024)
UI rate (Province)	-0.060*	(0.032)
Vacancy rate (Province)	3.453	(2.216)
Unemp. rate (Region)	0.019	(0.037)
Δ Unemp. rate (Region)	-0.104	(0.064)
Exclusion restrictions:		
Z^{bud}	-0.003***	(0.000)
Z^{pl}	-0.043***	(0.014)
Intercept	-2.068**	(0.880)

Note: */**/*** statistically significant at the 10/5/1 percent level.