Estimation of a Roy/Search/Compensating Differential Model of the Labor Market

by Christopher Taber, Rune Vejlin
Econometrica, Volume 88, Issue 3, May 2020, Pages 1031–1069

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 350, Spring 2022

- Model: captures key components of the Roy model;
  - A search model;
  - Compensating differentials; and
  - Human capital accumulation on-the job.
- Establish which components of the model can be non-parametrically identified and which ones cannot.
- Estimate the model and use it to assess the relative contribution of the different factors for overall wage inequality.
- Variation in premarket skills (the key feature of the Roy model) is the most important component to account for the majority of wage variation.

- 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).

- Goal of this study is to uncover the contribution of these different components and determine how they interact to produce overall wage inequality.

- Substantial interaction between the other components:
- 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.

1. Introduction

- Roy model: absolute and comparative advantage, in which some workers earn more than others as a result of different skill levels at labor market entry.

- In 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 their preferred job.

- Labor market frictions can lead to heterogeneity in wages for two different reasons:
    1. Some workers may work for firms for which they are a better match and earn higher wages; and
    2. Monopsony.

- Bargaining position of the worker depends on their outside option
- Two equally skilled workers at the same firm may earn different wages.
- 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.
- 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.

2. Relation to Other Work

- Paper builds on the literature: adds non-pecuniary aspects of jobs and much more heterogeneity in premarket skills.

- 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).

- We add search frictions to this model and consider mobility using panel data.

- 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

Basic Environment: Firms and Workers

- There are a finite number of job types, $j = 1, \ldots, J$
- $j = 0$ denotes nonemployment.
- 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.
- Allow for a large number of individuals in the economy.
- N individuals indexed by $i = 1, \ldots, N$.
- Focus on a generic individual in the data and use the i subscript to make it clear which variables vary across individuals.

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

Learning by Doing (LBD)

- LBD human capital takes on a discrete set of values $\psi_0, \dots, \psi_H$.

- 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.

- 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 u_{i0h}.

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.

Wage Determination

- When a worker receives an outside offer, wages are determined by a form of generalized Nash bargaining between the two firms.
- The object of negotiation is the human capital rental rate R, which is the price per unit of LBD human capital.
- When human capital is augmented, the wage is not renegotiated, but automatically rises from $R\psi_h$ to $R\psi_{h+1}$.
- $V_{ijh}(R)$: 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.

- A match is formed if $V_{ijh}(\pi_{ij}) > V_{i0h}$.

- 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 $R_{ij\ell h_0}\psi_h$.

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

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

where β is the worker's bargaining power.

- 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

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

- 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

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

- 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.

- 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, calculate the value functions $V_{ijh}(R)$ and V_{i0}, as there are no closed form solutions for the wage.

- Define

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

$$\Lambda^{n}_{i0h} \equiv \sum_{\{\ell : V_{i\ell h}(\pi_{i\ell}) > V_{i0h}\}} \lambda^{n}_{\ell},$$

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.

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:

$$\begin{aligned}
&\left(\rho + \delta_i + \lambda_h + \Lambda^e_{ijh}(R)\right) V_{ijh}(R) \\
&\quad = u_{ij}(R\psi_h) + \left(\sum_{\{\ell : V_{ijh}(R) < V_{i\ell h}(\pi_{i\ell}) \le V_{ijh}(\pi_{ij})\}} \lambda^e_\ell \left[\beta V_{ijh}(\pi_{ij}) + (1-\beta) V_{i\ell h}(\pi_{i\ell})\right] \right) \\
&\qquad + \left(\sum_{\{\ell : V_{i\ell h}(\pi_{i\ell}) < V_{ijh}(\pi_{ij})\}} \lambda^e_\ell \left[\beta V_{i\ell h}(\pi_{i\ell}) + (1-\beta) V_{ijh}(\pi_{ij})\right] \right) \\
&\qquad + \delta_i V^*_{i0h} + \lambda_h \max\left(V_{ijh+1}(R), V_{i0h+1}\right).
\end{aligned}$$

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

$$(\rho + \delta_i + \Lambda^e_{ijh}(R))V_{ijH}(R)$$
$$= u_{ij}(R\psi_H) + \left(\sum_{\{\ell : V_{ijH}(R) < V_{i\ell H}(\pi_{i\ell}) \leq V_{ijH}(\pi_{ij})\}} \lambda^e_\ell [\beta V_{ijH}(\pi_{ij}) + (1-\beta) V_{i\ell H}(\pi_{i\ell})] \right)$$
$$+ \left(\sum_{\{\ell : V_{i\ell H}(\pi_{i\ell}) < V_{ijH}(\pi_{ij})\}} \lambda^e_\ell [\beta V_{i\ell H}(\pi_{i\ell}) + (1-\beta) V_{ijH}(\pi_{ij})] \right) + \delta_i V^*_{i0H}.$$

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

$$(\rho + \Lambda^n_{i0h})V_{i0h} = u_{i0h} + \sum_{\{\ell : V_{i\ell h}(\pi_{i\ell}) > V_{i0h}\}} \lambda^n_j [\beta V_{i\ell h}(\pi_{i\ell}) + (1-\beta) V_{i0h}].$$

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

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

- 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.
- 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

- 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 nonemployment.

(d) If a worker is indifferent in terms of two options, 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.

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 $\bar{T}$.

- **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.

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

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

- **ASSUMPTION 4**:

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

- The point of this assumption is to keep the model interesting.

- **THEOREM 1**: Under Assumptions 1–4, with the data generated by the model exposited in Section 3, we can identify λ_A^n, λ_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.

4.2. Wage Components

- For any worker who is currently working, there are four different states which are relevant for their wages (denoted here as functions of the individual and time): their current employer $j(i,t)$, their current level of LBD human capital $h(i,t)$, the outside option when their current rental rate was negotiated $\ell(i,t)$, and the level of LBD human capital when the current rental rate was negotiated $h_0(i,t)$.

- 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.

- 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.
- **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., $\bar{T} > 8$).

- **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.

4.3. Non-Identification of β

- 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, does 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:

If, with probability 1 for the relevant groups:

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

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.

(continued on next slide)

- **THEOREM 3, CONT'D**:

  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 generically 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.

- 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 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

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

- 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.

5. Econometric Specification/Parameterization

- Even though the model is mostly non-parametrically identified, estimating it nonparametrically 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

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

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.

- 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:

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

- 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:

$$\mu_j^u = f_u\left[U_1(j) + f_{u,p} U_2(j)\right],$$
$$\mu_j^p = f_p\left[f_{u,p} U_1(j) + U_2(j)\right],$$

where $U_1(j)$ and $U_2(j)$ are distributed as discretely uniform across $[-1,1]$.

- Human capital evolves as

$$\log(\psi_h) = b_1 h + b_2 h^2 + b_3 h^3,$$

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.

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

- 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.

- We take a simple specification for the value of non-employment by assuming

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

with $\nu_{i0}^n \sim N(0, \sigma_\nu^2)$.

- 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.

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

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

6. Data and Danish Institutional Features

6.1. Sample Selection Criteria

- 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.

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

6.2. Descriptive Statistics

- 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.
- The workers have 12 years of education on average. However, this moderately changes over the sample period, since entering workers are better 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

7. Auxiliary Model

- **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.

- **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 $\widetilde{w}_{it}$ as an average wage residual at the given firm relative to the same individual at other firms.

- **Wage Dynamics**: We run a log wage regression on experience, experience squared, and tenure squared with individual × establishment fixed effects.

- **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.

- 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.

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_{υ^p}	Skills
Fraction Wage Drops	Weight on log wage: α	Non-pecuniary aspects
$E(\widetilde{S}_{i\ell j}\widetilde{r}_{-i\ell j})$	Firm utility parameter: f_u	
$E(\widetilde{w}_{i\ell j}\widetilde{r}_{-i\ell j})$	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
$E(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j})$	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_ν	
Cov($\overline{w}$, non-emp dur)	Worker ability cont to flow utility: γ_θ	

8. Results

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.
- Second, the standard deviation of the match productivity term, σ_{v^p}, is estimated to be 0.211, which implies that match effects are important in our model.
- 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.

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_{υ^p}	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×productivity parameter	0.467	(0.532)
σ_ν	Std. dev. of idiosyncratic non-employment utility	0.351	(0.012)
γ_θ	Worker ability contribution to flow utility	−0.282	(0.030)

- 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

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
$E(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}) \times 100$	0.77 (0.004)	0.77
$E(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}) \times 100$	0.69 (0.005)	0.69
$\text{cov}(\widetilde{r}_{-i\ell j}, S_{i\ell j}) \times 100$	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
Coef Tenure^2 $\times$ 1000	−0.460 (0.003)	−0.460
Var(Non-employment)	16,000 (50.39)	16,150
Cov($\overline{w}_i$, 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

8.2. Statistical Decomposition of Wage Variation

- 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

$$\begin{aligned}\log(w_{it}) &= \log(R_{ij(it)}) + \log(\psi_{h(it)}) + \xi_{it} \\ &= \left[\log(R_{ij(it)}) - \log(\pi_{ij(it)})\right] + \log(\pi_{ij(it)}) + \log(\psi_{h(it)}) + \xi_{it} \\ &= \left[\log(R_{ij(it)}) - \log(\pi_{ij(it)})\right] + \theta_i + \mu^p_{j(it)} + v^p_{ij(it)} + \log(\psi_{h(it)}) + \xi_{it}.\end{aligned}$$

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

$$\begin{aligned}\mathrm{Var}(\log(w_{it})) = {} & \mathrm{Cov}(\log(w_{it}), \log(R_{ij(it)}) - \log(\pi_{ij(it)})) \\ & + \mathrm{Cov}(\log(w_{it}), \theta_i) + \mathrm{Cov}(\log(w_{ijt}), \mu^p_{j(it)}) \\ & + \mathrm{Cov}(\log(w_{it}), v^p_{ij(it)}) + \mathrm{Cov}(\log(w_{it}), \log(\psi_{h(it)})) \\ & + \mathrm{Cov}(\log(w_{it}), \xi_{it}),\end{aligned}$$

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.

- 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 $v^{p}_{ij(it)}$.

- The other components are non-trivial but clearly smaller.

Table V. Linear Wage Variance Decomposition

$\text{Var}(\log(w_{it}))$	0.124
$\text{Cov}(\log(w_{it}), \log(R_{ij(it)}) - \log(\pi_{ij(it)}))$	0.005
$\text{Cov}(\log(w_{it}), \theta_i)$	0.044
$\text{Cov}(\log(w_{ijt}), \mu^p_{j(it)})$	0.008
$\text{Cov}(\log(w_{it}), v^p_{ij(it)})$	0.042
$\text{Cov}(\log(w_{it}), \log(\psi_{h(it)}))$	0.006
$\text{Cov}(\log(w_{it}), \xi_{it})$	0.019

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.
- We simulate four different sequences of decompositions (A)–(D) in Table VI.

Table VI. Counterfactual Decomposition of Variance of Log Wages

(A)

Total	0.104
No Learning by Doing	0.096
No Monopsony	0.093
No Premarket Skill Variation Across Jobs	0.050
No Premarket Skill Variation at All	0.008
No Search Frictions	0.007

(B)

Total	0.104
No Learning by Doing	0.096
No Monopsony	0.093
No Search Frictions	0.086
No Premarket Skill Variation Across Jobs	0.049
No Premarket Skill Variation at All	0.007

(C)

Total	0.104
No Learning by Doing	0.096
No Monopsony	0.093
No Non-Pecuniary Aspects of Jobs	0.087
No Premarket Skill Variation Across Jobs	0.048
No Premarket Skill Variation at All	0.006

(D)

Total	0.104
No Learning by Doing	0.096
No Monopsony	0.093
No Non-Pecuniary Aspects of Jobs	0.087
No Search Frictions	0.061
No Premarket Skill Variation Across Jobs	0.047

- 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., $\lambda^e, \lambda^n \to \infty$).

  - Variation in premarket skills by eliminating variation in wages within job type (i.e., $\sigma_\theta = \sigma_{v^p} = 0$), but we hold the preference ordering across jobs exactly the same.

  - Non-pecuniary aspects of jobs by assuming workers choose among acceptable jobs only by comparing wages (i.e., $v_{ij}^u = \mu_j^u = 0$).

- 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.

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.

- 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:

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

- Table VII presents the results of this decomposition.
- They differ substantially from the wage decomposition.
- 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.

Table VII. Counterfactual Decomposition of Variance of Flow Utility

(A)	
Total	0.234
No Learning by Doing	0.219
No Monopsony	0.210
No Premarket Skill Variation Across Jobs	0.177
No Premarket Skill Variation at All	0.143
No Search Frictions	0.047

(B)	
Total	0.234
No Learning by Doing	0.219
No Monopsony	0.210
No Search Frictions	0.081
No Premarket Skill Variation Across Jobs	0.088
No Premarket Skill Variation at All	0.047

(C)	
Total	0.234
No Learning by Doing	0.219
No Monopsony	0.210
No Non-Pecuniary Preferences for Jobs	0.087
No Premarket Skill Variation Across Jobs	0.048
No Premarket Skill Variation at All	0.006

(D)	
Total	0.234
No Learning by Doing	0.219
No Monopsony	0.210
No Non-Pecuniary Preferences for Jobs	0.087
No Search Frictions	0.061
No Premarket Skill Variation Across Jobs	0.047

8.5. Other Aspects of the Labor Market

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
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \mu^p_{j(it)})$	0.007
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \upsilon^p_{ij(it)})$	0.037
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \log(\psi_{h(it)}))$	0.011
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \mu^n_{j(i,t)}/\alpha)$	0.067
$\text{Cov}(\tilde{u}_{ij}(R_{ij(it)}\psi_{h(it)}), \upsilon^n_{ij(it)}/\alpha)$	0.065

9. Robustness and Identification in Practice

9.1. Alternative Auxiliary Parameters

- 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.

Table IX. Estimation Under Alternative Auxiliary Parameters: Model Fit

	Baseline	Alternative 1 Alt Aux	Alternative 1 Sim
Avg. Length Emp. Spell or Avg. Hazard Emp. $\times 1000$	377	2.18	2.13
Avg. Length Non-emp. Spell or Avg. Hazard Non-emp. $\times 1000$	91.4	7.64	7.74
Avg. Length Job Spell or Avg. Hazard Job $\times 1000$	106	5.41	5.48
Sample Mean $w_{i\ell jt}$	4.50		4.56
Between Persons $\times 100$	8.03		8.03
Between Jobs $\times 100$	2.87		2.82
Within Job $\times 100$	1.49		1.53
$E(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}) \times 100$ or $E(\Delta\overline{w_{i\ell j}} \mid Q1 \to Q4)$	0.77		0.79
$E(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}) \times 100$ or $\frac{E(\Delta\overline{w_{i\ell j}} \mid Q4 \to Q1)}{E(\Delta\overline{w_{i\ell j}} \mid Q1 \to Q4)}$	0.69		0.69
$\mathrm{cov}(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$ or $E(\widetilde{S}_{i\ell j}\widetilde{h}_{-i\ell j}) \times 100$	8.18		8.08
Fraction Wage Drops or $E(\Delta\overline{w_{i\ell j}} \mid JJ)$	0.40		0.41
Coeff Exper $\times 100$ or Coeff Exper $\times 100$	2.48		2.64
Coeff Exper^2 $\times$ 1000 or Coeff Exper^2 $\times$ 1000	−0.291		−0.278
Coeff Tenure^2 $\times$ 1000 or Coeff Tenure $\times 100$	−0.460		−0.450
Var(Non-employment) or Never Work	16,000	0.077	0.080
Cov($\overline{w}_i$, Non-employment) or Cov(D^n_{ik}, $\overline{w}_i$)	−3.42	−2.97	−2.85
Var(Employment Dur) or Never Job to Job	102,000	0.51	0.54
Invol Job to Job	0.205		0.202

Table IX. Estimation Under Alternative Auxiliary Parameters: Model Fit, Cont'd

	Alternative 2		Alternative 3	
	Alt Aux	Sim	Alt Aux	Sim
Avg. Length Emp. Spell or Avg. Hazard Emp. $\times 1000$		380		380
Avg. Length Non-emp. Spell or Avg. Hazard Non-emp. $\times 1000$		90.7		91.6
Avg. Length Job Spell or Avg. Hazard Job $\times 1000$		108		107
Sample Mean $w_{i\ell jt}$		4.50		4.46
Between Persons $\times 100$		8.10		8.01
Between Jobs $\times 100$		2.86		2.89
Within Job $\times 100$		1.49		1.52
$E(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}) \times 100$ or $E(\Delta\overline{w_{i\ell j}} \mid Q1 \rightarrow Q4)$	0.35	0.37		0.78
$E(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}) \times 100$ or $\frac{E(\Delta\overline{w_{i\ell j}} \mid Q4 \rightarrow Q1)}{E(\Delta\overline{w_{i\ell j}} \mid Q1 \rightarrow Q4)}$	−0.93	−0.93		0.69
$\mathrm{cov}(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$ or $E(\widetilde{S}_{i\ell j}\widetilde{h}_{-i\ell j}) \times 100$	3.58	3.64		8.15
Fraction Wage Drops or $E(\Delta\overline{w_{i\ell j}} \mid JJ)$		0.40	0.037	0.037
Coeff Exper $\times 100$ or Coeff Exper $\times 100$		2.43	1.61	1.65
Coeff Exper^2 $\times$ 1000 or Coeff Exper^2 $\times$ 1000		−0.294	−0.296	−0.287
Coeff Tenure^2 $\times$ 1000 or Coeff Tenure $\times 100$		−0.463	0.409	0.405
Var(Non-employment) or Never Work		16,047		16,161
Cov($\overline{w}_i$, Non-employment) or Cov(D^n_{ik}, $\overline{w_i}$)		−3.39		−3.39
Var(Employment Dur) or Never Job to Job		101,337		103,154
Invol Job to Job		0.210		0.204

- Table X presents the counterfactual decomposition corresponding to panel (A) of Table VI.
- We see that the decompositions are almost unaffected, with only minor differences present compared to the baseline results.
- We also performed the decompositions in other orders with a similar result.

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

9.2. Other Results

- **Estimate Restricted Versions of the Model**: We estimate seven restricted versions of the model, where various model components are taken out. The two most important ones are the fraction of wage drops and the correlation across workers in preferences of jobs $E(\tilde{S}_{i\ell j}\, \tilde{r}_{-i\ell j})$.

- **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.

- **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.

- **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.

10. Conclusions