

ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic
Theory in its Relation to Statistics and Mathematics*

<https://www.econometricsociety.org/>

Econometrica, Vol. 88, No. 3 (May, 2020), 1031–1069

ESTIMATION OF A ROY/SEARCH/COMPENSATING DIFFERENTIAL MODEL OF THE LABOR MARKET

CHRISTOPHER TABER

Department of Economics, University of Wisconsin-Madison

RUNE VEJLIN

Department of Economics and Business Economics, Aarhus University and IZA

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website <http://www.econometricsociety.org> or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

ESTIMATION OF A ROY/SEARCH/COMPENSATING DIFFERENTIAL MODEL OF THE LABOR MARKET

CHRISTOPHER TABER

Department of Economics, University of Wisconsin-Madison

RUNE VEJLIN

Department of Economics and Business Economics, Aarhus University and IZA

In this paper, we develop a model that captures key components of the Roy model, a search model, compensating differentials, and human capital accumulation on-the-job. We establish which components of the model can be non-parametrically identified and which ones cannot. We estimate the model and use it to assess the relative contribution of the different factors for overall wage inequality. We find that variation in premarket skills (the key feature of the Roy model) is the most important component to account for the majority of wage variation. We also demonstrate that there is substantial interaction between the other components, most notably, that the importance of the job match obtained by search frictions varies from around 4% to around 29%, depending on how we account for other components. Inequality due to preferences for non-pecuniary aspects of the job (which leads to compensating differentials) and search are both very important for explaining other features of the data. Search is important for turnover, but so are preferences for non-pecuniary aspects of jobs as one-third of all choices between two jobs would have resulted in a different outcome if the worker only cared about wages.

KEYWORDS: Search, compensating differentials, Roy model, wage inequality.

1. INTRODUCTION

IT IS WELL-KNOWN that variation in wages across observably similar workers is high. There are several competing theories as to why this is the case. Four of the most important models of post-schooling wage determination are the Roy model, the search model, the compensating differentials model, and the human capital model.¹ All four lead to wage heterogeneity. While separating human capital accumulation from the others is quite common, we know remarkably little about the relative importance and interactions of the other three sources of inequality. The goal of this paper is to quantify the effect of each of these factors on overall wage inequality and to investigate how they interact.

The key features of the Roy model are absolute and comparative advantage, in which some workers earn more than others as a result of different skill levels at labor market entry. In the canonical Roy model, workers choose the job for which they achieve the highest level of wages. In search models, workers may have just had poor luck in finding

Christopher Taber: ctaber@ssc.wisc.edu

Rune Vejlin: rvejlin@econ.au.dk

We would like to thank participants at various seminars for helpful comments and suggestions. We especially thank Joe Altonji, Jesper Bagger, Jim Heckman, John Kennan, Rasmus Lentz, Greg Veramendi, and four referees for their extremely helpful comments, and Carl Sanders for his comments and research assistance. Taber thanks the NSF for its support, and Vejlin acknowledges financial support from the Danish Social Sciences Research Council (Grant 09-066745). Appendices A–C are available on the *Econometrica* website, and Appendices D and E are on our own websites: Rune Vejlin: <https://sites.google.com/site/econrunevejlin/home> and Christopher Taber: <https://www.ssc.wisc.edu/~ctaber>. All remaining errors are our own.

¹One large class of models, which built around asymmetric or imperfect information such as Jovanovic (1979) and Altonji and Pierret (2001), is missing. This is a fruitful avenue for future research.

their preferred job. Labor market frictions can lead to heterogeneity in wages for two different reasons. First, some workers may work for firms for which they are a better match and earn higher wages. The second is monopsony. In the framework we use, the bargaining position of the worker depends on their outside option, so two equally skilled workers at the same firm may earn different wages. In a compensating wage differentials model, a worker is willing to be paid less to work on a job that they enjoy more. This means that workers with identical skills and job opportunities can earn different wages. Finally, in a human capital model, workers who have accumulated more human capital while working earn higher wages than less experienced workers due to higher productivity. In short, one worker may have higher wages than another because the individual has (a) more talent at labor market entry (Roy model), (b) had better luck in finding a good job and receiving outside offers (search frictions), (c) chosen a more unpleasant job (compensating differentials), or (d) accumulated more human capital while working (human capital). The goal of this study is to uncover the contribution of these different components and determine how they interact to produce overall wage inequality.

Combining these four prominent theories into one coherent framework and investigating the relative importance in various dimensions is important not only from a theoretical point of view but also from a practical standpoint. An understanding of the underlying causes of inequality is essential for policymakers concerned with wage inequality. Moreover, it is important for deciding the extent to which wage inequality is a problem and potentially how to mitigate it. If the primary driver behind wage inequality is compensating differentials, inequality might be less problematic since it simply reflects different preferences for jobs. If the primary driver is Roy model heterogeneity (productivity differences), search frictions, or human capital, then addressing wage inequality involves a more traditional efficiency-equity tradeoff. However, the appropriate approach depends on the relative importance of the different channels—if the primary source is Roy model inequality, that would suggest focusing on premarket factors (like education), but if human capital or search frictions are the primary drivers, post labor market policies would be more effective.

In this paper, we develop a structural model of wage determination that contains the essential elements of all four models. The model is estimated on Danish matched employer-employee data from 1985 to 2003. We use the estimated parameters to decompose overall wage inequality into the four components in various ways. The four components are captured as skill upon labor market entry (loosely the Roy model), search frictions, heterogeneity in preferences for non-pecuniary aspects of jobs (loosely compensating differentials), and human capital acquired while working. We find that while all four components are contributors to overall wage inequality, variation in premarket skills is the most important one and accounts for between 59% and 82% of wage inequality. The magnitude of the effects depends on the way we implement the decomposition, as there are important interactions between the different components of the model. The most striking component is search, where the monopsony component of search frictions explains roughly 3% of wage variation. The importance of the direct source of search frictions (workers work at different firms due to frictions) varies from around 1% to around 26%.

While search and non-pecuniary aspects are less important for wage inequality than premarket skills, we show that they are both essential for explaining other features of the data. For example, they strongly impact the level of wages, and we show that without them, workers would receive wages that are 0.22 and 0.20 log points higher for search and preferences for non-pecuniary aspects, respectively. Together, they also explain the bulk of the variation in utility across people. Furthermore, preferences for non-pecuniary

aspects are important in determining job choices. Roughly one-third of all choices would be different if workers only cared about wages and not the non-pecuniary aspects of a job. We conclude that all four of these components are important aspects of the labor market in Denmark and should be considered to achieve a full understanding of wage inequality and job turnover.

The paper proceeds as follows. Section 2 briefly discusses the relationship between this study and the previous literature while Section 3 describes the model and the decomposition. Next, Section 4 discusses identification and then Section 5 presents the econometric specification, where obtaining the right data is crucial. Ideally, matched employer-employee data, a long panel on workers, and detailed information on job-to-job transitions are required. Section 6 describes the data and institutional features of the Danish labor market and is followed by Section 7 which presents the auxiliary model used. Finally, Section 8 presents the results, Section 9 considers robustness, and Section 10 concludes.

2. RELATION TO OTHER WORK

There is a vast amount of research on the Roy model, search models, compensating differentials, and human capital acquired on the job. A full review of all of this research is beyond the scope of this paper. See Roy (1951) for the original model or Heckman and Taber (2008) for a discussion on the Roy model. Eckstein and Van den Berg (2007) and Rogerson, Shimer, and Wright (2005) provided a thorough discussion of empirical search models, while Rosen (1987) provided an excellent discussion of compensating differentials models. For post-schooling human capital models, see Weiss (1986) and Heckman, Lochner, and Todd (2006). Rather than engaging in a broad discussion, we focus on the relationship between our work and other important key papers from different literatures.

Two of the most important related literatures were started by Abowd, Kramarz, and Margolis (1999) (henceforth AKM) and Postel-Vinay and Robin (2002), where the former estimated a two-way fixed effect model. Subsequent studies that have arisen using this methodology found that variance in firm effects plays an important role for wages. Important examples are Gruetter and Lalive (2004), Andrews, Gill, Schank, and Upward (2008), Sørensen and Vejlin (2013), Card, Heining, and Kline (2013), and Card, Cardoso, and Kline (2015). More recently, there has been some focus on the limitations of AKM; see, for example, Abowd, McKinney, and Schmutte (2018) and Bonhomme, Lamadon, and Manresa (2019).² The fact that some firms are able to survive while paying lower wages is often attributed to search frictions. However, in our model, both search frictions and compensating differentials could result in differences in average wages across firms. Postel-Vinay and Robin (2002) decomposed wage inequality into a search component, a firm productivity component, and an ability-related component. As such, their research question is highly related to ours. The main components in our model that are not included in theirs are human capital, compensating differentials, and comparative advantage in jobs. Various papers build on the Postel-Vinay and Robin (2002) framework. Cahuc, Postel-Vinay, and Robin (2006) extended it to account for more general bargaining, and Bagger, Fontaine, Postel-Vinay, and Robin (2014) allowed for the accumulation of general human capital while working; see also Bagger and Lentz (2018), Lise, Meghir, and Robin (2016), and Robin (2011) for other applications. Our paper, in essence, builds

²Lentz, Piyapromdee, and Robin (2018) built on this framework and estimated the model using Danish data.

on this literature by adding the non-pecuniary aspects of jobs and much more heterogeneity in premarket skills.

The concept of compensating differentials dates back to [Smith \(1776\)](#). More recently, [Keane and Wolpin \(1997\)](#) initiated important related research by estimating a model that introduces compensating differentials into a model of human capital and Roy model inequality.³ Similarly to us, they found that premarket skills are the main drivers of earnings inequality. However, their model is very different, in that they did not explicitly incorporate search frictions and did not make use of matched employer-employee data, because their focus was on occupations rather than firms. [Becker \(2009\)](#) used a framework similar to ours, in that it incorporates compensating differentials into a search model but focuses more on unemployment insurance than wage inequality. [Dey and Flinn \(2005, 2008\)](#) estimated search models with a particular type of non-wage characteristics: health insurance.

In the vast literature on human capital models, the work by [Huggett, Ventura, and Yaron \(2011\)](#) is most directly related to ours. They estimated a life-cycle model with idiosyncratic shocks to human capital, heterogeneity in ability to learn, initial human capital, and initial wealth. Their focus was quite different from ours in that they focused on differences in lifetime earnings, utility, and wealth. They found that most of the variation is due to differences in initial conditions, which is in accordance with our finding that most of the differences in instantaneous wage variation are due to differences in workers' ability.

The work of [Sullivan and To \(2014\)](#) aligns more closely with our setup, in that they estimated a job search model with a general form of non-wage job characteristics. Although their model includes search and compensating differentials, our papers contain many differences. First, they specified output as being only match-specific, while we allow workers and firms to have constant ability and productivity across matches, in addition to being match-specific. Second, their model is only partial equilibrium in the sense that workers draw a wage and a non-wage component, but there is no negotiation between firms and workers. Finally, [Sorkin \(2018\)](#) distinguished between search and compensating differentials using a type of revealed preference argument similar to ours. Other than that, our models are very different as he specified the wage function in a more reduced form and focused on the distinction between search frictions and compensating differentials in understanding the firm's fixed effect in a wage regression, and mostly disregarded comparative advantage and human capital.⁴ By contrast, we use a firm random effect approach and focus on wage inequality.

The identification results in our paper can be thought of as an extension of the identification of the Roy model from [Heckman and Honoré \(1990\)](#) and, more largely, as being related to identification of selection models described in [Heckman \(1990\)](#); see, for example, [French and Taber \(2011\)](#) for a survey. We add search frictions to this model and consider mobility using panel data. This step relates to the identification of models of firms and workers that includes [Abowd, Kramarz, and Margolis \(1999\)](#) and [Hagedorn, Law, and Manovskii \(2017\)](#). Our model is quite different from that of [Hagedorn, Law, and Manovskii \(2017\)](#), in that they depicted workers skills and firm productivity as only being one-dimensional. [Abowd, Kramarz, and Margolis \(1999\)](#) essentially assumed orthogonal match effects. More recent work by [Bonhomme, Lamadon, and Manresa \(2019\)](#)

³[Hoffman \(2018\)](#) also used similar elements, including search frictions, into the basic [Keane and Wolpin \(1997\)](#) framework. [Hoffman \(2018\)](#) also looked at earnings inequality using German data and found that initial skills are very important.

⁴He also used Longitudinal Employer-Household Dynamics data, which are quarterly. While his paper studies a larger country, his data make distinguishing job-to-job transitions from job-to non-employment-to-job transitions much more difficult than in our data.

assumes a finite number of firm types, which is similar to our framework, and deals with selection. A major difference is that wages in our model are determined by bargaining, which means that there are many different wages that a worker can receive at the same firm. We identify this distribution.

3. THE MODEL

The model is in continuous time, and wages are determined similarly to Cahuc, Postel-Vinay, and Robin (2006), Dey and Flinn (2005), and Bagger et al. (2014).⁵ We formally treat agents in our model as infinitely lived, though, in practice, we think of it as a life-cycle model, where the date of retirement is far away enough that it can be ignored.

Basic Environment: Firms and Workers

There are a finite number of job types indexed $j = 1, \dots, J$ with $j = 0$ denoting non-employment. We assume that the economy consists of a very large number of potential employers who offer the jobs. Each job is tied to an establishment in the data in the sense that each establishment offers the worker one of the job types, and a worker must switch establishments in order to switch job type.

A substantive difference between our paper and most of the search literature is that we assume that there is a finite number of job types. In estimation, this is similar to the finite number of types of individuals, as is common in much of the structural labor literature; see, for example, Heckman and Singer (1984), or Keane and Wolpin (1997), and similar to Bonhomme, Lamadon, and Manresa (2019) and Lentz, Piyapromdee, and Robin (2018).

We allow for a large number of individuals in the economy. In the data, we have N individuals indexed by $i = 1, \dots, N$. In describing the model, we focus on a generic individual in the data and use the i subscript to make it clear which variables vary across individuals. The key elements of the model are:

- The productivity of individual i at a job type j at labor market entry is π_{ij} .
- The flow utility of individual i at a job type j with human capital ψ_h and human capital rental rate R is $u_{ij}(R\psi_h)$.

We are very flexible in both of these dimensions and allow for both absolute and comparative advantage. The fact that utility depends on j is also an important aspect of our model, which accommodates compensating differentials: workers care about jobs above and beyond the wage that they earn or expectations about future wages.

We treat production on the job as if it is linear, so hiring worker i does not crowd out the hiring of other workers or affect the productivity of current workers. The value of a vacancy is zero, and the flow value is the productivity of the worker on the job (defined below). We also assume that workers and firms have complete information, in the sense that they know all distributions and the specific utilities and productivities upon meeting each other.

We do not allow for borrowing or lending. Individuals begin their working lives non-employed.

⁵Note that we assume bargaining over wages as opposed to wage posting. Hall and Krueger (2012) showed that there is mixed evidence regarding the wage determination process. In their survey, around one-third of all workers report having bargained over their wage. Another third reports that they had precise information about the wage before meeting the employer, which is a sign of wage posting.

Learning by Doing

To distinguish this type of human capital from human capital acquired prior to labor market entry, we typically label it learning by doing (LBD) human capital.

- LBD human capital takes on a discrete set of values ψ_0, \dots, ψ_H .
- When individuals are employed, LBD human capital appreciates randomly to the next level (ψ_h to ψ_{h+1}) at rate λ_h .
- $\pi_{ij}\psi_h$ is the productivity of worker i at job type j when the worker has LBD human capital level h .
- We normalize $\psi_0 = 1$, so π_{ij} is the productivity at labor market entry.
- The flow utility of worker i with LBD human capital h when the worker is non-employed is u_{i0h} .

As LBD human capital is not a major focus of this paper, we have kept the model simple, and human capital is general in the traditional sense that it is fully transferable across jobs. In [Bagger et al. \(2014\)](#), human capital evolves deterministically, while in our model it is stochastic. We do not view this difference as important but, if anything, our specification tends to make human capital more important in explaining wage variation. Note that human capital does not accumulate when people are not working.⁶ We also allow the flow value of utility in non-employment to potentially depend on human capital to accommodate home production.

Job Destruction and Arrival Rates

We model frictions in the market as follows:

- A job of type j arrives at rate:
 - λ_j^n for non-employed workers.
 - λ_j^e for employed workers.
- A job is destroyed at rate δ_i .
- With probability P^* a worker receives another offer immediately after job destruction without having to enter non-employment. The worker can either accept the job, or reject it and enter non-employment.
 - The relative probability of receiving a job from each job type is the same as for non-employment.

The inclusion of an immediate offer into the model reconciles involuntary job-to-job transitions, which are shown in survey data which ask whether the employment relationship was terminated on the employer's initiative. If this is the case, we view it as involuntary.

One can see that we allow destruction rates to vary across individuals, and arrival rates to vary across firms. As a practical matter, allowing destruction rates to vary across jobs would make our model much more complicated, which is why the main specification does not do this. Our initial exploration of the importance of this indicated that this is not a major issue.⁷ Not allowing the job arrival rates to vary with i is essentially a normalization, as [Flinn and Heckman \(1982\)](#) showed, the arrival rate cannot be identified separately from the fraction of jobs below the reservation wage. Since we allow the utility of non-employment to vary across individuals, arrival rates cannot.

⁶Allowing human capital to depreciate when out of the labor force could easily be embedded into this framework, though we have no reason to think it would substantially alter our counterfactuals.

⁷In particular, we split firm histories by the median year in which they were in the sample and calculate the correlation between the firm-specific job-destruction hazards before and after the median year. The correlation in the data was very low, and matching this in the model suggested that variation in δ across j was not important for our results. Specifically, we want the structural parameter, δ_j , to pick up something systematic about the establishment, rather than being large due to a single large layoff. We calculate the hazards both before and

Wage Determination

Following [Bagger et al. \(2014\)](#), a key aspect of this model is that, when a worker receives an outside offer, wages are determined by a form of generalized Nash bargaining between the two firms. In this case, the object of negotiation is the human capital rental rate. That is, the employer and worker agree on a rental rate, R , which is the price per unit of LBD human capital, which is fixed until the next negotiation. This means that when human capital is augmented, the wage is not renegotiated, but automatically rises from $R\psi_h$ to $R\psi_{h+1}$.⁸ This form of wage setting leads to efficient turnover from the perspective of the two firms and the worker.

We define $V_{ijh}(R)$ to be the value function for worker i , with the rental rate R , working in job $j > 0$, and having LBD human capital level h . Workers who are non-employed have value function V_{i0h} . We let V_{i0h}^* denote the value function immediately after a match is destroyed. The difference between V_{i0h}^* and V_{i0h} is that the former incorporates the possibility of receiving an offer immediately. We will derive these explicitly below, but first show how they are used to determine wages.

A match is formed if $V_{ijh}(\pi_{ij}) > V_{i0h}$. Note that this is efficient in the sense that the joint surplus between the worker and the matched firm is maximized. A major issue is determination of the wage. Both the worker and the firm would be willing to form a match with rental rate R as long as $V_{i0h} \leq V_{ijh}(R) \leq V_{ijh}(\pi_{ij})$. The issue is that there are many such values of R . We denote the equilibrium rental rate as $R_{ij\ell h_0}$ for worker i , at current job j , with the best outside option ℓ , and units of human capital, h_0 , at the time of negotiation. We need to introduce the new notation h_0 , because human capital can evolve on the job, while the wage is not renegotiated. Thus, individual i at job j , with the best outside option l and current human capital h , but human capital level h_0 when the wage was last renegotiated, will have wage $R_{ij\ell h_0}\psi_h$.

after the median year using the following (see Appendix C for notation):

$$h_f^{bm} \equiv \frac{\sum_{i=1}^N \sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} 1[f_{i\ell j} = f]1[e_{i\ell j} < 52 * \text{med}(N(f))]1[j = J_{i\ell}]}{\sum_{i=1}^N \sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} 1[f_{i\ell j} = f][\min(e_{i\ell j}, 52 * \text{med}(N(f))) - b_{i\ell j}]}$$

$$h_f^{am} \equiv \frac{\sum_{i=1}^N \sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} 1[f_{i\ell j} = f]1[e_{i\ell j} \geq 52 * \text{med}(N(f))]1[j = J_{i\ell}]}{\sum_{i=1}^N \sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} 1[f_{i\ell j} = f][e_{i\ell j} - \max(b_{i\ell j}, 52 * \text{med}(N(f)))]}$$

It turns out that the covariance between these hazards is 1.53×10^{-6} . We also calculate $\text{cov}(h_{-i\ell j}, \tilde{w}_{i\ell j}) = -5.33 \times 10^{-5}$ and $\text{cov}(h_{-i\ell j}, \tilde{S}_{i\ell j}) = 3.63 \times 10^{-5}$ with $\tilde{S}_{i\ell j}, \tilde{w}_{i\ell j}$ is defined in Appendix C. For computational reasons, we were unable to re-estimate the model, but we did obtain a reasonable fit by hand, and we find that these covariances are sufficiently small enough to make very little difference to the model.

⁸Another natural way to do this would be to assume that the absolute wage is agreed upon and that the return to human capital only accrues when workers receive outside offers. This would be similar in nature to [Postel-Vinay and Turon \(2010\)](#). Human capital does not play a major role in this analysis, and we expect the alternative contracting to give very similar results, though to be somewhat messier to implement.

The negotiated rental rate, R_{ij0h_0} , for a worker coming out of non-employment is determined by

$$V_{ijh_0}(R_{ij0h_0}) = \beta V_{ijh_0}(\pi_{ij}) + (1 - \beta)V_{i0h_0}, \tag{1}$$

where β is the worker’s bargaining power.⁹ Since the bargaining position for a worker who has just been laid off is non-employment, this will also be the rental rate for workers who experience job destruction but are then immediately hired by a new firm, j . Note that when $\beta = 1$, the worker has all of the bargaining power and extracts full rent, $R_{ij0h_0} = \pi_{ij}$. When $\beta = 0$, the firm has all of the bargaining power and pays the value of R_{ij0h_0} , which makes the worker indifferent to accepting the offer or staying non-employed.

Now suppose that worker i , with human capital h and current rental rate $R_{ij\ell h_0}$, is working in job type j and receives an outside offer from job type q . As in Postel-Vinay and Robin (2002), one of three things can happen. First, the new job offer could dominate the old one, $V_{iqh}(\pi_{iq}) > V_{ijh}(\pi_{ij})$. In this case, the worker will switch to the new job and the new rental rate, R_{iqjh} , will be determined by

$$V_{iqh}(R_{iqjh}) = \beta V_{iqh}(\pi_{iq}) + (1 - \beta)V_{ijh}(\pi_{ij}). \tag{2}$$

If $V_{iqh}(\pi_{iq}) \leq V_{ijh}(\pi_{ij})$, then the worker has the option to renegotiate the wage. If renegotiation is chosen, the new rental rate will be determined by

$$V_{ijh}(R_{ijqh}) = \beta V_{ijh}(\pi_{ij}) + (1 - \beta)V_{iqh}(\pi_{iq}). \tag{3}$$

If $V_{ijh}(R_{ij\ell h_0}) < V_{iqh}(\pi_{iq})$, the worker will want to renegotiate.

Note that our notation is a bit loose, in that we use the notation $R_{ij\ell h_0}$ to denote the rental rate that worker i , with human capital ψ_{h_0} at the time of negotiation, would receive from job type j when their outside option was job type ℓ . As equations (2) and (3) show, the result is the same regardless of whether the worker started at job type ℓ and moved to j , or if the worker started at j and then used an outside offer from job type ℓ to renegotiate their wage.

Solving the Model

To solve the model, we need to calculate the value functions $V_{ijh}(R)$ and V_{i0} , as there are no closed form solutions for the wage as in Cahuc, Postel-Vinay, and Robin (2006) and Bagger et al. (2014).

It is convenient to define

$$\Lambda_{ijh}^e(R) \equiv \sum_{\{\ell: V_{ijh}(R) < V_{i\ell h}(\pi_{i\ell})\}} \lambda_{\ell}^e,$$

$$\Lambda_{i0h}^n \equiv \sum_{\{\ell: V_{i\ell h}(\pi_{i\ell}) > V_{i0h}\}} \lambda_{\ell}^n,$$

⁹We do not explicitly derive this from a bargaining game like Cahuc, Postel-Vinay, and Robin (2006) but treat it as a functional form assumption: wages are indeterminate, and this will give a wage that both parties will agree to. It has the nice property that if there is surplus in the match, when $\beta = 1$, the worker will get all of the surplus, and when $\beta = 0$, the firm gets all of the surplus, and when $0 < \beta < 1$, the surplus will be split between them.

which, respectively, for employed and non-employed workers are the sums of arrival rates that will lead to some reaction, either renegotiation or switching jobs. Thus, for worker i , with human capital h , who is currently employed at job type j , with rental rate R , this is the arrival rate of some outside offer that will change behavior.

We can write the value function for worker i , with human capital h , who is currently employed at job j , with rental rate R , as:¹⁰

$$\begin{aligned} & (\rho + \delta_i + \lambda_h + \Lambda_{ijh}^e(R))V_{ijh}(R) \\ &= u_{ij}(R\psi_h) + \left(\sum_{\{\ell:V_{ijh}(R) < V_{i\ell h}(\pi_{i\ell}) \leq V_{ijh}(\pi_{ij})\}} \lambda_\ell^e [\beta V_{ijh}(\pi_{ij}) + (1 - \beta)V_{i\ell h}(\pi_{i\ell})] \right) \\ &+ \left(\sum_{\{\ell:V_{i\ell h}(\pi_{i\ell}) < V_{ijh}(\pi_{ij})\}} \lambda_\ell^e [\beta V_{i\ell h}(\pi_{i\ell}) + (1 - \beta)V_{ijh}(\pi_{ij})] \right) \\ &+ \delta_i V_{i0h}^* + \lambda_h \max(V_{ijh+1}(R), V_{i0h+1}). \end{aligned}$$

Consider the different components on the right-hand side of this equation. The first, $u_{ij}(R\psi_h)$, is the flow utility that the worker receives. The second component denotes outside offers that will lead the worker to renegotiate their wage but ultimately stay at their current job. If $V_{i\ell h}(\pi_{i\ell}) \leq V_{ijh}(R)$, the outside offer will not be useful for renegotiating, and if $V_{ijh}(\pi_{ij}) < V_{i\ell h}(\pi_{i\ell})$, then the worker will leave and take the next job. The component in brackets in this expression represents the value function of the renegotiated rental rate, as described in equation (2). The next term denotes outside offers that lead the worker to leave the current job. Again, the component in brackets denotes the value function under the negotiated rental rate, as described in equation (3). The final two terms, $\delta_i V_{i0h}^*$ and $\lambda_h \max(V_{ijh+1}(R), V_{i0h+1})$, represent the events in which the worker is laid off (including the potential value of an immediate offer) and in which human capital is augmented, respectively.

When $h = H$, we get an identical expression, except that it no longer contains the possibility of augmenting human capital:

$$\begin{aligned} & (\rho + \delta_i + \Lambda_{ijH}^e(R))V_{ijH}(R) \\ &= u_{ij}(R\psi_H) + \left(\sum_{\{\ell:V_{ijH}(R) < V_{i\ell H}(\pi_{i\ell}) \leq V_{ijH}(\pi_{ij})\}} \lambda_\ell^e [\beta V_{ijH}(\pi_{ij}) + (1 - \beta)V_{i\ell H}(\pi_{i\ell})] \right) \\ &+ \left(\sum_{\{\ell:V_{i\ell H}(\pi_{i\ell}) < V_{ijH}(\pi_{ij})\}} \lambda_\ell^e [\beta V_{i\ell H}(\pi_{i\ell}) + (1 - \beta)V_{ijH}(\pi_{ij})] \right) + \delta_i V_{i0H}^*. \end{aligned}$$

The value function for a non-employed worker is much simpler:

$$(\rho + \Lambda_{i0h}^n)V_{i0h} = u_{i0h} + \sum_{\{\ell:V_{i\ell h}(\pi_{i\ell}) > V_{i0h}\}} \lambda_\ell^n [\beta V_{i\ell h}(\pi_{i\ell}) + (1 - \beta)V_{i0h}].$$

¹⁰Note that increases in human capital could trigger a decision to quit. We abstract from this case when estimating the model but then verify that this constraint does not bind in our simulations for any individual at any time under our estimated parameter values.

The first term is the flow utility and the second denotes the outcome in which an offer is received, which dominates non-employment. The term in brackets represents the value function under the renegotiated rate.

Finally, the value function for workers immediately after their match is destroyed is

$$V_{i0}^*(h) = P^* \frac{\sum_{\ell: V_{i\ell h}(\pi_{i\ell}) > V_{i0h}} \lambda_{\ell}^n [\beta V_{i\ell h}(\pi_{i\ell}) + (1 - \beta)V_{i0h}]}{\sum_{\ell} \lambda_{\ell}^n} + \left(1 - P^* \frac{\Lambda_{i0t}^n}{\sum_{\ell} \lambda_{\ell}^n}\right) V_{i0h}.$$

The first term is the result of an acceptable offer, while the second is the result of either no offer or an unacceptable offer.

This is the full model. Note that unlike many other search models, job ladders in this model are individual-specific. This is due to both comparative advantages and preferences for non-pecuniary aspects of the job. This makes the model computationally harder to solve than, for example, [Bagger et al. \(2014\)](#). Obviously, there are many other features in the labor market that we have disregarded. This is intentional. Our goal is not to devise the most complicated model that is computationally feasible, but rather to devise the simplest model that captures the essence of our four models and allows us to distinguish between them.

4. IDENTIFICATION

In this section, we discuss non-parametric identification of our model and show which aspects of the model can and cannot be identified. Doing so is important in that we cannot credibly simulate counterfactuals that are not identified from the data. We respect this in our simulations, except in two explicitly noted cases. The proofs and additional results are available in Appendix D on our websites.¹¹

Proving general non-parametric identification of the model when the number of job types, J , is very large seems overly tedious, which is why we focus on a simpler case to illustrate identification. We simplify the model as follows:

ASSUMPTION 1: (a) *There are two job types ($J = 2$), which we label A and B .*

(b) *LBD human capital takes on two values ($h = \{0, 1\}$).*

(c) *LBD human capital does not change the preference ordering across jobs and non-employment.*

(d) *If a worker is indifferent in terms of two options, we assume that (a) when the choice is between working and not working, they work, (b) when it is between an A and B firm, they choose the A firm, and (c) when a worker receives an outside offer from an identical firm type, they stay at the current firm.*

This simplification of the model is very much in the spirit of [Heckman and Honoré \(1990\)](#), who focused on a Roy model with two choices, as did [Roy \(1951\)](#). As in their model, we fully expect all of the basic results to extend beyond this simple case.¹²

¹¹Rune Vejlin: <https://sites.google.com/site/econrunevejlin/home>. Christopher Taber: <https://www.ssc.wisc.edu/~ctaber>.

¹²The length of the panel required for identification, however, would also increase substantially.

4.1. *Transition Components*

Workers begin their working life non-employed and we assume that we observe all data on them from time 0 to \bar{T} . We consider two different versions of the model. In one, we allow for heterogeneity in δ_i . However, in an effort to identify the full distribution of δ_i non-parametrically, we have only been able to show identification when we observe all completed spells, which is why we assume that we can let \bar{T} go to some arbitrarily large number. This is an unattractive assumption, so for our main case, we show that if we do not allow for heterogeneity in δ_i , then we can identify the model with a finite \bar{T} . Appendix D, available on our websites, shows the heterogeneous δ_i /infinite \bar{T} case. We continue here with the homogeneous δ case, where we make the following assumptions:

ASSUMPTION 2: *The econometrician*

- (a) *observes the history of job-type spells, with start and stop dates, and can identify the job type, j , at each job until point \bar{T} ,*
- (b) *does not record job switches within job type.*

ASSUMPTION 3: *There is no heterogeneity in δ_i .*

Assumption 2(a) assumes that we can identify the job type. Clearly, we do not actually observe these directly from the data, but we could use the procedure in [Bonhomme, Lamadon, and Manresa \(2019\)](#) to estimate the job types, under the assumption of many workers per firm. In the empirical specification below, we do not specifically identify establishments but choose the model appropriately for consistent estimates.

Assumption 2(b) implies that when a worker switches from an A type job to another A type job, the econometrician does not see this switch, and consequently does not use this information for identification. We use this restriction to tie our hands, because, in reality, (a) workers would be indifferent in terms of the two jobs, which is why we do not have a theory of movement, and (b) the finite number of jobs is intended to be a simplification and approximation of the real world, not something that provides identifying information. Of course, these switches would actually be seen in the data, but our goal is to show that we can identify the model without relying on these types of transitions.

Without loss of generality, we can categorize workers by their preference ordering using C_i , as¹³

$$C_i \equiv \begin{cases} 0 & \text{if } V_{iAh}(\pi_{iA}) < V_{i0h} \text{ and } V_{iBh}(\pi_{iB}) < V_{i0h}, \\ B0 & \text{if } V_{iAh}(\pi_{iA}) < V_{i0h} \leq V_{iBh}(\pi_{iB}), \\ A0 & \text{if } V_{iBh}(\pi_{iB}) < V_{i0h} \leq V_{iAh}(\pi_{iA}), \\ BA & \text{if } V_{i0h} \leq V_{iAh}(\pi_{iA}) < V_{iBh}(\pi_{iB}), \\ AB & \text{if } V_{i0h} < V_{iBh}(\pi_{iB}) \leq V_{iAh}(\pi_{iA}). \end{cases}$$

Since there are a finite number of jobs, C_i indicates the workers' order of preference across the jobs. It is important to note that we put no restrictions on the workers' productivity or intensity of preferences for workers with the same preference ordering.

We also assume the following:

¹³Note that Assumption 1 guarantees that these orderings do not depend on h .

ASSUMPTION 4:

$$\Pr(C_i = AB) + \Pr(C_i = BA) > 0.$$

The point of this assumption is to keep the model interesting. This is a model with on-the-job search and, if this assumption did not hold, there would be no AB or BA type workers and, thus, no worker would ever voluntarily switch jobs.

THEOREM 1: *Under Assumptions 1–4, with the data generated by the model expositied in Section 3, we can identify $\lambda_A^n, \lambda_B^n, P^*$, the distribution of C_i , and δ . If $\Pr(C_i = AB) > 0$, we can identify λ_A^e and, if $\Pr(C_i = BA) > 0$, we can identify λ_B^e .*

PROOF: Appendix D, available on our websites, contains the proof. *Q.E.D.*

The exceptions are not surprising. For turnover decisions, λ_A^e is only relevant for the AB types, so if $\Pr(C_i = AB) = 0$, then λ_A^e is not identified from these data. A similar argument holds for λ_B^e and the BA types.

The proof follows a random effect type of argument. We first show that we can identify λ_A^n and λ_B^n from the transitions out of non-employment. Next, following the patterns and timing of job-to-job switching, we can identify $P^*, \delta, \lambda_B^e, \lambda_A^e$, and $P(BA)/P(AB)$. Intuitively, these parameters are all identified from the rate at which people make job transitions. An important aspect is that the aggregate empirical hazard will be duration dependent due to unobserved heterogeneity. If we see a worker moving directly from a B firm to an A firm, it can be a voluntary switch for an AB worker or an involuntary switch for either a BA or an AB worker. Thus, as the duration increases, the fraction of AB workers falls. From the speed at which it falls, we can identify $\lambda_A^e, \delta P^*$, and $P(BA)/P(AB)$. From the transition to non-employment, we can separate δ from P^* . And from the transition from A to B , we can identify λ_B^e . Given these, we can infer the distribution of C_i from the patterns of workers at A and B firms over time.

4.2. Wage Components

We now incorporate information from wages. As a reminder of the notation, for any worker who is currently working, there are four different states which are relevant for their wages: their current employer, their current level of LBD human capital, the outside option when their current rental rate was negotiated, and the level of LBD human capital when the current rental rate was negotiated. We denote these as functions of the individual and time as $j(i, t), h(i, t), \ell(i, t)$, and $h_0(i, t)$, respectively. Then, for each time, t , at which the agent is working and wages are measured, we observe

$$\log(R_{ij(i,t)\ell(i,t)h_0(i,t)}) + \log(\psi_{h(i,t)}) + \xi_{it},$$

where ξ_{it} is i.i.d. measurement error. The distinction between $h(i, t)$ (the current level of LBD human capital) and $h_0(i, t)$ (the level of LBD human capital when wages were last negotiated) is visible here.

We now augment our Assumption 2 to include wage information. Since job-to-job transitions within a job type can be observed, we no longer assume part (b) of Assumption 2.¹⁴ Also, the nature of the Danish data we use is that we observe wages

¹⁴The goal of this was to tie our hands and to not use this for identification of job destruction and immediate new offer rates. This is no longer necessary, since we have showed that these are identified without this information.

once a year (at the end of November). We will mimic this by assuming that we only observe wages periodically and, for simplicity, we assume that it is at the integers, that is, at time 1.0, 2.0, etc.

ASSUMPTION 2': *The econometrician observes:*

(a) *The history of job-type spells, with start and stop dates, as well as the value of j at each job until point \bar{T} (including job switches within job type);*

(b) *If the individual is working, wages at the integers 1.0..., 2.0..., up until the largest integer less than \bar{T} ;*

(c) *We observe these for at least eight periods (i.e., $\bar{T} > 8$).*

We need an additional assumption, which is standard for the type of deconvolution argument we make.

ASSUMPTION 5: *The characteristic functions of the measurement error and of $\log(R_{iA00})$ (for workers who would work at an A type firm) do not vanish, and the logs of all random variables have finite first moments.*

The finite first moment could be avoided, but seems innocuous to us. The choice of R_{iA00} was also arbitrary; we could have chosen job B or another wage instead.

Finally, while, in principle, we could use the wage data to identify λ_A^e or λ_B^e when $\Pr(C_i = BA) = 0$ or $\Pr(C_i = AB) = 0$, we abstract from these special cases by assuming they are identified.

ASSUMPTION 6: *λ_A^e and λ_B^e are identified.*

THEOREM 2: *Under Assumptions 1, 2', and 3–6, with the data generated by the model presented in Section 3, we can identify the distribution of measurement error, ξ_{it} , human capital, ψ_1 , and the joint distributions of $(R_{iA00}, R_{iAB0}, \pi_{iA}, R_{iB00}, \pi_{iB}, R_{iA01}, R_{iAB1}, R_{iB01})$ conditional on $C_i = AB$ if $\Pr(AB) > 0$, $(R_{iA00}, \pi_{iA}, R_{iBA0}, R_{iB00}, \pi_{iB}, R_{iA01}, R_{iBA1}, R_{iB01})$, conditional on $C_i = BA$ if $\Pr(BA) > 0$, $(R_{iA00}, \pi_{iA}, R_{iA01})$, conditional on $C_i = A0$ if $\Pr(A0) > 0$, and $(R_{iB00}, \pi_{iB}, R_{iB01})$, conditional on $C_i = B0$ if $\Pr(B0) > 0$.*

PROOF: Appendix D, available on our websites, contains the proof.

Q.E.D.

The identification argument relies on the panel. Assumption 2' requires at least eight periods of data, which we need, because we identify up to an eight-dimensional object of conditional wages. The first complication is measurement error for which we use a deconvolution argument. The other complication is that we do not observe when human capital accumulates or when workers get outside offers. This means that the observed wages are determined by a mixture of distributions of unobserved state variables. However, we can calculate the distribution of the unobserved state variables conditional on the wage history. From this we know the mixture probabilities and can use an argument analogous to Bayes's theorem to infer the underlying distributions.

What cannot be identified is also important. For example, the $C_i = A0$ workers would never work at a type B job, so we cannot identify anything about their wages in those jobs. This is intuitively obvious, but it is important to keep in mind when we perform the counterfactuals, as this limits what can be credibly simulated.

4.3. *Non-Identification of β*

The final part of our identification result considers β . First, note that there are essentially two different possibilities. Either wages are never renegotiated for any worker or sometimes they are. For example, the results above show that we can observe the joint distribution of $(R_{iA00}, \pi_{iA}, R_{iA01})$ for $C_i = A0$. If wages are never renegotiated, then $R_{iA00} = \pi_{iA}$ and $R_{iA01} = \pi_{iA}\psi_1$, with probability 1. This will occur if either $\beta = 1$ (so the worker extracts the full surplus from the beginning and wages, then, do not respond to outside offers) or all $A0$ workers are indifferent between being employed and non-employed in which case there is no surplus to split. One cannot distinguish between these cases.

However, if there is some renegotiation in the model, we know that $\beta < 1$. In what follows, we show that this is generically all that we know about β . In particular, even in a restricted version of the model, for any other $0 \leq \beta^* < 1$, we can find unobserved preference components to rationalize the data (wages and job orderings).

THEOREM 3: *Under the assumptions of Theorem 2:¹⁵
If, with probability 1 for the relevant groups:*

$$R_{iA00} = R_{iAB0} = R_{iA01} = R_{iAB1} = \pi_{iA} \quad \text{and} \quad R_{iB00} = R_{iBA0} = R_{iB01} = R_{iBA1} = \pi_{iB},$$

then either $\beta = 1$ or all workers are indifferent in terms of all viable options.

If this is not the case, we know $\beta \in [0, 1)$ and that not all workers are indifferent in terms of all states. In this case, β is not generically identified. Moreover, in the special case of the separable model for $j \in \{A, B\}$:

$$u_{ij}(R) = \log(R) + v_{ij}^u$$

the model puts no restrictions on β . Specifically, for any $\tilde{\beta} \in [0, 1)$, we can generically¹⁶ find alternative preferences:

$$\tilde{u}_{ij}(R) = \log(R) + \tilde{v}_{ij}^u(\tilde{\beta}),$$

which is consistent with the distribution of the observed data in terms of wages and job choices.

PROOF: Appendix D, available on our websites, contains the proof. *Q.E.D.*

To provide some intuition for this problem, consider an even simpler case with a single job type A , no LBD human capital, no job destruction, the arrival rate of jobs λ is the same for employed and non-employed workers, no measurement error, and that everyone prefers employment to non-employment.

The flow utility from employment is $\log(R) + v_{iA}^u$, and the flow value of being non-employed is u_{i0} .

This model contains two wages for any given worker: the wage the worker receives right out of employment (which we call R_{iA0}) and the wage received when they get an outside

¹⁵This proof also works for the heterogeneous δ_i case.

¹⁶Linear equations of knowns and unknowns are required to show this, hence use of the word “generically”. Except for special cases in which the equations are linearly dependent, we can show non-identification.

offer, (R_{iAA}) . Since all firms are identical, the outside offer will be the competitive wage, $R_{iAA} = \pi_{iA}$. It is straight forward to show that

$$\begin{aligned} & \log(R_{iA0}) \\ &= \log(\pi_{iA}) - (\rho + \lambda)(1 - \beta) \left[\frac{\log(\pi_{iA}) + v_{iA}^u}{\rho} - \frac{u_{i0} + \lambda\beta \frac{\log(\pi_{iA}) + v_{iA}^u}{\rho}}{\rho + \beta\lambda} \right]. \end{aligned} \quad (4)$$

Using the same argument for identification as above, we can identify the joint distribution of $(\log(R_{iA0}), \log(\pi_{iA}))$. Equation (4) shows the lack of identification of β . For any value of β , we can find an alternative value of (v_{iA}^u, u_{i0}) that matches $\log(R_{iA0})$ (and, perhaps a bit less obviously, that does not alter the work decision). Thus, β is fundamentally unidentified. One cannot separate the bargaining parameter β from the intensity of preferences, (v_{iA}^u, u_{i0}) .

The theorem states that this general property holds for the more complicated model. The proof shows explicitly that, for any β , we can generically find alternative values of the idiosyncratic utility that precisely matches the turnover and wage data.

5. ECONOMETRIC SPECIFICATION/PARAMETERIZATION

Even though the model is mostly non-parametrically identified, estimating it non-parametrically is not feasible. In this section, we present our empirical specification, where we try to be flexible. We assume that log productivity of individual i at job type j is specified as

$$\log(\pi_{ij}) = \theta_i + \mu_j^p + v_{ij}^p, \quad (5)$$

where θ_i is the same for individual i at all jobs, μ_j^p is the same for all individuals at job j , and v_{ij}^p is the match-specific component. Thus, we allow for worker and firm heterogeneity in productivity and for match-specific (worker-firm) productivity. We will impose that the unconditional distribution of the three error terms are uncorrelated with each other. There will be substantial selection bias in our model so, while these components are unconditionally uncorrelated with each other, they will be correlated conditionally on the chosen jobs. This is a fairly restricted version of the underlying production function. For example, we do not allow for log complementarities. The reason is that we want to devise a simple model, and we think that this accomplishes that goal. In Appendix E and briefly discussed in Section 9, we discuss our exploration of a more general production function, which allows for sorting based on the production function. The model with the more general production function gives similar results.

The flow utility for individual i at job type j , with human capital rental rate R , and human capital level h , is

$$u_{ij}(R\psi_h) = \alpha \log(R\psi_h) + \mu_j^u + v_{ij}^u,$$

where α is the weight workers put on log consumption compared to the non-pecuniary aspects of a job. μ_j^u reflects common worker preferences across jobs, while we think of v_{ij}^u as heterogeneity across workers for the non-pecuniary aspects of a job. We expect v_{ij}^u to arise from the fact that different workers value different characteristics of the job.

In our model, the choice between any two jobs will be determined by the flow utility at each job evaluated at full surplus extraction with the wage $\pi_{ij}\psi_h$. We can rewrite the flow utility as the sum of three terms:

$$\begin{aligned} u_{ij}(\pi_{ij}\psi_h) &= \alpha(\theta_i + \mu_j^p + v_{ij}^p) + \alpha \log(\psi_h) + \mu_j^u + v_{ij}^u \\ &= \alpha(\theta_i + \log(\psi_h)) + (\alpha\mu_j^p + \mu_j^u) + (\alpha v_{ij}^p + v_{ij}^u). \end{aligned}$$

The first term is common across jobs, so all that matters for job-to-job turnover is the latter two terms. The second term pertains to firm-specific factors, and the last term pertains to match-specific factors. A special feature of log utility is that when human capital increases, the income and substitution effects balance, so that the preference ordering across jobs does not change. This greatly simplifies the computation. Given that human capital takes a relatively minor role, it likely makes little difference for the final results.¹⁷

We assume that θ_i is normally distributed with mean E_θ and variance σ_θ^2 and that we observe log wages with classical i.i.d. normally distributed measurement error ξ_{it} , which has mean zero and variance σ_ξ^2 .

The joint distributions (μ_j^p, μ_j^u) and (v_{ij}^p, v_{ij}^u) are ex ante independent of each other.¹⁸ The bargaining makes normalizations different from the standard case. Disregarding bargaining and non-employment, there would be two normalizations to make. The first is the standard scale normalization, which we impose by setting $\text{var}(v_{ij}^u) = 1$. The second issue is that α cannot be separately identified from the covariance components. That is, focusing on the idiosyncratic part, all that matters for the model is the joint distribution of v_{ij}^p and $(\alpha v_{ij}^p + v_{ij}^u)$. As a result, we cannot separately identify $\text{cov}(v_{ij}^u, v_{ij}^p)$ from α , so we normalize $\text{cov}(v_{ij}^u, v_{ij}^p) = 0$ and estimate α along with σ_{vp} , the standard deviation of v_{ij}^p .¹⁹ Given these two normalizations, we do not need additional ones for the (μ_j^p, μ_j^u) terms; we have already normalized the scale and the other normalization determines the size of α .

With bargaining, the argument is more subtle. We formally show that the bargaining parameter β is not non-parametrically identified. The issue is that one cannot separate the scale of the utility function from β . The two normalizations above set the scale of the utility function, and once that is done, β is identified. This implies that the level of β has no direct interpretation and cannot be compared to other measures in the literature; its level depends on the other “normalizations.” It is also important to point out that the normalizations are not precisely normalizations. Non-parametrically, they would be normalizations, but once we have restricted the distribution of the error terms to be jointly normal, they are not formally innocuous. Informally, it seems innocuous. To verify this, in Appendix E and discussed in Section 9, we use a very different “normalization.” We instead fix $\beta = 0$ and $\alpha = 1$ and estimate $\text{var}(v_{ij}^u)$ together with $\text{cov}(v_{ij}^u, v_{ij}^p)$. The estimated parameters differ, but the main conclusions from the wage decompositions are remarkably similar.

¹⁷We are able to solve the model without the assumption of log utility, but it complicates the numerical problem. Also, the income effect associated with human capital would be quite different if we allowed workers to borrow against future human capital growth. It also might be quite different if human capital was not perfectly general. For the question at hand, this does not seem particularly important, so we keep it simple, but allowing for a richer model of human capital is an important avenue for future research.

¹⁸As a result of selection, they will be correlated conditional on the jobs actually chosen.

¹⁹To see this, note that even after a scale normalization, what we can hope to identify is $\text{cov}(\alpha v_{ij}^p + v_{ij}^u, v_{ij}^p) = \alpha \text{var}(v_{ij}^p) + \text{cov}(v_{ij}^u, v_{ij}^p)$. We cannot separate the first component from the second. Thus, setting $\text{cov}(v_{ij}^u, v_{ij}^p) = 0$ is an innocuous normalization for us, though there are some counterfactuals for which it would not be innocuous.

We tried to choose a relatively parsimonious functional form for the distribution of (μ_j^p, μ_j^u) , which is a discrete distribution. With no obvious parametric alternative, we use the following:

$$\begin{aligned} \mu_j^u &= f_u[U_1(j) + f_{u,p}U_2(j)], \\ \mu_j^p &= f_p[f_{u,p}U_1(j) + U_2(j)], \end{aligned}$$

where $U_1(j)$ and $U_2(j)$ are distributed as discretely uniform across $[-1, 1]$. In our specification, we allow U_1 and U_2 to each take ten different values, and we assume these are unrelated to each other, giving us 100 different firm types. Essentially, f_u governs the variance of μ_j^u , f_p governs the variance of μ_j^p , and $f_{u,p}$ governs their correlation. Note that we deal with the normalizations with the idiosyncratic variables; once we have done that, we can leave these establishment-specific distributions flexible.

Human capital evolves as

$$\log(\psi_h) = b_1h + b_2h^2 + b_3h^3,$$

with the constraint that the profile is flat at the end:

$$\frac{\partial \log(\psi_H)}{\partial h} = 0.$$

In what follows, we treat b_1 and b_2 as free parameters and think of b_3 as then determined by the constraint.

The time period is assumed to be one year and we fix $\lambda_h = 1$.

As mentioned above, we allow the job destruction rate to vary across individuals; we specify it as

$$\log(\delta_i) \sim N(\mu_\delta, \sigma_\delta^2).$$

Allowing δ_i to vary across establishments in a way that is correlated with job types makes the model much more difficult to solve, and our preliminary investigation of this suggests that it is unlikely to change the main results.²⁰

In terms of specification of the arrival rates of jobs λ^e and λ^n , there are two important considerations. The first is the identification issue discussed in [Flinn and Heckman \(1982\)](#). With a continuous number of jobs, we cannot separately identify the arrival rate of jobs from the reservation utility. We address this issue by estimating the λ 's, but by restricting the location of the utility of non-employment. We take a simple specification for the value of non-employment by assuming

$$u_{i0} = \alpha[E(\theta_i) + \gamma_\theta(\theta_i - E(\theta_i)) + \nu_{i0}^n],$$

with $\nu_{i0}^n \sim N(0, \sigma_v^2)$. When $\theta_i = E(\theta_i)$ and $\nu_{i0}^n = 0$, then $u_{i0} = \alpha E(\theta_i)$. This means that when they have no LBD human capital, average workers' acceptance rate of jobs from non-employment would be roughly 50%. We think of this as a normalization; choosing a different value would lead to different estimates of λ . This should be taken into account when comparing the estimates' arrival rates to others in the literature.

²⁰See footnote 6.

For the same reason that we cannot separately identify the level of λ^e and λ^n from the level of u_{i0} , we cannot separate individual heterogeneity in λ^e and λ^n from heterogeneity in u_{i0} . As shown, we incorporate this as heterogeneity in u_{i0} , and do not allow for individual-specific heterogeneity in the λ 's which we think of as a normalization.

The second issue is allowing for job-specific arrival rates, that is, j subscripts on λ_j^e and λ_j^n . We parameterize the model somewhat differently. We estimate λ^e and λ^n , which are the arrival rate of *any* job. We then estimate the distribution from which those jobs are drawn (and assume this is the same regardless of whether it comes from employment or non-employment). Thus, instead of allowing for differences in arrival rates across job types and assuming a mass of one of each type, we allow the mass of job types to differ but assume the same arrival rate. In absence of a theory about firm size, we view the two ways to be isomorphic.

Finally, we fix the discount rate at $\rho = 0.05$. This leaves a total of 18 parameters to be estimated:

$$[\mu_\delta, \lambda^n, \lambda^e, E_\theta, \sigma_\theta, \sigma_\xi, \sigma_{vp}, \alpha, f_u, f_p, f_{u,p}, b_1, b_2, \beta, P^*, \sigma_\delta^2, \sigma_v^2, \gamma_\theta].$$

6. DATA AND DANISH INSTITUTIONAL FEATURES

We use Danish, matched employer-employee data that consist of two types. The first type is weekly spell data, which cover all individuals aged 15–74 in Denmark from 1985 to 2003.²¹ The version of the spell data used in this paper is generated from various raw registries maintained by Statistics Denmark. It consists of a worker identifier, firm and establishment identifiers, start and end date of the spell, and a state variable. The states are employed, unemployed, self-employed, retired, and non-participation. Unemployment is defined by receiving either unemployment insurance benefits or social assistance. Non-participation is a residual state in the sense that it means that we do not observe the worker in any of the available registries. The second type of data is annual cross-section data from the Danish register-based, matched employer-employee data set IDA (Integrated Database for Labor Market Research) and other annual data sets.²² In addition to containing socioeconomic information on workers and background information on employers, IDA covers the entire Danish population age 15 to 74.²³ Thus, the unit of observation in our data set is a worker-week-labor market state-establishment (if employed).

We choose the empirical analogue of a job type to be an establishment and not a firm. We differ on this point from most of the empirical search literature, which uses firms as the employer unit. Using establishments has at least three advantages in the current setting. The firm identifier in the Danish data is not well defined over time, since it is based on a mixture of tax reporting numbers and legal units. Firms might change both the tax reporting number and the legal unit (and hence the firm identifier) without changing anything else. The establishment identifier, however, is consistent over time.²⁴ Second, when

²¹See [Bobbio and Bunzel \(2018\)](#) for more detailed descriptions.

²²Both IDA and the other annual data sets are constructed and maintained by Statistics Denmark.

²³IDA contains the annual average hourly wage for the job occupied in the last week of November. However, as shown in [Lund and Vejlin \(2016\)](#), the hourly wage measure developed by Statistics Denmark has several drawbacks. As a result, we use the hourly wage measure suggested by [Lund and Vejlin \(2016\)](#), which improves the original measure in multiple ways.

²⁴The establishment is constructed by Statistics Denmark and is the same across years if one of three criteria is met: same owner and industry; same owner and workforce; same workforce and either same address or same industry.

considering preferences for non-pecuniary aspects, which is an integral part of the model, the most appropriate unit seems to be the establishment and not the legal firm, since many amenities are determined locally. Third, using establishments provides a more convenient way to think about government jobs, which are included in our sample. Treating the government as one firm is problematic. Since different government establishments often have very different responsibilities and do not coordinate with each other, thinking about them as separate units appears to be the best approximation.

We aggregate unemployment and non-participation into non-employment. How to think about non-participation is an open question. Looking at the data, it is clear that non-participation in the data is not an absorbing state in the sense that the hazard rate into employment from non-participation is around half of the hazard from unemployment. Our model allows for workers who do not take any job, so including workers who are not active in the labor market is not a big issue, since it will be captured by a high flow utility from non-employment.

We categorize a job-to-job transition as a transition between two establishments potentially separated by an up to two week intermediate non-employment spell (see details in Appendix A in the Supplemental Material (Taber and Vejlín (2020))). From the identification strategy, it is clear that we need to distinguish between voluntary job-to-job transitions and involuntary job-to-job transitions in order to interpret them as a revealed preference. It is obvious that not all job-to-job transitions in the data are indeed voluntary. In Denmark, an average notice period is 2–3 months.²⁵ This means that, on average, a displaced worker has around 2–3 months to find a new job before the old job stops. If the worker finds a job in this period and starts the new job before the old one stops, it will look like a job-to-job transition in the data. Using only population-wide register data, we are unable to credibly distinguish such a case from a true voluntary job-to-job transition. As shown above, we can identify P^* (immediate offer after job destruction) without direct information on the status of the job-to-job transition. However, we chose to augment the standard data from Statistics Denmark with survey data from two representative samples of workers in 1995 and 2000.²⁶ Both surveys contain questions on whether the job spell in the last week of November five years ago terminated on the initiative of the employer or the employee. The survey data are matched with the register data and used later to form an auxiliary parameter that identifies the structural parameter, P^* . We believe that using survey data is an improvement over the common practice of inferring the involuntary job-to-job transition rate based on model specification. However, using survey data is not without problems. First, there are the usual problems using survey data and backward looking questions, such as recall bias. Second, the question answered in the survey is not accurate enough. Ideally, we would want to know if the worker who quit did so due to their firm's financial problems or if they would have quit regardless of the economic state of the firm. Since our model precludes firm dynamics, we cannot simulate such behavior. Still, we believe that using survey data is an improvement over using the more standard approaches.

6.1. Sample Selection Criteria

To select an appropriate estimation sample, we use the following sample selection criteria involving these main steps:²⁷ first, we censor workers after age 55 to avoid retirement

²⁵The notice period typically depends positively on tenure and also varies across sectors.

²⁶The survey is the "The Danish Work Environment Cohort Study."

²⁷See Appendix A for a more thorough description.

TABLE I
SUMMARY STATISTICS: POOLED CROSS SECTIONS

	Mean	Std. Dev.
Number of years in sample	11.06	6.27
Number of establishments per worker	2.70	1.85
Female	0.49	
Preparatory educations	0.30	
High school	0.04	
Vocational education	0.42	
Short further education	0.04	
Medium-length further education/Bachelor's degree	0.14	
Master's degree and Ph.D.	0.05	
Average years of education	11.69	3.18
Age	38.31	9.63
Employed	0.83	
Publicly employed	0.32	
Missing establishment ID	0.01	
Real experience in years	13.46	9.47
Log hourly wages	4.50	0.35

related issues. We also disregard spells before labor market entry (or age 19), defined as the time of highest completed education (and not observed in education later). Since the model does not contain retirement or self-employment, we censor workers when they enter those two states. Thus, the sample selection criteria are not strict, and we obtain an almost balanced sample in the sense that it consists of around 1.8 million workers in 1985 and 1.65 million in 2003.

Since the model is cast in steady state, which means cross-sectional wages have no trend, we detrend wages in logs by sex-educational groups conditional on experience. We condition on experience since the composition of workers changes over the sample period due to an aging workforce.

Since job-to-job transitions play a vital role in the identification of our model, we ignore transitions from two types of establishments. The first are transitions for workers to or from establishments with missing ID (0.5%; cf. Table I). We also ignore job-to-job transitions from closing establishments or establishments with mass layoffs.²⁸

6.2. Descriptive Statistics

In this section, we present different descriptive statistics for the estimation sample. The number of years in sample and the number of establishments for each worker are important for the identification of the model. Table I shows statistics for these measures, together with other descriptive statistics. The worker is, on average, in the sample for 11 years and is employed in 2.7 different establishments. There are almost as many women as men in the sample. This is because we are not censoring or deleting public employees, many of whom are women. The workers have 12 years of education on average. However, this moderately changes over the sample period, since entering workers are better

²⁸A mass layoff is defined as the establishment having more than 15 workers in year t and less than 70% left in year $t + 1$. We define an establishment as closed if we no longer observe it in the data. The auxiliary parameters are not sensitive to the definition of a mass layoff.

educated than those leaving the sample. The average age is 38. A total of 83% are employed in general, while 32% are employed in the public sector. The fact that only 83% are employed is intentional and a result of the mild sample selection that we impose. The average labor market experience is 13 years. Finally, in a given cross section, the establishment identifier is missing for 0.5% of all employment observations.

6.3. *Institutional Setting*

The Danish labor market is characterized by having a so-called flexicurity system comprising three main pillars: a generous social security system, flexible hiring and firing rules, and an active labor market policy with a focus on job search and re-employment possibilities; see [Andersen and Svarer \(2007\)](#) for a more thorough review. The generous social security system consists of two main parts: unemployment insurance benefits and social assistance, which is means tested. The former is heavily subsidized by the government, with members of unemployment insurance funds paying around one-third of the total cost in membership fees. If an individual is ineligible for unemployment insurance, they can apply for social assistance. This is means tested, and eligibility requires, for example, that the worker does not own much of value and is available to work. The flexible hiring and firing rules and the low degree of employment protection result in a highly fluid labor market; see, for example, [Bertheau, Bunzel, Hejlesen, and Vejlin \(2016\)](#). These first two pillars were already in place in the 1970s. However, in the 1970s and the 1980s, Denmark performed relatively poorly in international comparisons. A series of reforms in the early 1990s shifted focus from passive to active labor market policies, shortened the maximum unemployment duration, and expanded the eligibility criteria. This combined approach has become known as the flexicurity model.

A full-time job in Denmark is around 1700 hours per year, which is lower than in the United States, but the participation rate is rather high at around 75–80% for both men and women. Furthermore, as indicated by the sample above, the public sector is large, which is typical for a Scandinavian welfare state.

Clearly, the model presented makes some strong assumptions, most notably regarding the wage-setting mechanism. The fact that wages can be renegotiated when credible outside offers present themselves might sound dubious in a labor market that is as highly unionized as Denmark's. Note, however, that [Caldwell and Harmon \(2019\)](#), using coworker networks in Denmark, found that when a worker's outside option improves (higher demand within the coworker network), it generates within-job wage increases, which is what sequential bargaining would predict. On a more structural level, Denmark underwent a transformation in the 1980s and the early 1990s, from having a centralized wage bargaining system to having an increasingly decentralized wage-setting system; see, for example, [Dahl, le Maire, and Munch \(2013\)](#) and [Boeri, Brugiavini, and Calmfors \(2001\)](#). With this change in mind, centralized bargaining is still an important part of the Danish labor market, especially for the public sector. It is clear that for most public sector jobs, wage renegotiation does not take place when an outside offer appears. However, there is wide use of tenure contracts in both the public and private sectors, in the sense that wages go up with tenure at a specific pace. This is especially true for contracts negotiated in the more centralized system. There are several papers, starting with [Burdett and Coles \(2003\)](#), that study wage-tenure contracts theoretically. In these models, tenure effects arise as a result of firms trying to reduce their quit rate by backloading wages. This is essentially the same underlying mechanism as in our framework, where tenure effects also arise as a result of firms trying to retain workers. With obvious caveats, tenure contracts could be interpreted as firms realizing that workers, over time, accumulate outside

offers but without being able or willing to renegotiate wages. Thus, the same underlying economic mechanism is at play.

7. AUXILIARY MODEL

We estimate our model using indirect inference based in [Gourieroux, Monfort, and Renault \(1993\)](#). We simulate the model 1,580,000 times to produce the estimates drawing from the empirical distribution of start and stop times of the data window. Our approach is to use the formal, non-parametric identification results as a guide to which aspects of the data identify the different parameters. To keep the relationship between the structural parameters and the data as transparent as possible, we focus on the just-identified case. In particular, for each structural parameter, we choose one auxiliary parameter that we think is useful for identifying it.²⁹ We find this approach to be highly beneficial in understanding the mapping between the structural parameters and the data. This is also relevant because identification of some of our parameters is more subtle than others. The most extreme example is β versus E_θ . We argue that the coefficient on tenure squared in the fixed effect regression plays a key role in identifying β while mean wages play the primary role in identifying E_θ . The former coefficient is very imprecisely estimated compared to the latter parameter. We suspected that if we included many more auxiliary parameters in the data, the coefficient on tenure squared would receive very little weight, and identification of β would come more from higher-order moments of the wage distribution rather than the patterns of wage growth on the job.

The details for construction of the auxiliary parameters and their standard errors can be found in Appendix C of the Supplemental Material. Here, we briefly describe the auxiliary model we use.

Transition Data

We use duration data on the average length of a non-employment spell, the average length of an employment spell, and the average length of a job spell. The difference between the last two is that a job spell is for a specific establishment, while an employment spell is the length of time at all employers between two non-employment spells. To identify heterogeneity in these objects, we include the variance of employment spells, the variance of non-employment spells, and the covariance between average wages and non-employment duration.

Basic Wage Information

We use the mean wage and also use a three-way variance decomposition that decomposes total variance into within-establishment, between-establishment/within-person, and between-person variances.

Firm Information

To identify the relative importance of establishment types, we construct \tilde{w}_{it} as an average wage residual at the given firm relative to the same individual at other firms. We

²⁹While we use this language, this is not precisely how the estimation works. In practice, all the auxiliary parameters are useful for identifying all of the structural parameters.

construct \tilde{S}_{itj} and \tilde{r}_{itj} , which are measures of revealed preference (through job-to-job transitions) for the establishment for the worker and the worker's coworkers, respectively. There are three parameters involving firms, essentially picking up the variances and covariance of (μ_j^p, μ_j^u) . We use three auxiliary parameters to pick up the three structural parameters: First, the covariance between the workers' wage residuals and coworkers' wage residuals; second, the covariance between workers' revealed preferences and coworkers' revealed preferences; and, finally, the covariance between workers' wage residuals and their coworkers' revealed preferences.

Wage Dynamics

We run a log wage regression on experience, experience squared, and tenure squared with individual \times establishment fixed effects. Note that this is the first stage of the procedure in [Topel \(1991\)](#) where the coefficient on linear experience will pick up both the linear experience and linear tenure effects.³⁰ We expect the magnitude of the coefficient on tenure squared to be important for identifying the importance of outside offers influencing wage growth on the job beyond human capital effects. We also construct a measure of the probability that we see wages fall following a job-to-job transition.

Involuntary Job-to-Job

This variable comes from survey data rather than the administrative data and just picks up the fraction of time respondents report that job-to-job transitions were involuntary. This likely understates the number of true involuntary separations as some people might voluntarily find a new job because of concerns about future layoffs.

To give an overview, [Table II](#) provides what we call an Identification Map. The table lists each of the auxiliary parameters described above and also which structural parameter each one primarily helps identify. While all structural parameters are determined by all auxiliary parameters, and they interact in interesting ways, we present the table as an approximate illustration of how we think about identification. Taking everything else as given, the structural parameter in the model is driven by changes in the auxiliary parameter next to it. We also list the counterfactuals as determination of all of them is explicitly achieved by manipulating the structural parameters. In [Section 9](#), we discuss our simulations that show that the Identification Map is a reasonable approximation of how identification works in practice in our model.

8. RESULTS

We estimate the model using indirect inference with the auxiliary model described above. [Appendix B](#) of the Supplemental Material contains details on how we compute the model. Our objective function is the sum of the squared deviation between the simulated model and the data weighted by the inverse of the absolute value of the estimated parameters.³¹ The covariance matrix of the auxiliary parameters has been calculated by bootstrapping the sample 500 times. [Appendix C.3](#) discusses the details. This section presents

³⁰That is, with individual \times establishment fixed effects, tenure will be perfectly colinear with experience.

³¹We do this instead of the most common approach, which is to normalize weight by the inverse of the standard errors (which is related to the efficient weighting matrix). Our goal is to match all the auxiliary parameters, but the standard errors are very different for the different auxiliary parameters, and often very small. We worry that the standard approach might put essentially no weight on some auxiliary parameters. Given the fit of the model, the weighting matrix matters little in practice.

TABLE II
IDENTIFICATION MAP

Auxiliary Parameter	Structural Parameter	Counterfactual
Coefficient on Experience	Coef on linear term (human capital): b_1	Learning by doing
Coefficient on Experience ²	Coef on quadratic term (human capital): b_2	Learning by doing
Coefficient on Tenure ²	Bargaining power: β	Monopsony
Variance between Person	Std. dev. of worker productivity: σ_θ	Premarket
Variance between Jobs	Std. dev. of match productivity: σ_{v^p}	Skills
Fraction Wage Drops	Weight on log wage: α	Non-pecuniary aspects
$E(\tilde{S}_{itj}\tilde{r}_{-itj})$	Firm utility parameter: f_u	
$E(\tilde{w}_{itj}\tilde{r}_{-itj})$	Firm utility \times productivity parameter: $f_{u,p}$	
Avg Length Job Spell	Employment job offer arrival rate: λ^e	Search frictions
Avg Length Non-emp Spell	Non-employment job offer arrival rate: λ^n	
Variance within Job	Std. dev. of measurement error: σ_ξ	Measurement error
$E(\tilde{w}_{itj}\tilde{w}_{-itj})$	Firm productivity parameter: f_p	
Avg Length Emp Spell	Mean of log job destruction distribution: μ_δ	
Var Employment Duration	Std. dev. of log job destruction distribution: σ_δ	
Sample Mean w_{it}	Mean worker productivity: E_θ	
Involuntary Job-to-Job	Prob of imm offer upon job destruction: P^*	
Var Nonemp Duration	Std. dev. of idio. non-employment utility: σ_v	
Cov(\bar{w} , non-emp dur)	Worker ability cont to flow utility: γ_θ	

the results from the main specification, while Section 9 discusses results from alternative specifications.

8.1. *Fit and Estimates*

Table III presents the structural parameters of the model. The magnitude of the structural parameters is easier to judge in the context of their contribution to the counterfactuals, but we want to comment on a few of them here. First, we focus on the job offer arrival rates, λ^e and λ^n . We find a much higher value in the former compared to the latter. This is driven by the fact that we are looking at non-employment rather than unemployment. The auxiliary parameters show that non-employment spells tend to be similar in length to job spells, but switching jobs from employment should happen at a lower rate, since workers with jobs are presumably much pickier. Whether this is due to heterogeneity in arrival rates or reservation values is hard to identify, and we have made a certain normalization through heterogeneity in reservation utility rather than heterogeneity in arrival rates. The primary goal of this project is to explain wages rather than employment, so we do not view this issue as first order for this paper but worth exploring for other papers that are more concerned with explaining unemployment. Also note that the arrival rates cannot be directly compared to other estimates in the literature since, in our model, workers only accept around 50% of all jobs out of unemployment.³²

³²Bagger et al. (2014) estimated a similar model on the same data and found that the monthly job offer arrival rate when employed is roughly 6%. We find that the monthly probability is $\frac{1-e^{-2.08}}{12} = 0.16$. However, since workers only accept 50% of all jobs, then the comparable rate is 0.08. The difference is likely driven by two differences. First, we use establishments instead of firms, and second, Bagger et al. (2014) truncated labor market histories when workers enter the public sector, which tends to select non-mobile workers.

TABLE III
PARAMETER ESTIMATES

Parameter	Description	Estimate	Standard Error
E_θ	Mean worker productivity	4.26	(0.001)
σ_θ	Std. dev. of worker productivity	0.217	(0.001)
σ_{vp}	Std. dev. of match productivity	0.211	(0.001)
α	Weight on log wage	3.575	(0.043)
β	Bargaining power	0.844	(0.008)
P^*	Probability of immediate offer upon job destruction	0.394	(0.019)
λ^n	Non-employment job offer arrival rate	0.989	(0.002)
λ^e	Employment job offer arrival rate	2.079	(0.011)
μ_δ	Mean of log job destruction distribution	-2.96	(0.026)
σ_δ	Std. dev. of log job destruction distribution	2.262	(0.009)
$b_1 \times 100$	Coefficient on linear term (human capital)	0.262	(0.103)
$b_2 \times 100$	Coefficient on quadratic term (human capital)	0.087	(0.006)
σ_ξ	Std. dev. of measurement error	0.139	(0.001)
f_u	Firm utility parameter	2.163	(0.169)
f_p	Firm productivity parameter	0.142	(0.003)
$f_{u,p} \times 100$	Firm utility \times productivity parameter	0.467	(0.532)
σ_v	Std. dev. of idiosyncratic non-employment utility	0.351	(0.012)
γ_θ	Worker ability contribution to flow utility	-0.282	(0.030)

Second, the standard deviation of the match productivity term, σ_{vp} , is estimated to be 0.211, which implies that match effects are important in our model. Recently, [Card, Heining, and Kline \(2013\)](#) (CHK) have argued that there are almost no match effects using German data. They based this on two findings. First, they found almost no match effects running an AKM model.³³ Second, they showed that the average gain (loss) from moving up (down) the firm ladder is symmetric, which they interpreted as evidence suggesting that there cannot be large match effects. At first sight, our result and CHK's results appear to conflict. However, we argue that they do not. First, σ_{vp} is a parameter from the offer distribution, and since there is a lot of selection into jobs in our model, it is not directly comparable to the realized distribution of wages. Second, in an AKM model, much of the variance in match effects is subsumed into the worker effects.³⁴ This is especially true in shorter panels, as used in CHK. As will become clear in the next section, setting $\sigma_{vp} = 0$ decreases the variance of wages significantly, suggesting that match effects are important. The reason that this is not necessarily at odds with the symmetry results found in CHK is that the workers in our model move based on utility and not just productivity. In Section 9, which looks at robustness, we explicitly show that our model is able to

³³Sørensen and Vejlin (2013) found that match effects explain around 7% of the variance of log wages using Danish data, which is comparable to the CHK results on German data.

³⁴To illustrate this, we estimated AKM on our 19-year-long sample. Afterwards, we replaced real wages with simulated wages. The simulated wages comprise two terms: first, a normally distributed error term with a variance of 0.019 as estimated in our AKM model; second, a normally distributed match effect with a variance of 0.104, such that the total variance is the variance we observe in the data. We then ran an AKM model on this partly simulated data and used the worker-firm allocations, as seen in the data, but the simulated wages instead of the real wages. The variances of the estimated worker, firm, and match effects are 0.068, 0.010, and 0.043, respectively. This illustrates that a high proportion of the variance of the true match effects can be soaked up by worker effects and, to a smaller extent, firm effects. This highlights that the moments from our model are not directly comparable to the moments estimated by AKM.

reproduce the symmetry found in CHK between moving up/down the firm ladder while maintaining a high value of σ_{vp} .

Finally, we comment on the value of β . This is estimated to be 0.844, which is considerably higher than other studies find. However, the estimate is not comparable for several reasons. First, the value depends on the normalization ($\text{var}(v_{ij}) = 1$). If we picked a different value for $\text{var}(v_{ij})$, we would get a different level of β . We explore this in Appendix E, which can be found on our webpages, and show that our model gives very similar result with a different “normalization” in which we set $\beta = 0$. Second, more intuitively, incorporating preferences for non-pecuniary aspects of jobs into the model fundamentally changes the bargaining game, since utility in our model is also derived from non-pecuniary differences and not just pecuniary differences across jobs. As can be seen from the parameter estimates (and the utility variance decomposition presented later), preferences for non-pecuniary aspects are important for utility. This means that, for explaining the same amount of wage increase within a job, one would expect to find a considerably higher β , since a low β would generate huge within-job wage increases, because outside offers vary greatly in utility terms and, to compensate, the incumbent firm would need to increase the wage significantly. Third, we are identifying β in a different way than other papers. When $\beta = 1$, the bargaining process is unimportant in determining wages. For a given $\text{var}(v_{ij})$, the further β is from 1, the more important the bargaining process is. If the bargaining process matters a great deal, this means that tenure should be important in determining wages; the longer a worker has been at the firm, the more outside offers they have received and the larger their wages will be. We think of β as identified primarily through the coefficient on tenure squared (since the coefficient on the level is not identified). We do not know of other papers that have done this. Note that since the model disregards job-specific human capital, one could argue that we overstate the importance of tenure, which would lead to an underestimate of β . For all three reasons, we do not think our value of β is comparable to other papers in the literature.³⁵

Table IV shows the auxiliary parameters from the sample and model. The fit is excellent. This is perhaps not surprising, because we have as many free parameters as we do auxiliary parameters to match. However, the model is nonlinear, so there is no guarantee of a match.³⁶

Note that wage drops at job-to-job transitions are not at all uncommon, roughly 40%. Some of this will be explained by the fact that many of these transitions, 20%, are involuntary. Some will be explained by measurement error, while others will be explained by the mechanism described in Postel-Vinay and Robin (2002), in which workers will take wage cuts when moving to more productive firms in anticipation of future wage increases. The rest will be explained by compensating differentials.

8.2. Statistical Decomposition of Wage Variation

Our main goal is to decompose wage variation into its economic factors. Before doing so, we begin with a simpler statistical decomposition that helps shed light on those results.

Many papers using search models make linear decompositions of the variance of wages (see, e.g., Postel-Vinay and Robin (2002)). This is not, however, possible in our model,

³⁵As previously mentioned, we present a very different normalization fixing $\beta = 0$ and $\alpha = 1$, and then estimating $\text{var}(v_{ij}^u)$ together with $\text{cov}(v_{ij}^u, v_{ij}^p)$ in Section 9/Appendix E. The fit and estimated parameters differ. However, the main conclusions from the wage decompositions are the same.

³⁶It is also possible that there are multiple solutions of the model that all fit the data. While one can never guarantee a global optimum, we have found no evidence of this.

TABLE IV
AUXILIARY MODEL AND ESTIMATES

Aux Parameter	Data (Std. Err.)	Model
Avg Length Emp. Spell	377 (0.193)	377
Avg Length Non-emp. Spell	91.4 (0.094)	91.2
Avg Length Job Spell	106 (0.102)	106
Sample Mean w_{it}	4.50 (0.000)	4.50
Between Persons $\times 100$	8.03 (0.012)	8.00
Between Jobs $\times 100$	2.87 (0.006)	2.88
Within Job $\times 100$	1.49 (0.003)	1.49
$E(\tilde{w}_{itj}\tilde{w}_{-itj}) \times 100$	0.77 (0.004)	0.77
$E(\tilde{r}_{-itj}\tilde{w}_{itj}) \times 100$	0.69 (0.005)	0.69
$\text{cov}(\tilde{r}_{-itj}, \tilde{S}_{itj}) \times 100$	8.18 (0.013)	8.21
Fraction Wage Drops	0.400 (0.000)	0.392
Coeff Exper $\times 100$	2.48 (0.006)	2.47
Coeff Exper ² $\times 1000$	-0.291 (0.001)	-0.292
Coeff Tenure ² $\times 1000$	-0.460 (0.003)	-0.460
Var(Non-employment)	16,000 (50.39)	16,150
Cov(\bar{w}_i , Non-employment)	-3.42 (0.030)	-3.43
Var(Employment Dur)	102,000 (72.82)	102,666
Invol Job to Job	0.205 (0.011)	0.205

since the mechanisms interact in different ways. It is nevertheless possible to do a different decomposition which is similar in spirit. Recall that, in our model, the wage is equal to the rental rate on human capital times the level of human capital, $w(R, \psi) = R\psi$. R is a complicated nonlinear function of the other components of the model. Based on equation (5), we can rewrite our wage equation as

$$\begin{aligned} \log(w_{it}) &= \log(R_{ij(it)}) + \log(\psi_{h(it)}) + \xi_{it} \\ &= [\log(R_{ij(it)}) - \log(\pi_{ij(it)})] + \log(\pi_{ij(it)}) + \log(\psi_{h(it)}) + \xi_{it} \\ &= [\log(R_{ij(it)}) - \log(\pi_{ij(it)})] + \theta_i + \mu_{j(it)}^p + v_{ij(it)}^p + \log(\psi_{h(it)}) + \xi_{it}. \end{aligned}$$

With this in mind, we do the following linear decomposition of the log of wage variance:

$$\begin{aligned} \text{Var}(\log(w_{it})) &= \text{Cov}(\log(w_{it}), \log(R_{ij(it)}) - \log(\pi_{ij(it)})) \\ &\quad + \text{Cov}(\log(w_{it}), \theta_i) + \text{Cov}(\log(w_{it}), \mu_{j(it)}^p) \\ &\quad + \text{Cov}(\log(w_{it}), v_{ij(it)}^p) + \text{Cov}(\log(w_{it}), \log(\psi_{h(it)})) \\ &\quad + \text{Cov}(\log(w_{it}), \xi_{it}), \end{aligned}$$

where $h(it)$ and $j(it)$ are the human capital level and job type of individual i at time t , and ξ_{it} is measurement error. The first term in bracket picks up the importance of the bargaining, that is, if $\beta = 1$, then $R_{ijt} = \pi_{ijt}$.

Table V shows the result from the decomposition and previews many of the main results from our nonlinear decomposition. The two largest parts, by far, are the covariance with θ_i and the covariance with $v_{ij(it)}^p$. The other components are non-trivial but clearly smaller. Note that this does not inform us about the relative importance of the economic model components. In particular, $\text{Cov}(\log(w_{it}), v_{ij(it)}^p)$ depends strongly on selection determined

TABLE V
LINEAR WAGE VARIANCE DECOMPOSITION

Var(log(w_{it}))	0.124
Cov(log(w_{it}), log($R_{ij(it)} - \log(\pi_{ij(it)})$))	0.005
Cov(log(w_{it}), θ_i)	0.044
Cov(log(w_{ijt}), $\mu_{j(it)}^p$)	0.008
Cov(log(w_{it}), $v_{ij(it)}^p$)	0.042
Cov(log(w_{it}), log($\psi_{h(it)}$))	0.006
Cov(log(w_{it}), ξ_{it})	0.019

by initial skills, search frictions, and preferences for non-pecuniary aspects of jobs. In the next subsection, we attempt to uncover the relative contribution of these different factors.

8.3. Model Decomposition of Wage Variation

Table VI presents the decomposition of total log wage variance. We sequentially eliminate the different sources of wage inequality and document their effect on inequality. Prior to the decomposition in the table, we eliminate measurement error by setting $\sigma_\xi^2 = 0$. The total variance of log wages in the model and in the raw data is 0.124, but it falls to 0.104 without the measurement error.

There are many ways to decompose wages, depending on the order in which we eliminate the sources of wage inequality. We use different orders to highlight different features of the model.³⁷ Note that even though we have endogenous wage determination,

TABLE VI
COUNTERFACTUAL DECOMPOSITION OF VARIANCE OF LOG WAGES

(A)		(B)	
Total	0.104	Total	0.104
No Learning by Doing	0.096	No Learning by Doing	0.096
No Monopsony	0.093	No Monopsony	0.093
No Premarket Skill Variation Across Jobs	0.050	No Search Frictions	0.086
No Premarket Skill Variation at All	0.008	No Premarket Skill Variation Across Jobs	0.049
No Search Frictions	0.007	No Premarket Skill Variation at All	0.007
(C)		(D)	
Total	0.104	Total	0.104
No Learning by Doing	0.096	No Learning by Doing	0.096
No Monopsony	0.093	No Monopsony	0.093
No Non-Pecuniary Aspects of Jobs	0.087	No Non-Pecuniary Aspects of Jobs	0.087
No Premarket Skill Variation Across Jobs	0.048	No Search Frictions	0.061
No Premarket Skill Variation at All	0.006	No Premarket Skill Variation Across Jobs	0.047

³⁷We consider what we are doing a nonlinear extension of an Oaxaca/Blinder decomposition. As an example, suppose that there are two sets of structural parameters ζ and v (e.g., one set representing search frictions and the other non-pecuniary aspects of jobs). The model produced variance of log wages $V(\zeta, v)$, and we know that $V(\zeta_0, v_0) = 0$ (e.g., when the parameters take these values, search frictions and compensating differentials disappear). Let $\hat{\zeta}$ and \hat{v} be the estimated parameters so the model predicts variance $V(\hat{\zeta}, \hat{v}) - V(\zeta_0, v_0)$.

the model does not endogenize job offer arrival rates or the distribution of job types. This is important to keep in mind—they are partial equilibrium simulations. Allowing for equilibrium effects is an interesting and important extension of our work.

We begin with the variance of 0.104 in an attempt to determine which factors contribute to it. We simulate four different sequences of decompositions (A)–(D) in Table VI. Note that we eliminate sources sequentially. For example, in column (A), the first row presents results where we eliminate human capital. The second row presents results where we eliminate human capital and monopsony. In the third, we eliminate, for example, human capital, monopsony, and variation in premarket skills across jobs, etc.

In all four panels, the first two steps are identical. We begin by eliminating variation in human capital by setting it to the maximum level, ψ_H . We do this first since it is well known to have little explanatory power (i.e., the R^2 in a Mincer model does not change much when experience and experience squared are dropped), so we view this as less interesting for the purpose of explaining variation in wages. We show that the small explanatory power is true and that human capital explains about 8% of wage variation. When eliminated at later stages of the process, it is even smaller.³⁸

The second step is to eliminate the variation in monopsony powers that firms have over workers reflected in the bargaining process by setting $\beta = 1$. We set $\beta = 1$ prior to the other counterfactuals for reasons discussed in Sections 4.3 and 5. The level of β is determined by normalizations on other parameters (mostly the scale of preferences), so it does not make sense to hold it fixed while changing the other parameters. Setting $\beta = 1$ eliminates the variation in the model that comes from negotiation by giving all of the bargaining power to the worker. This lowers the variance of wages by about another 3%. The fundamental source in the model that leads to this heterogeneity is search frictions, that is, in a perfectly competitive environment, firms would have no monopsony power. Thus, in an accounting sense, the 3% should be attributed to search frictions. Next, we eliminate the remaining parts: search frictions, non-pecuniary aspects of jobs, and differences in premarket skills. These are done in different orders as the order does matter considerably. As discussed in Section 4, the selection problem makes eliminating non-pecuniary aspects of jobs the most tenuous of these. Thus, the most reliable simulations are (A) and (B), where we eliminate differences in premarket skills and search frictions first. More specifically we eliminate:

- Search frictions by allowing workers to find the most preferred job immediately (i.e., $\lambda^e, \lambda^n \rightarrow \infty$). Note that this also implicitly eliminates variation in job destruction rates, since when a job is destroyed a new equivalent one will be found instantly. We find that heterogeneity in δ_i plays a minor role in inequality, so we do not show it explicitly.³⁹

This can be rewritten as $V(\widehat{\zeta}, \widehat{v}) - V(\zeta_0, v_0) = (V(\widehat{\zeta}, \widehat{v}) - V(\zeta_0, \widehat{v})) + (V(\zeta_0, \widehat{v}) - V(\zeta_0, v_0))$. Similar to an Oaxaca/Blinder decomposition, we could have changed the baseline by simply switching ζ and v .

³⁸Human capital interacts with the bargaining process, since it affects the utility for non-employment, and interacts with search frictions, since it only accumulates when people work. Specifically, if we interacted monopsony first and then human capital, we would get 0.100 for monopsony and continue to get 0.093 after eliminating both. We could redo Table VI and eliminate search right before human capital. The result for search in panels (A)–(D) would be 0.010, 0.088, 0.011, and 0.063, leading to importance of human capital of 0.04, 0.02, 0.05, and 0.02.

³⁹We could have eliminated search frictions in steps like we did with premarket skills, first by setting $\text{var}(\delta_i) = 0$ and then setting $\lambda^e, \lambda^n \rightarrow \infty$. When we do this, it has no effect (to significant digits) in panel (A) or (C) of Table VI. In (B), it would have lowered inequality from 0.093 after eliminating monopsony to 0.092 when then eliminating variation in δ_i . It makes the biggest difference in the panel (D) experiment, where it would have lowered inequality from 0.087 after eliminating non-pecuniary aspects of jobs to 0.082 when then limiting the variation in δ_i .

- Variation in premarket skills by eliminating variation in wages within job type (i.e., $\sigma_\theta = \sigma_{vp} = 0$), but we hold the preference ordering across jobs exactly the same. We hold the preference ordering for jobs constant since we cannot non-parametrically identify how these would change when we change productivity as discussed in Section 4 on identification. Note that, for this counterfactual, one can see the necessity of matched employer-employee data. We still allow for variation in μ_j^p , the variation across job types that is common to all workers. In the case where we only eliminate differences across jobs (i.e., $\sigma_{vp} = 0$), we replace v_{ij}^p with its mean value across jobs.⁴⁰

- Non-pecuniary aspects of jobs by assuming workers choose among acceptable jobs only by comparing wages (i.e., $v_{ij}^u = \mu_j^u = 0$). However, for the reasons discussed above, we condition on acceptable jobs, by which we mean any job the worker would accept from non-employment in the real world. That is, any job rejected from non-employment in the full model would always be rejected in this counterfactual.

Immediately note in Table VI that in all four simulations, variation in premarket skill is the most important accounting for the vast majority of the variation in every decomposition—and this is roughly evenly split between the across-job component and the remaining part. We find that roughly 60% of the initial skills component comes from the common component.

The relative importance of search frictions and non-pecuniary aspects varies considerably across the four simulations, so this is clearly not an orthogonal decomposition. We do not view this feature as a weakness of our decomposition, but rather as a way of demonstrating how the aspects of the model interact. Most interesting is search frictions. Recall that the monopsony aspect explains 3% of the variation in every case and that monopsony power arises in the this model because of search frictions. The remaining amount explained by search frictions varies considerably across the specifications. It is about 1% in (A), 7% in (B), 6% in (C), and 26% in (D). This occurs because the order of the decomposition fundamentally alters the aspects of jobs for which workers are searching. In the base case, which corresponds to (B), workers are searching for good matches in four different dimensions: firm-specific productivity (μ_j^p), firm-specific non-pecuniary aspects (μ_j^u), match (individual \times firm type) productivity (v_{ij}^p), and match (individual \times firm type) utility (v_{ij}^u). When workers are searching for all four of these aspects, it leads to 7% variation in wages (0.093–0.086 in panel (B)). In experiment (A), we first eliminate initial skills. This means that we have eliminated the individual \times firm type productivity match (v_{ij}^p), which is a very important source of wage inequality. In this case, workers are searching for a good match in terms of the non-pecuniary aspects (both match-specific and firm-specific) and firm productivity only. Perhaps unsurprisingly, these aspects are not particularly important for wage inequality, and search only explains 1% of the inequality. In (C), we eliminate both the search for non-pecuniary aspects and the individual \times firm type productivity match, so that only firm-type productivity is being searched for. In this case, search explains roughly 6% of the variation. Experiment (D) is quite different in that we first eliminate non-pecuniary aspects of the job, which means we are eliminating both the firm-specific and individual \times firm-specific non-pecuniary aspects (μ_j^u and v_{ij}^u). In this case, workers are searching only for pecuniary aspects of the job and search frictions turn out to be very important. They explain 26% of the variation (or 29% if monopsony is included).

⁴⁰We include the mean value because this cannot be separately identified from θ_i , so this counterfactual is non-parametrically identified, while one for which $v_{ij}^p = 0$ would not be. We can still simulate this counterfactual and it gives a similar result.

Despite the other interesting results, the main takeaway remains that most of the variation in wages is explained by premarket differences in skills. This is comparable to Keane and Wolpin (1997), who estimated a model with compensating differentials, human capital accumulation, and Roy model inequality on U.S. data. They found that unobserved endowment heterogeneity accounts for 90% of the variance in lifetime utility. More recently, Bonhomme, Lamadon, and Manresa (2019) proposed a mixture model to estimate earnings distributions. Using Swedish data, they found that 80% of the explained wage variance is due to differences in worker components.⁴¹ Finally, our results are fairly different from Postel-Vinay and Robin (2002), who estimated a sequential auction search model on French data, where firms compete, in Bertrand style, over workers. They found that worker differences explain, at most, 40% of wage variation.

8.4. Model Decomposition of Utility Variation

Table VI quantifies the amount of variation in log wages. However, workers care about more than just wages. Another way of quantifying inequality is to look at variation in utility rather than just wages. We should emphasize that unlike the main decomposition of Table VI, this decomposition is not non-parametrically identified because we can only non-parametrically identify ordinal utility and not cardinal utility. Thus, this exercise should be viewed with some caution. For example, while obtaining some bounds on utility (measured in units of foregone wages) is possible by observing the wage cuts workers take, the results would be sensitive to tail behavior at the most favored job, which cannot be bounded non-parametrically. Despite these limitations, the results show how potentially important the other two channels might be. For comparability with the wage decomposition, we only use employed workers and we normalize utility to log wage equivalent units. Recall that flow utility is defined as

$$u_{ij}(R\psi_h) = \alpha \log(R\psi_h) + \mu_j^n + v_{ij}^n.$$

To put utility in the same units as log wages, we can rescale simply by renormalizing by dividing by α :

$$\tilde{u}_{ij}(R\psi_h) \equiv \log(R\psi_h) + \left(\frac{\mu_j^n + v_{ij}^n}{\alpha} \right).$$

Table VII presents the results of this decomposition. They differ substantially from the wage decomposition. This can be seen purely by looking at the variance in the first row—variance in preferences for non-pecuniary aspects contributes more to the variance than does variance in wages: The variance of log wages was 0.104, but the variance of log wages plus the non-pecuniary component is 0.234. Thus, the overall log wage variance accounts for only about 45% of the variation in utility.

Unlike the wage variance, where the premarket skill explained most of the variation, most of the variation in utility is explained by the interaction between compensating differentials and search. In this case, it does not matter much how we do it, but getting rid of either search or compensating differentials eliminates most of the variation in flow utility.

We also estimate a version of the statistical decomposition for the utility, which shows that the common firm component in utility is much larger than the common firm component in wages—both in an absolute sense and relative to the match component of each.

⁴¹Bødker (2019) estimated the model by Bonhomme, Lamadon, and Manresa (2019) on Danish data and found very similar results.

TABLE VII
COUNTERFACTUAL DECOMPOSITION OF VARIANCE OF FLOW UTILITY

(A)		(B)	
Total	0.234	Total	0.234
No Learning by Doing	0.219	No Learning by Doing	0.219
No Monopsony	0.210	No Monopsony	0.210
No Premarket Skill Variation Across Jobs	0.177	No Search Frictions	0.081
No Premarket Skill Variation at All	0.143	No Premarket Skill Variation Across Jobs	0.088
No Search Frictions	0.047	No Premarket Skill Variation at All	0.047
(C)		(D)	
Total	0.234	Total	0.234
No Learning by Doing	0.219	No Learning by Doing	0.219
No Monopsony	0.210	No Monopsony	0.210
No Non-Pecuniary Preferences for Jobs	0.087	No Non-Pecuniary Preferences for Jobs	0.087
No Premarket Skill Variation Across Jobs	0.048	No Search Frictions	0.061
No Premarket Skill Variation at All	0.006	No Premarket Skill Variation Across Jobs	0.047

Table VIII also shows how important the non-pecuniary aspect is for total utility as the last two terms explain roughly half of the variance. Interestingly, they are similar in magnitude.

8.5. Other Aspects of the Labor Market

The fact that search frictions and compensating differentials are not that important for wage inequality does not mean that they are not important for the labor market more generally. This can be seen from Table VII, but we also quantify their importance in a few other ways. Non-pecuniary aspects of a job are important for turnover. In roughly one-third of the competing offers in the simulation of our full model, the workers would have made a different choice if they only cared about wages. The consequences are large as workers earn a wage that is about 0.20 log points lower as a result of these choices.

In our model, search frictions are essential for explaining turnover as there would be no turnover without them. To quantify, wages would be about 22% higher in the absence of search frictions. Of this 22%, roughly 10% is due to the negotiation, and 12% is due to the improved job match.

TABLE VIII
LINEAR UTILITY VARIANCE DECOMPOSITION

$\text{Var}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}))$	0.234
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \log(R_{ij(it)}) - \log(\pi_{ij(it)}))$	0.008
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \theta_i)$	0.039
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \mu_{j(it)}^p)$	0.007
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), v_{ij(it)}^p)$	0.037
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \log(\psi_{h(it)}))$	0.011
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \mu_{j(i,t)}^n/\alpha)$	0.067
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), v_{ij(it)}^n/\alpha)$	0.065

9. ROBUSTNESS AND IDENTIFICATION IN PRACTICE

While we formally prove identification above, there is also the question of what identifies the model in practice. A closely related question is how robust the results are to alternative assumptions about the parameterizations. In this section, we provide insight into both questions through several exercises to further understand the main results. We have:

1. re-estimated the model using alternative auxiliary models,
2. estimated restricted versions of the model, where we eliminate various components,
3. measured the sensitivity of the auxiliary and counterfactuals to the structural parameters,
4. carried out alternative normalization of the model,
5. allowed for more complementarity between firms and workers.

Due to space restrictions, only the first exercise is presented in the actual article, while the rest are discussed in Appendix E, which is available on our websites. We discuss the first exercise in the next subsection and then briefly summarize the results of the other exercises following that.

9.1. *Alternative Auxiliary Parameters*

Specification tests are common with this type of model, where over-identifying auxiliary parameters are used to test it. Because our model is highly stylized and the data set extremely large, we would almost certainly reject any formal over-identification test. This does not mean that our stylized model does not do a good job capturing the key features of our models. We think the more interesting question is not whether we can fit alternative auxiliary parameters, but rather, would a model that targeted alternative parts of the data lead to different results. With this in mind, we do a sensitivity analysis, where we continue to use an exactly identified model, but replace some of the targeted auxiliary parameters with others. After fitting the model, we show that the results are robust to the alternative specifications. Given the computational time in doing this, we use fewer worker simulations. In our main estimates we simulate 1,580,000 worker histories, but in these cases we use 158,000. The simulation error is small relative to the differences in changing the counterfactuals. For this reason, the fit and counterfactuals for the “Baseline” in Tables IX and X do not exactly correspond to the Base model in Tables IV and VI, but they are very close.

Table IX shows the fit from the three models using the alternative auxiliary parameters. For all the models, the fit to the data is very good.

In Alternative 1, we use different auxiliary parameters related to turnover rates. Rather than use durations for job, employment spells, and non-employment spells, we use hazards. We construct the covariance between wages and non-employment durations by weighting by spell rather than by person. Rather than using the variance of non-employment and job-to-job transitions, we look at the tails of these distributions—the fraction of people with neither. Supplemental Material Appendix C contains detailed descriptions of the new alternative auxiliary parameters, while Table IX presents that the fit of the model is excellent.⁴² Table X presents the counterfactual decomposition corresponding to panel (A) of Table VI. We see that the decompositions are almost unaffected,

⁴²Appendix E, available on our websites, contains the parameter estimates.

TABLE IX
ESTIMATION UNDER ALTERNATIVE AUXILIARY PARAMETERS: MODEL FIT

	Baseline		Alternative 1		Alternative 2		Alternative 3	
	Alt Aux	Sim	Alt Aux	Sim	Alt Aux	Sim	Alt Aux	Sim
Avg. Length Emp. Spell or Avg. Hazard Emp. $\times 1000$	377	2.18	2.13	380		380		380
Avg. Length Non-emp. Spell or Avg. Hazard Non-emp. $\times 1000$	91.4	7.64	7.74	90.7		90.7		91.6
Avg. Length Job Spell or Avg. Hazard Job $\times 1000$	106	5.41	5.48	108		108		107
Sample Mean w_{it}^r	4.50		4.56	4.50		4.50		4.46
Between Persons $\times 100$	8.03		8.03	8.10		8.10		8.01
Between Jobs $\times 100$	2.87		2.82	2.86		2.86		2.89
Within Job $\times 100$	1.49		1.53	1.49		1.49		1.52
$E(\tilde{w}_{it}^r \tilde{w}_{-it}^r) \times 100$ or $E(\Delta \tilde{w}_{it}^r Q1 \rightarrow Q4)$	0.77		0.79	0.35		0.37		0.78
$E(\tilde{r}_{-it}^r \tilde{w}_{it}^r) \times 100$ or $\frac{E(\Delta \tilde{w}_{it}^r Q1 \rightarrow Q4)}{E(\Delta \tilde{w}_{it}^r Q1 \rightarrow Q4)}$	0.69		0.69	-0.93		-0.93		0.69
$\text{cov}(\tilde{r}_{-it}^r, \tilde{S}_{it}^r) \times 100$ or $E(\tilde{S}_{it}^r \tilde{r}_{-it}^r) \times 100$	8.18		8.08	3.58		3.64		8.15
Fraction Wage Drops or $E(\Delta \tilde{w}_{it}^r JJ)$	0.40		0.41	0.40		0.40		0.037
Coeff Exper $\times 100$ or Coeff Exper $\times 100$	2.48		2.64	2.43		2.43		1.65
Coeff Exper ² $\times 1000$ or Coeff Exper ² $\times 1000$	-0.291		-0.278	-0.294		-0.294		-0.287
Coeff Tenure ² $\times 1000$ or Coeff Tenure $\times 100$	-0.460		-0.450	-0.463		-0.463		0.405
Var(Non-employment) or Never Work	16,000	0.077	0.080	16,047		16,047		16,161
Cov(\bar{w}_i , Non-employment) or Cov(D_{ik}^n , \bar{w}_i)	-3.42	-2.97	-2.85	-3.39		-3.39		-3.39
Var(employment Dur) or Never Job to Job	102,000	0.51	0.54	103,337		103,337		103,154
Invol Job to Job	0.205		0.202	0.210		0.210		0.204

TABLE X
ESTIMATION UNDER ALTERNATIVE AUXILIARY PARAMETERS: SENSITIVITY OF COUNTERFACTUALS

	Baseline	Alternative 1	Alternative 2	Alternative 3
Full Model	0.105	0.105	0.105	0.105
No Learning by Doing	0.096	0.095	0.096	0.099
No Monopsony	0.093	0.092	0.093	0.097
No Premarket Across	0.049	0.052	0.050	0.055
No Premarket Total	0.008	0.008	0.006	0.007
No Search	0.007	0.007	0.006	0.007

with only minor differences present compared to the baseline results. We also performed the decompositions in other orders with a similar result.

Our next specification, Alternative 2, uses different auxiliary parameters to characterize the establishment-level variables. The first two are motivated by the symmetry result in [Card, Heining, and Kline \(2013\)](#) who found that the wage gains associated with moving from low paying firms to high paying firms are similar to the wage losses associated with moving from high paying firms to low ones. In particular, our estimated baseline model implies a high degree of match effects.⁴³ However, [Card, Heining, and Kline \(2013\)](#) concluded that match effects are not very important in wage equations using an AKM type of model. We want to show that we can reconcile our model with their findings. This depends on two things: the magnitude of the firm effects (related to f_p) and the amount of selection on wages in the model (related to α). That is, if $\alpha = 0$, selection would depend only on preferences for non-pecuniary aspects of the job, and we would expect to see symmetry. We divide our firms into four groups and estimate the wage change in moving from the highest to the lowest and the wage change in moving from the lowest to the highest. We match on the latter and on the ratio between them. As a comparison, the data moment for the ratio is -0.93 in Alternative 2, while in the baseline simulation, it is -0.42 . While this is not very close to symmetry, it is negative, which might seem surprising given that we find an important role for the premarket skill variation across jobs. Compensating differences is an important explanation for this. The fact that people are switching jobs, in large part for non-pecuniary reasons, makes the pecuniary match variable less important. For the other auxiliary parameter, we replace the parameter picking up preference for the establishment \tilde{r}_{-itj} , with a reasonable alternative, \tilde{h}_{-itj} . Again, Supplemental Material Appendix C describes the three new auxiliary parameters in detail. As [Table IX](#) shows, the model is able to match the two statistics from [Card, Heining, and Kline \(2013\)](#). The counterfactual decompositions shown in [Table X](#) do not change much from the base case. In particular, the importance of the premarket skills across jobs decreases only very slightly.

Finally, in Alternative 3, we change the set of auxiliary parameters used to estimate the return to human capital (b_1 and b_2), bargaining power (β), and the weight on wage (α). Rather than estimate a fixed effect regression on experience, experience squared, and tenure squared, we follow [Altonji and Shakotko \(1987\)](#) and estimate a model on experience, experience squared, and a linear tenure effect. Again, details are provided in Supplemental Material Appendix C. The fit is very good and, in general, the counterfactual decompositions do not change much, but there are some changes as human

⁴³Recall that match effects accounted for one-third of the variance of log wages in our statistical decomposition and were of similar magnitude in the model decomposition.

capital becomes less important; most importantly, the monopsony effect becomes even smaller.

Note that [Hornstein, Krusell, and Violante \(2011\)](#) pointed out that search models without on-the-job search have a hard time matching both the observed wage distribution and the fact that unemployed workers rapidly transition out of unemployment. Intuitively, if workers transition rapidly out of unemployment in a model without on-the-job search, it implies that the value of waiting for a better offer is low, which means that there cannot be much variation in wages across jobs. However, they also noted that this can be resolved if compensating differentials are important and are negatively correlated with wages. This is precisely the result we get in Alternatives 2 and 3. The key statistic in their framework is the ratio between the mean acceptable wage and the minimum acceptable wage. In our model, this varies slightly across ages, but the average value across all potential workers in our simulations is approximately 1.67.⁴⁴

9.2. Other Results

Appendix E, available on our websites, presents the details of the other robustness exercises described at the beginning of this section. We will briefly summarize the results here.

Estimate Restricted Versions of the Model

We estimate seven restricted versions of the model, where various model components are taken out. We think that the following are the main takeaways. First, eliminating Pre-market Skills Across Jobs ($\sigma_{vp} = 0$) causes the model to miss wildly on the between-job variance. The reason that the model does not generate more between-job variance by increasing f_p is that this would cause $E(\tilde{w}_{it}\tilde{w}_{-it})$ to overshoot. When we eliminate all variation in Pre-market Skills ($\gamma_\theta = \sigma_\theta = \sigma_{vp} = 0$), the model misses in both the between-job and between-person variances. Second, eliminating non-pecuniary aspects of jobs ($f_u = \text{var}(v_{ij}^u) = 0$ and $\sigma_{vp} = 0$) leads the model to miss in many dimensions. The two most important ones are the fraction of wage drops and the correlation across workers in preferences of jobs $E(\tilde{S}_{itj}\tilde{r}_{-itj})$.

Sensitivity to the Structural Parameters

The sensitivity of the auxiliary parameters and counterfactuals to the structural parameters confirms that, while there are interactions and some complications, [Table II](#) is a reasonable approximation of how identification works in practice.⁴⁵

Alternative Normalization of the Model

The main results are robust to the alternative parameterization, with the exception of a few results that have to change almost mechanically.

⁴⁴Note that some individuals would not accept any offers. We exclude these people from our sample. Otherwise it is unweighted across all potential entrants and calculates the wages based on the immediate job taken from non-employment conditional on alternative levels of human capital. Note that the minimum wage is the lowest such wage workers would take, which is not typically the wage at the lowest utility establishment. The ratio varies from 1.668 for the least experienced worker to 1.679 for the most experienced.

⁴⁵We also tried constructing the sensitivity matrix of structural parameters to auxiliary parameters following [Andrews, Gentzkow, and Shapiro \(2017\)](#). Unfortunately, this did not work well as our model is not differentiable, and different numerical approximations of the derivatives led to quite different results.

Complementarity Between Firms and Workers

Finally, we estimate the model with a more general production function, which imposes complementarity between worker ability (θ_i) and firm productivity (μ_j^p). The main conclusion is the same, that variation in premarket skills explains the bulk part of variation in wages, but other results differ slightly.⁴⁶

10. CONCLUSIONS

The goal of this paper is to estimate the primary drivers behind wage variation. For this purpose, we have developed and estimated a general labor market model. The model includes components from the Roy model (premarket productivity differences across workers), search frictions, compensating differentials (preferences for non-pecuniary aspects of jobs), and general human capital acquired on the job. Wages are determined endogenously through a bargaining process.

We add to the literature by investigating non-parametric identification of the model given revealed preferences from job-to-job transitions and wages. We show that almost all aspects of the model are identified. Two important exceptions are bargaining power and wages in jobs that workers would never take. The latter is expected, while the former was unanticipated by us and is due to the fact that revealed preferences can identify the preference order, but not the differences in utility.

We estimate the model on Danish matched employer-employee data using indirect inference. The model fit is very good. Using the estimated model, we show that premarket skills are the most important driver behind wage variation, explaining between 61% and 85% of the total. Search friction, LBD human capital, and compensating differentials matter to a lesser extent. These factors also interact in important ways.

Simulations show that non-pecuniary aspects of jobs and search frictions are both important for other aspects of the data. One-third of all encounters between two firms and a worker would have changed transition path if workers only cared about wages and not the non-pecuniary aspects of a job. On average, workers would earn wages that are 0.20 log points higher if they did not care about non-pecuniary aspects of the job but only about wages, and 0.22 log points higher if there were no search frictions. We show that variance in utility is primarily driven by search frictions and compensating differentials and not premarket skill heterogeneity. These conclusions generally hold across various specification and robustness checks.

We should also emphasize that we have intentionally kept the model very simple. Computation of this model is quite fast, so adding other features into this framework is straightforward computationally (although likely difficult to identify). Additional features that come to mind are job/occupation/tenure/task-specific human capital, non-pecuniary costs of switching jobs, learning about the match productivity, learning about absolute ability, relaxing the assumption that employers know the workers' preferences for non-pecuniary aspects of the job, and allowing these preferences to change over time.

We conclude that future research aimed at further understanding wage determination and labor market transitions should concern all of the four components incorporated in our model.

⁴⁶Most importantly, search frictions lower inequality. The reason is, by assumption, matching components are more important for high skilled workers. This means that the worse matches resulting from search frictions harm high skilled workers more than low skilled ones.

REFERENCES

- ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67 (2), 251–333. [1033,1034]
- ABOWD, J. M., K. L. MCKINNEY, AND I. M. SCHMUTTE (2018): "Modeling Endogenous Mobility in Earnings Determination," *Journal of Business & Economic Statistics*, 0 (0), 1–14. [1033]
- ALTONJI, J. G., AND C. R. PIERRET (2001): "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 116 (1), 313–350. [1031]
- ALTONJI, J. G., AND R. A. SHAKOTKO (1987): "Do Wages Rise With Job Seniority?" *Review of Economic Studies*, 54 (3), 437–459. [1065]
- ANDERSEN, T., AND M. SVARER (2007): "Flexicurity—Labour Market Performance in Denmark," CESifo Working Paper Series 2108, CESifo Group Munich. [1051]
- ANDREWS, I., M. GENTZKOW, AND J. SHAPIRO (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, 132 (4), 1553–1592. [1066]
- ANDREWS, M. J., L. GILL, T. SCHANK, AND R. UPWARD (2008): "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society A*, 171 (3), 673–697. [1033]
- BAGGER, J., AND R. LENTZ (2018): "An Equilibrium Model of Wage Dispersion and Sorting," *Review of Economic Studies*, 86 (1), 153–190. Working Paper. [1033]
- BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J.-M. ROBIN (2014): "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 104 (6), 1551–1596. [1033,1035-1038,1040,1054]
- BECKER, D. (2009): "Non-Wage Job Characteristics and the Case of the Missing Margin," Federal Trade Commission. [1034]
- BERTHEAU, A., H. BUNZEL, M. HEJLESEN, AND R. VEJLIN (2016): "Labor Market Transitions Over the Business Cycle," Unpublished Working Paper, Department of Economics, Aarhus University. [1051]
- BOBBIO, E., AND H. BUNZEL (2018): "The Danish Matched Employer-Employee Data," Economics Working Paper 2018-3, Department of Economics and Business Economics, Aarhus University. [1048]
- BØDKER, J. (2019): "Wage and Mobility Dynamics of the Danish Labor Market," Report. [1061]
- BOERI, T., A. BRUGIAVINI, AND L. CALMFORS (2001): *The Role of Unions in the Twenty-First Century*. Oxford University Press. [1051]
- BONHOMME, S., T. LAMADON, AND E. MANRESA (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87 (3), 699–739. [1033-1035,1041,1061]
- BURDETT, K., AND M. COLES (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71 (5), 1377–1404. [1051]
- CAHUC, C., F. POSTEL-VINAY, AND J.-M. ROBIN (2006): "Wage Bargaining With On-the-Job Search: Theory and Evidence," *Econometrica*, 74 (2), 323–364. [1033,1035,1038]
- CALDWELL, S., AND N. HARMON (2019): "Outside Options, Bargaining and Wages: Evidence From Coworker Networks," Report. [1051]
- CARD, D., A. CARDOSO, AND P. KLINE (2015): "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *The Quarterly Journal of Economics*, 131 (2), 633–686. [1033]
- CARD, D., J. HEINING, AND P. KLINE (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128 (3), 967–1015. [1033,1055,1065]
- DAHL, C. M., D. LE MAIRE, AND J. R. MUNCH (2013): "Wage Dispersion and Decentralization of Wage Bargaining," *Journal of Labor Economics*, 31 (3), 501–533. [1051]
- DEY, M. S., AND C. J. FLINN (2005): "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 73 (2), 571–627. [1034,1035]
- (2008): "Household Search and Health Insurance Coverage," *Journal of Econometrics*, 145 (1–2), 43–63. [1034]
- ECKSTEIN, Z., AND G. VAN DEN BERG (2007): "Empirical Labor Search," *Journal of Econometrics*, 136 (2), 531–564. [1033]
- FLINN, C., AND J. HECKMAN (1982): "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 18, 115–168. [1036,1047]
- FRENCH, E., AND C. TABER (2011): "Identification of Models of the Labor Market," in *The Handbook of Labor Economics*, Vol. 4A, ed. by D. C. O. Ashenfelter. [1034]
- GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118. [1052]
- GRUETTER, M., AND R. LALIVE (2004): "The Importance of Firms in Wage Determination," IEW—Working Papers 207, Institute for Empirical Research in Economics—IEW. [1033]

- HAGEDORN, M., T. H. LAW, AND I. MANOVSKII (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85, 29–65. [1034]
- HALL, R. E., AND A. B. KRUEGER (2012): "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search," *American Economic Journal: Macroeconomics*, 4 (4), 56–67. [1035]
- HECKMAN, J. (1990): "Varieties of Selection Bias," *American Economic Review, Papers and Proceedings*, 80, 313–318. [1034]
- HECKMAN, J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320. [1035]
- HECKMAN, J. J., AND B. E. HONORÉ (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58 (5), 1121–1149. [1034,1040]
- HECKMAN, J. J., AND C. R. TABER (2008): "Roy Model," Chapter Roy model. Palgrave Macmillan. [1033]
- HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, Vol. 1. Amsterdam: North Holland, 307–458. [1033]
- HOFFMAN, F. (2018): "An Empirical Model of Life-Cycle Earnings and Mobility Dynamics," Unpublished Manuscript. [1034]
- HORNSTEIN, A., P. KRUSELL, AND G. L. VIOLANTE (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *The American Economic Review*, 101, 2873–2898. [1066]
- HUGGETT, M., G. VENTURA, AND A. YARON (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101 (7), 2923–2954. [1034]
- JOVANOVIĆ, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87 (5), 972–990. [1031]
- KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *The Journal of Political Economy*, 105 (3), 473–522. [1034,1035,1061]
- LENTZ, R., S. PIYAPROMDEE, AND J.-M. ROBIN (2018): "On Worker and Firm Heterogeneity in Wages and Employment Mobility: Evidence From Danish Register Data." [1033,1035]
- LISE, J., C. MEGHIR, AND J.-M. ROBIN (2016): "Matching, Sorting and Wages," *Review of Economic Dynamics*, 19, 63–87. Special Issue in Honor of Dale Mortensen. [1033]
- LUND, C., AND R. VEJLIN (2016): "Documenting and Improving the Hourly Wage Measure in the Danish IDA Database," *The Danish Journal of Economics*, 1, 1–35. [1048]
- POSTEL-VINAY, F., AND J.-M. ROBIN (2002): "Equilibrium Wage Dispersion With Worker and Employer Heterogeneity," *Econometrica*, 70 (6), 2295–2350. [1033,1038,1056,1061]
- POSTEL-VINAY, F., AND H. TURON (2010): "On-the-Job Search, Productivity Shocks, and the Individual Earnings Process*," *International Economic Review*, 51 (3), 599–629. [1037]
- ROBIN, J.-M. (2011): "On the Dynamics of Unemployment and Wage Distributions," *Econometrica*, 79 (5), 1327–1355. [1033]
- ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, XLIII, 959–988. [1033]
- ROSEN, S. (1987): "The Theory of Equalizing Differences," in *Handbook of Labor Economics*, Vol. 1, 641–692. [1033]
- ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (2), 135–146. [1033,1040]
- SMITH, A. (1776): *An Inquiry into the Nature and Causes of the Wealth of the Nations*, ed. by W. Strahan and T. Cadell. [1034]
- SØRENSEN, T., AND R. VEJLIN (2013): "The Importance of Worker, Firm and Match Effects in the Formation of Wages," *Empirical Economics*, 45 (1), 435–464. [1033,1055]
- SORKIN, I. (2018): "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133 (2), 1331–1393. [1034]
- SULLIVAN, P., AND T. TO (2014): "Search and Nonwage Job Characteristics," *Journal of Human Resources*, 49 (2), 472–507. [1034]
- TABER, C., R. VEJLIN (2020): "Supplement to 'Estimation of a Roy/Search/Compensating Differential Model of the Labor Market'," *Econometrica Supplemental Material*, 88, <https://doi.org/10.3982/ECTA14441>. [1049]
- TOPEL, R. (1991): "Specific Capital, Mobility, and Wages: Wages Rise With Job Seniority," *The Journal of Political Economy*, 99 (1), 145–176. [1053]
- WEISS, Y. (1986): "The Determination of Lifecycle Earnings: A Survey," in *Handbook of Labor Economics*, Vol. 1, ed. by O. Ashenfelter and D. Card. Amsterdam: North-Holland, 603–640. [1033]

Co-editor Liran Einav handled this manuscript.

Manuscript received 31 May, 2016; final version accepted 2 December, 2019; available online 22 January, 2020.