I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16663

Non-Wage Job Values and Implications for Inequality

Tobias Lehmann

DECEMBER 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16663

Non-Wage Job Values and Implications for Inequality

Tobias Lehmann
Università della Svizzera italiana and IZA

DECEMBER 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Non-Wage Job Values and Implications for Inequality*

I study inequality in job values, both in terms of wages and non-wage values, in Austria over the period 1996 to 2011. I show that differences in non-wage job value between firms are non-parametrically identified from data on worker flows and wage differentials. Intuitively, firms with high non-wage value attract workers without paying a wage premium. I study the distribution of job value among workers and find a positive correlation between wage and non-wage value. Inequality in job value is thus considerably greater than wage inequality, reflected in the standard deviation of job value being more than twice as large as the standard deviation of wage. Job value inequality increases between 1996 and 2011, although wage inequality remains constant. An important reason is that, over time, dispersion of rents offered by firms increases, while compensating differentials lose importance.

JEL Classification: E24, J31, J32

Keywords: inequality, amenities, worker heterogeneity, firm heterogeneity, on-the-job search, wage dispersion, matched employer-employee data

Corresponding author:
Tobias Lehmann
Università della Svizzera italiana
Via Giuseppe Buffi 13
6900 Lugano
Switzerland
E-mail: tobias.lehmann@usi.ch

* I thank Josh Angrist, Jakob Beuschlein, Adrien Bilal, Matthias Doepke, Jonathan Cohen, Mitch Downey, Lea Fricke, Rustamdjan Hakimov, John Horton, Simon Jäger, Gregor Jarosch, Leo Kaas, Philipp Kircher, Nicole Maestas, Andreas Mueller, Arash Nekoei, Raphael Parchet, Laura Pilossoph, Heather Sarsons, Julian Schärrer, David Seim, Isaac Sorkin, David Strömberg, Chris Taber, Pinar Yildirim, and seminar participants at Goethe Frankfurt, ifo Munich, USI Lugano, University of Lausanne, MIT labor lunch, University of Montreal, Stockholm University, Vienna University, as well as conference audience at EALE, IZA summer school, SKILS, SSES, and YSEM for helpful comments. A particular thanks to Rafael Lalive and Camille Terrier for their patience in advising and supporting me, and to Josef Zweimüller and University of Zurich for granting access to data and IT-infrastructure. I gratefully acknowledge funding from the Swiss National Science Foundation grant 187353.

1 Introduction

"The ultimate desideratum is a grand measure of inequality in the returns to work that embodies all monetary and nonpecuniary returns." Hamermesh (1999)

Workers derive utility from their job's wage, and from its non-wage value. Recent experimental evidence shows that workers have high valuation for some non-wage characteristics, for example, schedule flexibility or the opportunity to telecommute (Mas and Pallais, 2017; Maestas et al., 2018; Wiswall and Zafar, 2018). Taber and Vejlin (2020) estimate that only half of the variance of utility workers derive from jobs comes from wage, while the other half is borne by non-wage values. Understanding inequality in workers' well-being thus requires consideration of both, wage and non-wage values of jobs.

While a blossoming literature discusses wage inequality (see Acemoglu and Autor (2011) and Card et al. (2018) for detailed reviews), there is remarkably little empirical evidence on inequality in non-wage values. Maestas et al. (2018), Marinescu et al. (2021), Dube et al. (2022) and Sockin (2022) show that non-wage characteristics tend to be worse in low-wage jobs, therefore exacerbating inequality in job value compared to wage inequality.[1] Hamermesh (1999) and Pierce (2001) show that inequality in fringe benefits and risk of injury grew stronger than wage inequality in the US in the 1980s and 1990s. While these studies document interesting patterns with respect to the *subset* of non-wage characteristics they consider, it is necessary to know the value of *all* non-wage characteristics of jobs for statements about inequality in workers' overall well-being. Few papers have studied how *all* non-wage characteristics affect inequality across comprehensive worker populations, finding that non-wage characteristics can both reduce (Sorkin, 2018a) and exacerbate (Taber and Vejlin, 2020; Lamadon et al., 2022; Berger et al., 2023) labor market inequality.

In this paper I propose a tractable framework for estimating the total non-wage value workers derive in their job. Combining wage and non-wage value allows me to study the evolution of inequality in total job value, and to compare it to the evolution of

[1] Maestas et al. (2018) consider the following job characteristics: set own schedule, telecommute, physical demands, fast paced/relaxed work, independence, 10-20 days paid time off, work in team, training opportunities, positive impact on society. Marinescu et al. (2021) focus on labor rights violations, and Dube et al. (2022) on a set of characteristics related to workplace dignity.

wage inequality. In my framework, workers consider wage and non-wage value when comparing job offers. I identify non-wage value as the residual that explains observed job choices after accounting for wage. My definition of non-wage value thus, by construction, captures the full set of workplace characteristics that contribute to workers' utility.

My analyses are based on Austria, a labor market more comparable to the US than others in Europe, for example, regarding the unemployment rate and labor turnover (Stiglbauer et al., 2003). I use employer-to-employer transitions in Austrian administrative data between 1996 and 2011. Two features of this matched employer-employee data make it attractive for my study. First, it provides daily information on people's employment status, allowing me to follow workers across firms.[2] Second, it provides me with an uncensored measure of earnings, which I can combine with information on whether one is a full-time worker to get a high-quality measure of workers' wage. In order to study developments over time I split the sample into two consecutive 8-year intervals. The 1996–2003 sample covers 800,000 workers at 4,500 firms, and the 2004–2011 sample covers 960,000 workers at 5,900 firms.

I measure voluntary employer-to-employer transitions, which are transitions that do not follow a layoff, or firm-level dynamics such as firm mergers and takeovers.[3] I then describe patterns of worker flows between employers. For example, I find that employers in the *manufacturing* and *public administration/education* industry attract more workers from other employers than they lose workers. I show wage differentials associated with employer-to-employer transitions. While employers in *manufacturing* pay a wage premium, this is not the case for employers in *public administration/education*, where many workers are willing to accept a wage decrease.[4] A possible explanation for this is that employers in *public administration/education* are attractive to workers for non-wage reasons.

I develop a structural interpretation of these reduced form patterns through an on-the-job search model in the vein of Burdett (1978). Workers search for job offers,

[2] I use the terms firm and employer interchangeably.

[3] I observe layoffs if workers apply for unemployment benefits. I account for unobserved layoffs using the procedure by Sorkin (2018a).

[4] This pattern of industry-wage differentials is also found in Krueger and Summers (1988) and Gruetter and Lalive (2009).

which they receive at Poisson rate. Firms' job offers consist of a wage, and a firm-specific non-wage value. The firm-specific non-wage value reflects the median worker's valuation for a firm after accounting for the offered wage. In addition, workers have an idiosyncratic valuation for each firm, reflecting worker-specific work arrangements and preference heterogeneity. When receiving an offer from an outside firm, workers compare it to the offer of their current firm, and transition to the outside firm if it offers them greater value than their current firm.

I show that under this parsimonious model firms' non-wage value offer is non-parametrically identified. Intuitively, for a given wage offered by firm A, the task is to find the wage at firm B at which half of workers choose firm A, and the other half choose firm B. The difference between the wage at firm B and the wage at firm A then reflects the difference in non-wage value (in money units) between the two firms. Providing non-parametric identification of firm non-wage value is important as it lends credibility to my inequality results, which are driven by firm non-wage value.[5]

While I can provide non-parametric identification for the core of my model, estimating it non-parametrically with a reasonable degree of precision is not feasible. To efficiently estimate my model I therefore impose two assumptions: First, the value of a job for a worker is an additive combination of the log-wage, the firm non-wage value, and the worker-firm idiosyncratic value. This assumption implies that workers' valuation for firms' non-wage value is proportional to wage. This is supported by Maestas et al. (2018) finding that workers' willingness to pay for non-wage characteristics is about the same fraction of wage for all quintiles of the wage distribution. Second, I assume that workers' idiosyncratic non-wage value follows a random normal distribution.

Under these assumptions my model gives rise to a simple probit-style likelihood function, where every likelihood contribution represents an employer-to-employer transition.[6] I account for differing firm sizes and the intensity with which firms make job offers to each other's employees by appropriately weighting each likelihood contribution.[7] I allow for heterogeneity between workers in two ways: First, I let the intensity

[5]See French and Taber (2011) for a detailed discussion of non-parametric identification of labor market models, and its importance.

[6]I show that employer-to-employer transitions observed in the data are sufficient for identification, which is necessary because I do not observe when a worker rejects a job offer from an outside firm.

[7]While I directly observe firm size in the data, I follow Bonhomme and Jolivet (2009) and Sorkin

with which workers receive offers from different firms depend on the worker's current firm.[8] Second, I allow for heterogeneity in non-wage value at the individual level through a worker-firm idiosyncratic value component.

I estimate two parameters through my model: The first is each firm's non-wage value.[9] The second parameter identifies the importance of wage, relative to non-wage value, for job value. With this parameter, I can convert non-wage value to a log-wage equivalent scale. This allows me to study inequality in non-wage job value, and to directly compare it to inequality in wage.

I estimate the search model separately for the 1996–2003 period and for the 2004–2011 period. I then combine the search model estimates with the wage data, which allows me to estimate the distribution of job value among all workers. I find a positive correlation between wage and non-wage value for both periods, reflecting sorting of workers with high wages to firms offering high non-wage value. Job value inequality is thus considerably greater than wage inequality, where the standard deviation of job value is a about 2.1 times as large as the standard deviation of wage.

I find that between the 1996–2003 and the 2004–2011 period, job value variance increases by 15 percent. Job value variance can increase for three reasons: wage variance, non-wage value variance, and their covariance. I find that the increase in the covariance between wage and non-wage value is an important driver of the increase in job value variance over time. To understand the sources of this increase, I decompose wage following Abowd et al. (1999) (AKM) into worker quality and firm wage premium. I find that the increase in job value variance is mainly due to a striking change in the covariance between firm wage premium and firm non-wage value. In the 1996–2003 period the covariance between firm wage premium and firm non-wage value is negative, whereas it is positive in 2004–2011.[10]

Economically, the covariance between firm wage premium and firm non-wage value

(2018a) and estimate the intensity with which firms make offers from the number of workers they hire from non-employment. I show that my results do not change when the offer distribution is estimated under alternative assumptions.

[8]Doing so, I allow for sorting of workers across firms.

[9]I actually estimate 4,500 (1996–2003) and 5,900 (2004–2011) parameters here, one for each firm in my sample.

[10]The correlation between firm non-wage value and the firm wage premium in 1996–2003 is close to the correlation Hall and Mueller (2018) find between the non-wage value and the wage of jobs offered to unemployed job seekers.

measures the importance of compensating differentials relative to firm-level rents (Robinson, 1933; Rosen, 1986). Intuitively, if firms fully compensate workers through wage for the quality of their non-wage characteristics, firm wage and non-wage value will be perfectly negatively correlated. If there are no compensating differentials, and dispersion of wage and non-wage value is purely due to firms offering rents, firm wage and non-wage value will be perfectly positively correlated. Obviously, the empirical reality lies somewhere in between. My results show that compensating differentials attenuated job value inequality in the 1996–2003 period. By 2004–2011, however, they have declined and dispersion of firm-level rents has increased, leading to an increase in job value inequality.

This paper contributes to the literature estimating job values in search environments (Bonhomme and Jolivet, 2009; Becker, 2011; Sullivan and To, 2014; Hall and Mueller, 2018; Sorkin, 2018a; Taber and Vejlin, 2020; Jarosch, 2021).[11] Most closely related are Sorkin (2018a) and Taber and Vejlin (2020), who also rely on worker flows between firms to identify total job values. The most important innovation relative to Sorkin is that my model identifies job value in log-wage units. This allows me to compare job value inequality to wage inequality, which is not possible in Sorkin's framework.[12] Another innovation relative to Sorkin is that I incorporate worker heterogeneity in wages, allowing to study sorting of workers with respect to firm wage and non-wage value. Taber and Vejlin (2020) also separate job value into a wage and a non-wage value part. They rely on a rich structural model in which parameters are only indirectly identified by the data.[13] In contrast, my model provides constructive identification based on worker flows and wage differentials. As a result, my model also allows for a direct mapping to the wage model of Abowd et al. (1999), which is not the case in Taber and Vejlin (2020).

[11] An alternative approach is to build an environment without search frictions, and instead follow the industrial organization literature on static differentiated products models by focusing on heterogeneity across workers in their valuation for jobs at different firms (Card et al., 2018; Lamadon et al., 2022; Manning, 2021). Closely related are Roussille and Scuderi (2022) and Azar et al. (2022) who estimate firm non-wage value using job board data.

[12] Sorkin cannot tell whether job value dispersion is greater or smaller than wage dispersion, because in his model, the scale of job value is unidentified.

[13] The richness of the model by Taber and Vejlin (2020) is driven by their ambition to decompose total labor market wage and utility variation into variation due to pre-market skills variation, learning by doing, preferences for non-pecuniary aspects, monopsony, and search frictions.

This paper also contributes to the literature attempting to explain wage inequalities with compensating differentials. Krueger and Summers (1988) find that differences in non-wage characteristics of jobs cannot explain inter-industry wage differentials.[14] Subsequent work has shown that search frictions (Hwang et al., 1998; Bonhomme and Jolivet, 2009) as well as idiosyncratic preferences of workers over firms (Card et al., 2018; Lamadon et al., 2022; Manning, 2021) can explain this result.[15] My model features both search frictions and idiosyncratic preferences of workers over firms. My results confirm that search frictions and idiosyncratic preferences can lead to rent dispersion among firms nullifying the inequality attenuating effect of compensating differentials.

The rest of the paper is organized as follows. The next section describes the data and provides descriptive evidence on patterns of employer-to-employer transition. Section 3 discusses identification of non-wage values. Section 4 presents the results. Robustness is considered in Section 5, and Section 6 concludes.

2 Background, Data and Descriptive Evidence

Background The Austrian labor market combines broad institutional regulation with high flexibility. Virtually all jobs are covered by collective bargaining agreements setting wage floors and minimum non-wage work arrangements (Glassner and Hofmann, 2019).[16] For most jobs, however, provisions from collective bargaining agreements are not binding. For example, Leoni et al. (2011) find that actual wages in manufacturing in the early 2000s were on average 20-30 percent higher than collective bargaining wage floors. The Austrian labor market thus maintains a high degree of flexibility. Job creation and job destruction rates in most industries are comparable to those in the US (Stiglbauer et al., 2003). Between 1996 and 2011, the Austrian labor market was characterized by relatively steady conditions. Unemployment was among the lowest

[14] Similarly, Katz et al. (1989) find a slight positive correlation between the industry wage premium and the quality of non-wage characteristics.

[15] An earlier literature emphasizes the role of unobserved worker heterogeneity (Hwang et al., 1992; Brown, 1980), which is of second order in studies relying on panel data and within-individual variation.

[16] Non-wage characteristics are, for example, dismissal protection or paid further training (Glassner and Hofmann, 2019).

in Europe, ranging from 3.5 percent in 2000 to 6.5 percent after the great recession in 2009. The wage structure was stable between 1996 and 2011.[17]

Data I use data from two administrative sources, which together allow me to follow workers across firms and observe their wages. The Austrian social security data (Zweimüller et al., 2009) provide matched employer-employee data on the universe of Austrian private sector employment and public sector employment under private labor law.[18] The social security data contain detailed daily information on worker labor market status (e.g., employed, unemployed, retired). Each employment spell is linked to a firm identifier and information on the firm's industry and location.[19]

The second data source is the Austrian wage tax data (Büchi, 2008). They cover the universe of private and public sector employment. The wage tax data are based on wage tax forms annually submitted by firms. They contain workers' uncensored gross labor earnings,[20] and since the year 2002 an indicator whether an individual is working full-time or part-time. Before 2002, more than 97 percent of working men were full-time employed.[21] When limiting attention to men and excluding part-time workers after 2002, gross earnings from wage tax data represent a high-quality measure of wage, as large variation in working hours is ruled out.[22]

Matched employer-employee panel I construct two consecutive 8-year panels of the Austrian workforce, from 1996 to 2003 and from 2004 to 2011, by combining employment information from the social security data with wage information from the wage tax data.[23] Individuals in my panel satisfy the following three conditions: (1) The per-

[17] My model does not assume a steady state, but allows firms to grow and shrink over time. Business cycles affect my results only through the change in the composition of jobs (w.r.t. wages and non-wage values) they induce. This is, however, exactly the main outcome captured by my model, and not a confounder.

[18] In 2004 34 percent of public sector employees were employed with private sector contracts and therewith part of the social security data (Bundeskanzleramt, 2021).

[19] Most establishments of multi-establishment firms in Austria have a common firm-identifier in the social security data (Fink et al., 2010).

[20] Including bonus payments.

[21] Only around 70 percent of women were employed full-time between 1996 and 2002.

[22] Industry-level averages of weekly working hours range from 39.8 hours in utilities to 44.4 hours in hotel and restaurant (own calculations based on the Austrian Microcensus).

[23] To the best of my knowledge, this is the first study on Austria to rely on wage information from Austrian wage tax data, while all previous studies on Austria have estimated earnings from the social security data (e.g., Card et al., 2007; Lalive and Zweimüller, 2009; Nekoei and Weber, 2017).

son is male and not a part-time worker, (2) he is working for the entire calendar year, and (3) holds only one single job. Condition (1) allows me to interpret earnings as wages. Condition (2) and (3) ensure that I can link a person-year observation in the social security data to the wage tax data. Apart from being required by the data, these conditions are also motivated by my framework. I interpret employer-to-employer transitions as the result of a worker's binary choice over two jobs. This is only suitable for workers holding one single job at a time. The condition that workers must work for the same employer for at least one entire calendar year excludes workers in seasonal employment, where the termination of an employment spell in most cases is caused by the end of the employer's business season, rather than following a worker's choice.

The model I will introduce in Section 3 is only identified for employers strongly connected by employer-to-employer transitions.[24] The restriction concerns the network of worker flows between employers. An employer is in a strongly connected set if it hires at least one worker from another employer in this strongly connected set, and has at least one of its workers hired by another employer in this strongly connected set.[25] To make sure my model estimator converges within reasonable time, I limit my sample to employers that have at least 5 employer-to-employer transitions with other employers in the strongly connected set.[26]

Table 1 shows descriptive statistics on the 1996 to 2003 and the 2004 to 2011 employment panel. Columns 1 and 3 show statistics on all employers, while columns 2 and 4 consider the sample of strongly connected employers. Panel A. shows that while there are much fewer employers in the strongly connected sample, columns 2 and 4 still cover more than half of the labor market when measured through the number of people-year observations. This reflects that the strongly connectedness condition is much more likely to be satisfied by medium-sized and large employers. Panel B. shows that while workers in my sample earn higher wages on average, wage dispersion is about the same in my sample as in the Austrian labor market overall.

[24]Technically, the strongly contentedness condition follows from the maximum likelihood estimator regularity condition that the identified parameter vector needs to be an interior point (see Section 3.3).

[25]In my sample I consider the largest strongly connected set, that is, the set containing most employers.

[26]My results are not sensitive to this restrictions. However, convergence of my estimator is very slow if one allows for firms with a lower number of transitions. I implement this restriction in a loop, where I sequentially drop firms with fewer than 5 employer-to-employer transitions with the strongly connected set, until every firm has at least 5 employer-to-employer transitions with the strongly connected set.

Table 1: POPULATION AND STRONGLY CONNECTED SAMPLE 1996–2003 & 2004–2011

	1996 – 2003		2004–2011	
	All	**Strongly connected**	**All**	**Strongly connected**
	(1)	(2)	(3)	(4)
A. Sample size				
People-years	9,526,421	4,513,833	9,906,446	5,480,901
People	1,621,545	797,492	1,712,585	964,635
Employers	193,633	4,544	182,811	5,944
B. Summary statistics				
Mean age	38.80	39.07	40.21	40.21
Share blue collar	0.48	0.43	0.43	0.39
Median monthly wage (2012 €)	3,048	3,345	3,196	3,481
Mean log monthly wage	8.09	8.19	8.14	8.23
Var log monthly wage	0.21	0.20	0.21	0.20
C. Industry shares				
Manufacturing	0.31	0.39	0.31	0.39
Utilities	0.02	0.03	0.02	0.03
Construction	0.10	0.05	0.10	0.06
Retail trade, cars	0.16	0.10	0.15	0.10
Transportation	0.07	0.07	0.07	0.08
Hotel and restaurant	0.02	0.00	0.02	0.00
Information and communication	0.02	0.02	0.03	0.03
Finance and insurance	0.06	0.08	0.05	0.06
Real estate	0.02	0.02	0.02	0.02
Prof./scientific/tech. services	0.05	0.05	0.04	0.03
Services	0.04	0.04	0.05	0.05
Public admin./education	0.10	0.13	0.10	0.13
Health and social	0.02	0.02	0.02	0.02
D. Employer-to-employer transitions				
Transitions	159,199	58,349	178,835	74,271
Share excess separations	0.49	0.54	0.48	0.47
Mean log wage increase	0.11	0.11	0.09	0.10
Mean log wage increase (adjusted)†	0.05	0.06	0.05	0.05
Share wage increase (adjusted wage)	0.58	0.60	0.59	0.60
Share both employers same industry	0.44	0.47	0.43	0.45

Notes: This table reports summary statistics on all male full time workers (columns 1 and 3) and those in the sample of strongly connected firms (columns 2 and 4). The industry classification is based on NACE Rev. 2 main sections. I combine section D & E (Utilities), O & P (Public admin./education) and N & S (Services). The following industries are not shown: Agriculture, forestry and fishing, Mining, Arts and entertainment, Households as employers, (All share people-years in 1996–2003 <0.01). All summary statistics on transitions (Panel D. after *Share excess separations*) are with observations weighted by their probability of being an excess separation as defined in the text.

† The wage at the old employer is observed in year t, and the wage at the new employer in year $t+2$. I substract time and experience effects from the wage at the new employer using the estimates from my AKM-regression (see Online Appendix G.2)

Concerns related to external validity may also arise because my sample restricts attention to male workers and to full-time workers. For example, one might be concerned that women and part-time workers are differently sorted across firms, and that they differ in their preferences over non-wage characteristics offered by firms. I address these concerns in Online Appendix J. I show that I obtain similar results when including women in the 2004–2011 panel, and that my sample well reflects the overall structure and dynamics of the Austrian labor market.

Employer-to-employer transitions A change of employer is classified as an employer-to-employer transition if there are at most 30 days of non-employment between two consecutive employment spells. Second, the worker must have been working for the old employer since the start of the calendar year preceding the transition, and he must work for the new employer until the end of the calendar year succeeding the transition.[27]

My model is built around the idea that employer-to-employer transitions are the outcome of a worker's choice between a job offer from his old employer and a job offer from his new employer. I therefore exclude all transitions that most likely are not the result of such a worker decision. Specifically, I exclude all transitions that follow a layoff recorded in the social security data.[28] I also exclude all transitions that follow firm-level dynamics such as firm renamings, takeovers, mergers, spin-offs, or firm closures.[29]

Even after removing these transitions, there are involuntary employer-to-employer transitions left in my sample. In particular, my data do not allow me to identify cases where a worker is laid off and finds a new job without an interrupting unemployment spell. Sorkin (2018a) proposes a probabilistic approach to correct for these transitions. The underlying idea is that these transitions are most likely to happen at contracting firms. I calculate the average employer-to-employer separation rate at expanding firms,

[27]The year of the transition is the year of the last day of employment at the old employer.

[28]Laid-off workers are eligible for unemployment benefits from the first day of unemployment. Workers who quit face a waiting period of 4 weeks. This implies that I can identify laid-off workers from the social security data to the extent that the lay-off leads to receiving unemployment benefit.

[29]Following (Fink et al., 2010) I identify firm-level dynamics from collective actions of groups of workers, as recorded in the social security data. For example, a firm takeover is identified if a firm-identifier disappears from the records and if at least two thirds of workers work for the same firm in the following quarter. See Online Appendix E for details.

which I use as an estimate for the expected separation rate from voluntary employer-to-employer transitions. When a firm is contracting and the separation rate is in excess of the expected rate, I consider these separations as exogenous due to an employer-level shock. I calculate the expected rates by industry, and then downweight separations at contracting firms with $(1 - \frac{\text{excess}}{\text{excess+expected}})$.[30,31]

Panel D. of Table 1 shows descriptive statistics on employer-to-employer transitions. 58,349 transitions occur between firms in my sample for 1996–2003 and 74,271 transitions between 2004–2011. In both periods, employer-to-employer transitions come on average with a log wage increase of about 0.05, and wage increases for around 60 percent of transitions. Table A.2 shows in detail how I obtain the transitions in Table 1 from all employment spells that end in the two sample periods.

Descriptive evidence on transitions, wage differentials, and non-wage values I will now discuss descriptive evidence on employer-to-employer transitions and wage differentials between firms, and illustrate how we can use them to learn about firms' non-wage values. I will use evidence aggregated on the industry-level for the 2004–2011 panel, noting that patterns look similar in the 1996–2003 panel.

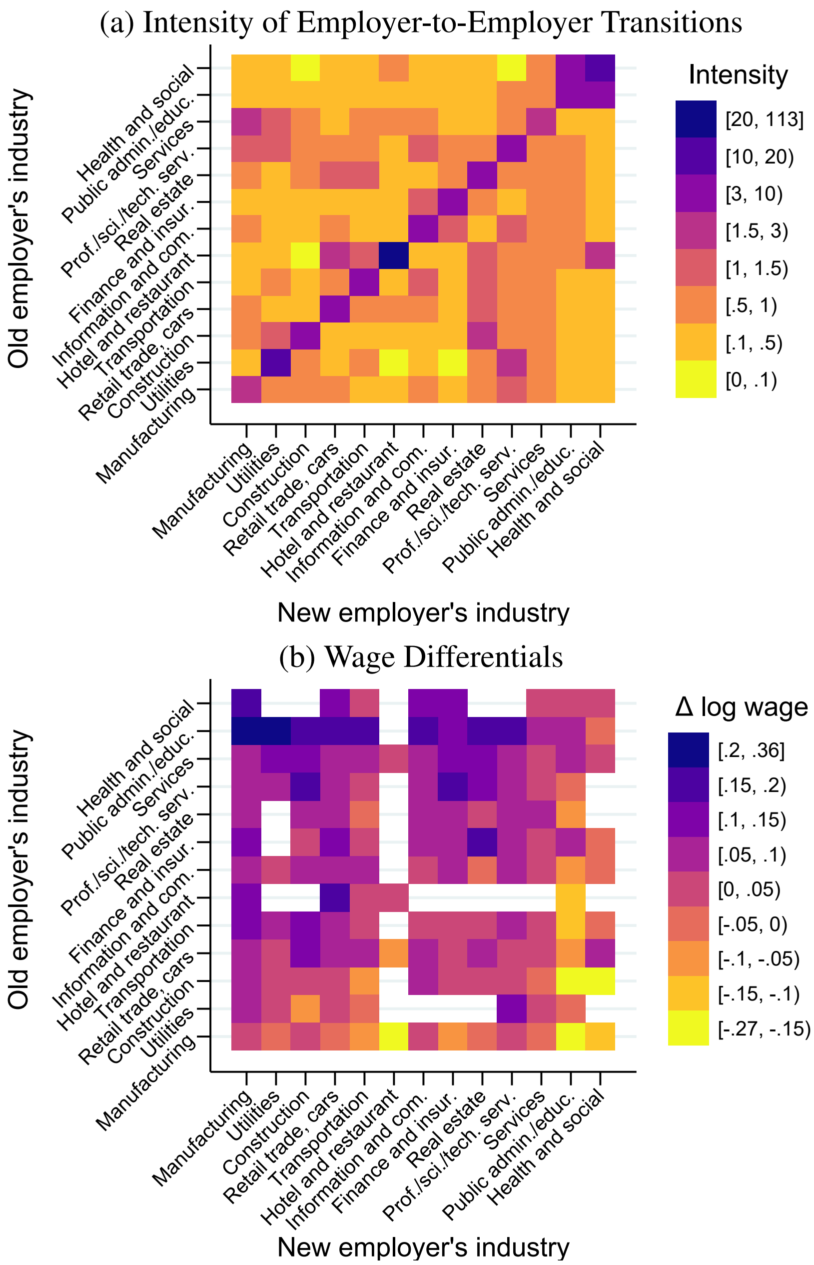
Figure 1a shows how workers transition between industries. Each cell measures the intensity of employer-to-employer transitions from an industry in the corresponding row to an industry in the corresponding column. The intensity measures how many employer-to-employer transitions actually happen from a row-industry to a column-industry, relative to how many would be expected to happen if mobility was random with respect to industries.[32] Thus the greater the value of a cell the more intensively workers transition from the corresponding row-industry to the corresponding column-industry. Values above 1 represent intensities above the random mobility counterfactual,

[30] Annual separation rates at expanding firms are highest in Services (3 percent) and lowest in Public administration/education and Utilites (1 percent). See Table A.3 for separation rates by industry.

[31] This approach corrects the ratio of firm-to-firm inflows and firm-to-firm outflows at contracting firm-years, but not the wage differentials associated with involuntary firm-to-firm transitions. I therefore repeat my analyses excluding all separations at contracting firms. The results are qualitatively identical and quantitatively similar.

[32] Each cell corresponding to row-industry j and column-industry k equals $\frac{\text{transitions}_{jk}}{\sum_{s\in J}\sum_{l\in J}\text{transitions}_{sl}} * (\frac{\sum_{l\in J}\text{transitions}_{jl}}{\sum_{s\in J}\sum_{l\in J}\text{transitions}_{sl}} * \frac{\sum_{s\in J}\text{transitions}_{sk}}{\sum_{s\in J}\sum_{l\in J}\text{transitions}_{sl}})^{-1}$, where transitions_{jk} denotes the number of employer-to-employer transitions between industry j and industry k, and J the set of all industries.

Figure 1: EMPLOYER-TO-EMPLOYER TRANSITIONS AND WAGE DIFFERENTIALS



Notes: Figure a shows the intensity of employer-to-employer transitions between industries over the period 2004–2011. If mobility was random, the intensity would be equal to 1 for each cell. Intensities above 1 indicates that there are more transitions from the row-industry to the column-industry than expected under random mobility. See text for a formal definition of the intensity. Figure b shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions with the old firm in the row-industry in the new firm in the column-industry. Missing cells in figure b contain fewer than 10 observations. Both figures are based on transitions between firms in the strongly connected 2004–2011 sample (column 4 in Table 1). See Figure A.2 for employer-to-employer transitions of all workers over the period 2004–2011.

and values below 1 intensities below. The large variation in intensities depicted in Figure 1a shows that mobility between industries is clearly non-random. Unsurprisingly, the intensities are largest along the diagonal, reflecting that most employer-to-employer transitions happen within the same industry. There are also systematic patterns between some industries, for example between *public administration/education* and *health and social services*, reflected by high intensities in the top-right corner cells in Figure 1a.

Table A.4 summarizes, by industry, the number of workers employers attract from other employers, and compares it to the number of workers they loose to other employers. Two industries, *manufacturing* and *public administration/education*, stand out because they attract around 20 percent more workers from other employers than they lose to other employers. This suggests that working in *manufacturing* and *public administration/education* is relatively attractive for workers, that is, workers are willing to give up their old job to join an employer in these two industries, but not as willing to give up their job in these two industries to work elsewhere.[33]

Figure 1b provides evidence on the extent to which *manufacturing* and *public administration/education* employers' attractiveness can be explained by wage premia. It shows the average wage increase that comes with a transition from an employer in the row-industry to an employer in the column-industry. We see rather dark colors in the column *manufacturing*, reflecting that workers who join *manufacturing* typically see their wage increase. On average, workers who join *manufacturing* see their wage increase by 6.9 percent.[34] In contrast, workers who leave manufacturing on average see their wage increase by only 0.5 percent, reflected by rather bright colors in the *manufacturing* row. The exact opposite picture arises for *public administration/education*. Workers who join *public administration/education* on average see their wage decline by 2 percent, while workers who leave it see their wage increase by 8.3 percent on average.

Overall, industry-level descriptive statistics suggest that while employers in *manufacturing* and *public administration/education* are attractive for workers, it is only in the case of manufacturing that this can at least in part be explained by manufacturing employers paying a wage premia. In *public administration/education*, however, there must be something other than the wage making it attractive for workers. This is exactly

[33]On the other hand, employers in *construction*, *real estate*, and *services* lose more workers to other employers than they hire from them. This suggests that employers in these industries are rather unattractive for workers. The *services* industry includes mostly industries providing low-skilled services (NACE Rev. 2 codes N & S).

[34]Table A.5 shows average wage differentials for employer-to-employer transitions by industry.

the intuition behind the identification of non-wage values in my model, which I will explain in the following section.

3 The Model

I will now construct an on the job search model in the vein of Burdett (1978).[35] The model is partial equilibrium, meaning that I take firm behavior as exogenously given. Firms post contracts that workers either accept or not, so there is no bargaining. The model incorporates search frictions in the form of a stochastic rate at which workers receive job offers. For ease of exposition I assume random search, while showing in Online Appendix F that identification results also hold in a directed search version of the model. I use the notation that capital letters denote random variables, and small letters denote realizations of random variables and parameter values.

3.1 Primitives

Firms Each firm $j \in J$ is fully characterized by the tuple $\langle \tilde{\psi}_j, a_j, g_j, f_j, \delta_j, \rho_j \rangle$. $\tilde{\psi}_j$ denotes the wage premium firm j offers and a_j denotes its non-wage value. One can think of $a_j = a(\mathbf{m}_j)$, where $\mathbf{m}_j$ is an arbitrary-dimensional vector containing characteristics other than the present wage that are valuable to a worker when working at firm j, and that are converted to a non-wage value through the function $a()$.[36] g_j denotes the size of the firm, that is, the number of employees of firm j. f_j denotes the share of all job offers made that are from firm j. δ_j denotes the job destruction rate, and ρ_j the rate at which, after job destruction, workers directly find a new firm (as opposed to being sent to non-employment).

Job offers A firm's job offer to worker i consists of a wage w_{ij} and the firm's non-wage value a_j. I assume firms' wage offer can be written as $ln(w_{ij}) = \tilde{\alpha}_i + \tilde{\psi}_j + \eta_{ij}$,

[35] This is essentially a partial-equilibrium version of the well-known Burdett and Mortensen (1998) model.

[36] Besides amenities that provide flow-utility to the worker, a_j also contains expectations about job security (through δ_j and ρ_j) and, depending on the definition of W_j, future wage growth. This is intentional as the aim of this article is to provide a comprehensive measure of compensation in the labor market that includes all aspects other than the contemporaneous wage. I discuss in Section 5 how these assumptions affect my results.

where $\tilde{\alpha}_i$ captures workers' productivity type and η_{ij} reflects a random wage component. Note that while the assumption on the structure of wage offers is not needed when providing identification of the model below, it is useful later to understand results on inequality and its evolution over time.

Workers Workers value a job offer as follows:

$$V_{ij} = V(w_{ij}, a_j) + \epsilon_{ij},$$

where ϵ_{ij} captures both idiosyncratic variation in non-wage value offered by firm j as well as idiosyncratic variation in worker i's preference for firm j. A worker employed at firm j has the following value function:[37]

$$\begin{aligned}
\underbrace{V(w_j, a_j)}_{\text{value of being at } j} &= \underbrace{v(w_j, a_j)}_{\text{flow payoff}} \\
+ \underbrace{\beta}_{\text{discounter}} \Bigg[& \underbrace{\delta_j(1-\rho_j)V^n}_{\substack{\text{exogenous employer-}\\\text{to-non-employment}}} + \underbrace{\delta_j \rho_j \sum_{k \in J} f_k \int_{W_k} \int_{E_k} (V(w_k, a_k) + \epsilon_k) dF(w_k) dF(\epsilon_k)}_{\text{exogenous employer-to-employer}} + \underbrace{(1-\delta_j)}_{\text{no exogenous transition}} \\
* \Bigg(& \underbrace{\lambda_E \sum_{k \in J} f_k \int_{W_k} \int_{W_j} \int_{E_k} \int_{E_j} \max\{V(w_k, a_k) + \epsilon_k, V(w_j, a_j) + \epsilon_j\} dF(w_k) dF(w_j) dF(\epsilon_k) dF(\epsilon_j)}_{\text{receive job offer and make binary choice}} \\
& + \underbrace{(1-\lambda_E)}_{\text{no job offer}} \int_{W_j} \int_{E_j} (V(w_j, a_j) + \epsilon_j) dF(w_j) dF(\epsilon_j) \Bigg) \Bigg],
\end{aligned} \tag{1}$$

meaning that the value of a worker employed at j consists of his flow payoff v and a continuation value which he discounts by β. The continuation value represents the combination of four mutually exclusive cases. With probability δ_j the worker's job is destructed. In that case he is sent to non-employment with probability $(1 - \rho_j)$, while he immediately finds a new job with probability ρ_j.

The case from which I draw the identifying variation in my framework is the case when the worker's job is not destructed and he receives a job offer from another firm, which happens with probability $(1-\delta_j)\lambda_E$. The probability the offer is from a particular firm k is f_k. When receiving a job offer from firm k the worker draws an idiosyncratic

[37] Following Arcidiacono and Ellickson (2011, p. 368) I write the value function as the value of being at firm k just before the first idiosyncratic draw η is revealed, which is why the idiosyncratic draw does not show up in the flow utility.

taste shock for the outside firm k and the current firm j, and makes a binary choice: If the outside firm offers him greater value, he switches to the outside firm, otherwise he stays with his current firm.[38]

The last line of Equation 1 describes the case when the worker does not receive a job offer from an outside firm, in which case he draws a new wage and idiosyncratic taste shock and stays with the current firm.

Non-employed workers' value is characterized by the following Bellman equation:

$$\underbrace{V^n}_{\text{value of non-empl.}} = b + \beta\Bigg(\underbrace{\lambda_{NE}\sum_{j\in J} f_j \int_{W_j}\int_{E_j}(V(w_j,a_j)+\epsilon_j)dF(w_j)dF(\epsilon_j)}_{\text{receive job offer}} + \underbrace{(1-\lambda_{NE})V^n}_{\text{no offer}}\Bigg), \tag{2}$$

The non-employed worker receives a flow payoff of b, and a continuation value. The continuation value consists of two parts: first, with probability λ_{NE}, the non-employed worker receives a job offer, which comes with probability f_j from firm j. I assume that the worker always prefers employment over non-employment, and thus accepts any job offer he receives. This will be useful later as it allows me to identify the offer probabilities f_j from observing at which firms non-employed workers get hired. The second part of the continuation value represents the case where the non-employed worker does not receive a job offer.

3.2 Identification

In the following I will establish which features of the model above can be non-parametrically identified. For the others I will show under which parametric assumptions they are identified. This is important because if identification relies on parametric assumptions, outcomes may depend on these assumptions rather than the distribution of the data (Heckman and Honore, 1990; French and Taber, 2011). I will discuss identification for the pair of firms j and k, noting that the results hold for any pair of firms $j \in J$ and $k \neq j \in J$, and thus for all firms in the sample.

I make the following assumptions on the data observed:

Assumption 1. *The analyst*

[38] The idiosyncratic taste shock captures both idiosyncratic variation in non-wage work arrangements at a firm and the worker's idiosyncratic taste for the non-wage work arrangements at a firm.

(a) observes all employer-to-employer transitions that are not the result of an exogenous separation.

(b) observes the wage offered by the new firm and the old firm for all employer-to-employer transitions.

(c) does not observe job offers that do not result in an employer-to-employer transition.

(d) observes all hires out of non-employment.

(e) observes the number of employees at a firm.

Assumption 1(a) implies that the econometrician sees when a worker makes a voluntary employer-to-employer transition, i.e., a transition where the worker chooses between a job offer from an outside firm and a job offer from the current firm. Assumption 1(b) assumes that the econometrician also observes the wages of job offers.

The idea behind Assumption 1(c) is to bring the data requirement for identification in line with what is typically observed in administrative data. Assumption 1(d), together with the model assumption that non-employed workers accept all job offers, allow me to identify the intensity with which firms make job offers to employed workers.

Next I will impose some assumptions on offered wages and the idiosyncratic taste shocks.

Assumption 2. *The idiosyncratic taste shocks $\epsilon_{ij}, \epsilon_{ik}$ are continuously distributed with distribution function G, support $\mathbb{R}^2$, and independent of W_j, W_k. The marginal distributions of ϵ_{ij}, ϵ_{ik} and $\epsilon_{ij} - \epsilon_{ik}$ have medians equal to zero.*

Assumption 3. *The support of W_j and W_k includes at least one point where $V(w_j, a_j) = V(w_k, a_k)$.*

Assumption 4. $\frac{\partial V(w,a)}{\partial w} > 0$ *and* $\frac{\partial V(w,a)}{\partial a} > 0$.

Assumption 2 assumes that the nuisance parameters are independent of the explanatory variables, and normalizes their distributions to have median zero. These assumptions are standard in non-parametric identification of choice models (Matzkin, 1992;

French and Taber, 2011). Assumption 3 implies that there exists at least one combination of wage offers of firms j and k where the worker is indifferent between the two firms.

Theorem 1. *Under Assumptions 1–4 and with data generated by the model in Section 3.1, for any w_j, w_k for which $V(w_j, a_j) = V(w_k, a_k)$ the compensating variation for the difference between a_j and a_k equals $w_k - w_j$, and is thus identified in monetary units from observing employer-to-employer transitions between firms j and k.*

Proof. Appendix A.1 □

Theorem 1 says that we can identify the compensating differential for working at firm j as opposed to working at firm k from observing employer-to-employer transitions for which $V(w_j, a_j) = V(w_k, a_k)$. To see the intuition behind Theorem 1, consider the case where firms j and k have equally many employees and make equally many job offers. If we then observe that equally many workers who are offered the wages (w_j, w_k) switch from firm j to firm k as workers offered wages (w_j, w_k) switch from firm k to firm j, then the additional utility derived from the difference in non-wage value between firms j and k must be negative one times the additional utility derived from the difference in wage between j and k.

Theorem 1 uses a simple intuition to account for heterogeneity between firms with respect to the number of employees and the intensity with which they make offers, which are both observed by Assumption 1(d)–(e). To see this, consider the case where firm j and k are of equal size but firm j makes twice as many job offers as firm k. In this case we will observe twice as many workers moving from firm k to firm j at the point $V(w_j, a_j) = V(w_k, a_k)$. Likewise, if firm j and firm k make equally many offers but firm k has twice as many employees as firm j, we will observe twice as many workers moving from firm k to firm j at the point $V(w_j, a_j) = V(w_k, a_k)$.

Assumption 5. *The function $V(w, a)$ assumes the form $V = \gamma ln(w) + ln(a)$*

Assumption 5 assumes that workers' utility from wage and non-wage value is a linear combination of log-wage and non-wage value.[39] This log-additive functional

[39]This functional form assumption can be relaxed with regard to Theorem 2 if one is willing to impose stronger assumptions on the support of W.

form is supported by Maestas et al. (2018, Figure 7) and Mas and Pallais (2017, p. 3754) finding that individuals with high vs. low wage are willing to give up the same fraction of wage for various amenities. The scale parameter γ converts log-wage to utility.

Assumption 6. *$supp(W_s) = \mathbb{R}_{>0}$ holds for at least one $s \in j, k$.*

Theorem 2. *Under Assumptions 1–6 and with data generated by the model in Section 3.1, the distribution of $\epsilon_{ij} - \epsilon_{ik}$ is identified.*

Proof. Appendix A.2 □

First note that γ can be normalized to any strictly positive real value, according to the desired scale of job value.[40] Because we know firm non-wage value by Theorem 1 and the substitutabiliy of wage and non-wage value by Assumption 5, we can identify the distribution of $\epsilon_{ij} - \epsilon_{ik}$ from observing the changes in probability of choosing firm j when varying W_j or W_k.[41] As the support of W_j or W_k is the positive real line, the support of $\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k))$ is the full real line, allowing for identification of the density of $\epsilon_{ij} - \epsilon_{ik}$ for any value on its support.

3.3 Estimation

I have shown that non-wage value differences between firms are non-parametrically identified, and that assuming Cobb-Douglas utility allows for identification of the distribution of $\epsilon_{ij} - \epsilon_{ik}$.[42] To estimate my model efficiently and with a reasonable degree of precision, I make the following assumption on the distribution of idiosyncratic preferences:

Assumption 7. *$\epsilon_{is} \sim i.i.d.\ N(0, \frac{1}{2})\ \forall\ s \in J$.*

Theorem 3. *Let $\Omega = ([j, k, \Delta ln(w)]_1, ..., [j, k, \Delta ln(w)]_S)$ be the set of all S employer-to-employer transitions between all firms $j, k \in J$ generated under the model above, let*

[40] An intuitive normalization is to set $\gamma = 1$ so that non-wage value is in log-wage units.

[41] Here, W_j and W_k act as exclusion restrictions as discussed in the context of the Roy model in Heckman and Honore (1990) and French and Taber (2011).

[42] Although firm non-wage job value is non-parametrically identified, estimating it non-parametrically is not feasible with a reasonable degree of precision.

Φ denote the standard normal cumulative distribution function, and suppose Assumptions 1–7 hold. The joint likelihood of all S transitions is:

$$\mathcal{L} = \prod_{s=1}^{S} \Phi[\gamma(ln(w_{ij}) - ln(w_{ik})) + ln(a_j) - ln(a_k)]^{\frac{1}{f_j}\frac{1}{g_k}},$$

and one can consistently estimate γ and $ln(a)$ with Maximum Likelihood.

Proof. Appendix A.3 □

Theorem 3 states that the likelihood the above model results in the set S of employer-to-employer transitions is simply the product of the likelihood contributions of the transitions, each of them appropriately weighted. To see the intuition behind Theorem 3, it is instructive to consider the case when all firms make equally many offers, so f_j is constant, and all firms are of equal size, so g_j is constant. In this case, Theorem 3 states that the likelihood of observing the S transitions is simply the product of the likelihood contributions of these S transitions. This holds true because for every pair of firms j and k the number of workers at firm j that receives a job offer from firm k, but rejects the offer, is equivalent to the number of workers at firm k that receives a job offer from firm j and accepts, and vice versa.

Starting from this, we can see the intuition for the likelihood-weight $\frac{1}{f_j}$, which is the inverse of the offer intensity of the firm the worker *joins*. Suppose firms j and k are otherwise exactly the same, but that firm j makes twice as many job offers as firm k. Consequently, we will observe twice as many employer-to-employer transitions from firm k to firm j as from firm j to firm k. This is, however, not because firm k offers any better non-wage value than firm j, but only because it recruits more intensively. By downweighting the likelihood contribution of every employer-to-employer transition from firm k to firm j by one half, the estimator accounts for the difference in offer intensity between these firms.

A very similar intuition applies for the likelihood weight $\frac{1}{g_k}$, which is the inverse of the number of employees at the firm the worker *leaves*. Consider two firms that are exactly the same, except that one has twice as many employees as the other. In this case, we will observe twice as many employees leaving the larger firm than the smaller. By downweighting the likelihood contribution of every employer-to-employer transition away from the larger firm by one half, the estimator accounts for the difference in size between these firms.

Incidental parameter bias The estimate of γ estimated through Theorem 3 will suffer an incidental parameter bias as the estimator is nonlinear, and the number of observations used to identify each firm non-wage value is limited (Greene, 2015, pp. 188–192). I use the jackknife correction in Hahn and Newey (2004) to correct for this bias. This correction builds on the insight by Neyman and Scott (1948) that the incidental parameter bias is of order $\frac{1}{T}$, where T reflects the number of observations available to identify each of the fixed effects. In my framework $T = \frac{2n(S)}{n(J)}$, where $n(S)$ denotes the number of employer-to-employer transitions and $n(J)$ denotes the number of firms. The approach then lies in discarding $\frac{n(J)}{2}$ different transitions in each jackknife replication, resulting in T estimates γ^{T-1} with bias O($\frac{1}{T-1}$). Then the difference between T times γ^T from the baseline estimate (with bias O($\frac{1}{T}$)) minus $T-1$ times the average of all jackknife estimates $\bar{\gamma}^{T-1}$ (with bias O($\frac{1}{T-1}$)) results in the jackknife-corrected γ_{JK} with bias reduced to O($\frac{1}{T^2}$). In Appendix B I provide more detail including Monte Carlo simulations showing that the jackknife correction reduces incidental parameter bias in the estimate of γ from 15 percent to 5 percent.

Non-wage job value in log-wage units Using Theorem 3 I can estimate $ln(a)$ in utility units, and I can estimate γ_{JK}, which converts utility to log-wage.[43] Thus I can estimate firm non-wage value in log-wage units as $ln(a^{€}) = \frac{ln(a)}{\gamma_{JK}}$, and the variance of idiosyncratic utility in log-wage units as $Var(\epsilon^{€}) = \frac{Var(\epsilon)}{\gamma_{JK}^2} = \frac{0.5}{\gamma_{JK}^2}$.

Accounting for estimation error Firms' non-wage values estimated through the likelihood function in Theorem 3 will be a combination of true non-wage value and some estimation error. Therefore, the variance of estimated firm non-wage values across firms will be a combination of the true variance and the variance of the estimation error.

[43] I use Stata's *ml* command and Newton-Raphson. I account for the probability a transition in my sample does not represent a worker-initiated employer-to-employer transition (see Section 2) by weighting each likelihood contribution by $(1-\rho_{kt})$, where ρ_{kt} represents the share of employer-to-employer transitions at firm k in year t that are excess separations.

I estimate the variance of each firms' non-wage value estimate π_j^2 by bootstrapping the likelihood function in Theorem 3 50 times (see Online Appendix D for details). The variance of firm non-wage value excluding estimation error variance is then given by the following formula (Sorkin, 2018b, Appendix H):

$$Var(ln(a^{€})) = \frac{\sum_J w_j \{ (\frac{n_J}{n_J - 1}) [ln(\hat{a}_j^{€}) - ln(\bar{a}^{€})]^2 - \hat{\pi}_j^2 \}}{\sum_J w_j}, \tag{3}$$

where $ln(\hat{a}_j^{€})$ is the estimate of firm j's non-wage value, $ln(\bar{a}^{€}) = \frac{\sum_J w_j ln(\hat{a}_j^{€})}{\sum_J w_j}$, and w_j is the number of person-year observations at firm j.

Table 2: SEARCH MODEL ESTIMATES

	Mean	Var	Min	Max
	1996 – 2003			
Parameters				
Firm size (person-years)	799	$2,136^2$	2	56,744
Hires from non-employment	60	134^2	1	4,918
Estimates				
$ln(a^{€})$	0	.43	-4.29	2.97
$Var(\epsilon^{€})$	.10			
Number of transitions	58,349			
	2004 – 2011			
Parameters				
Firm size (person-years)	727	$1,830^2$	5	53,569
Hires from non-employment	59	132^2	1	4,140
Estimates				
$ln(a^{€})$	0	.49	-4.39	4.39
$Var(\epsilon^{€})$	.12			
Number of transitions	74,271			

Notes: This table shows model parameters and estimates from applying the estimator in Theorem 2 on the samples in Table 1, column (2) and (4). The firm size is measured 1995 to 2002 (2003 to 2010). Hires from non-employment are measured 1996 to 2003 (2004 to 2011). The variance of firms' non-wage value reported is after subtracting variance from estimation error using equation 3. Firms' non-wage values are only identified relative to each other and thus standardized to have mean zero.

Search model estimates Table 2 shows the distribution of the two model parameters I use in estimation, separately for the 1996–2003 and the 2004–2011 panel. A firm's size g is reflected by its number of people-year observations. The hires from non-employment correspond to the total number of individuals hired that were non-employed for at least 30 days. Both measures are totals over 8 years.[44]

Firms' non-wage values $ln(a^{€})$ in Theorem 3 are identified relative to a base firm's non-wage value, which I select to be the firm with the most employer-to-employer transitions. Table 2 summarizes the estimates of firms' non-wage values. As each firm's non-wage value is only identified relative to a base firm, I standardize the distribution of firms' non-wage values to have mean zero.

Estimation of wage components Under my search model, wages assume the following AKM form:

$$ln(w_{it}) = \alpha_i + \psi_{\boldsymbol{J}(i,t)} + \boldsymbol{X}'_{it}\beta + r_{it}, \tag{4}$$

where α_i is a person fixed effect representing the fully portable component of wage capacity of individual i, and $\boldsymbol{X}'_{it}$ is a set of time-varying controls. [45,46] $\psi_{\boldsymbol{J}(i,t)}$ is the wage premium paid by firm j to every worker.[47] $\boldsymbol{J}(i,t)$ indicates the workplace of worker i in year t, and r_{it} is the residual. I estimate equation 4 separately for the 1996–2003 and 2004–2011 panel (columns 2 and 4 in Table 1), where I rely on the procedure by Kline et al. (2020) to calculate the (co)variances of the person and firm effects.[48]

[44] By the definition of employer-to-employer transitions, for the 2004–2011 panel only person-year observations from 2003 to 2010 are at risk of being hired by another firm because they need to work one full calendar year for the new firm after they are hired. Hence, the appropriate sample period for the calculation of g_j is 2003 to 2010. With regard to the hires made by firms, only full-time workers hired from other firms in the years 2004 to 2011 can enter my sample as employer-to-employer transition. Therefore, the appropriate time period to calculate the measure of firm offer intensity, f_j, is 2004 to 2011. The same reasoning applies for the 1996–2003 panel.

[45] The person fixed effect and the time-varying terms in X are only identified under a normalizing assumption. Following Card et al. (2018) I assume that $X'_{it}\beta = 0$ at age 40, that is, the person effects are measured as of age 40.

[46] The relation to $\tilde{\alpha}_i$ of my search model is: $\tilde{\alpha}_i = \alpha_i + X_{it}\beta$.

[47] The relation to $\tilde{\psi}_j$ of my search model is: $\tilde{\psi}_j = \psi_j - \mathbb{E}_k[\eta_{ij}|\text{accepted job offers}]$, i.e., the wage premium firm j offers equals its AKM wage premium minus the expected value of the wage residual among all offers made by firm j that are accepted. So by offering $\tilde{\psi}_j$ the firm ends up paying an average wage premium of ψ_j.

[48] (Co)variances of the person and firm effects when calculated using the OLS point estimates suffer from a bias due to sampling error, often referred to as *limited mobility bias* (Krueger and Summers, 1988; Andrews et al., 2008). Online Appendix G.2 provides details on the estimation of wage components.

Table 3: AKM VARIANCE ESTIMATES

	1996–2003	2004–2011
Variance of person effect	0.1538	0.1568
Variance of firm effect	0.0142	0.0127
Covariance of person and firm effect	0.0055	0.0055
Number of movers	118,942	153,418

Notes: This table reports the (co-)variances of person and firm effects from estimating the AKM wage regression using the procedure by Kline et al. (2020) on the samples in columns 2 and 4 of Table 1. See Tables A.6 and A.7 for a full decomposition of wage variance.

Table 3 summarizes the variation in worker and firm wage effects in the 1996–2003 and the 2004–2011 panel. Variance in person effects explains the largest share of variance in wage, while variance in firm effects is one order of magnitude smaller. The covariance between person and firm effects is positive, reflecting that high-wage workers are sorted to high-wage firms. In the following section, I show how we can combine the estimates from the AKM model with those from my search model to learn about job value inequality between workers, and about its evolution over time. Before that, I briefly discuss how the assumptions underlying the identification of equation 4 are reconciled with my search model.

Search model and AKM When estimating equation 4, I assume that worker mobility is uncorrelated with the time-varying residual component of wage (see Card et al. (2013) for a detailed discussion of this assumption).[49] In my search model, however, workers are the more likely to move to an outside firm the higher the residual component of the wage offer of the outside firm. I show in Online Appendix G.3 that under a condition on firm offer intensity, the identification assumptions of the AKM model can be reconciled with my search model. Intuitively, the reason is that the AKM model identifies firm wage premia from *all* transitions between firms, including those with an interrupting non-employment spell, while my search model only uses *voluntary and direct* transitions between firms for identification.

[49] I show in Online Appendix G.1 that there is no evidence that worker mobility is correlated with the time-varying residual component of wages.

4 The Evolution of Non-Wage Job Values and Implications for Inequality

I will now combine the non-wage value estimates from my search model with information on wages to obtain an estimate of the total job value of each worker in my sample. I will then study the distribution of total job value among workers in the 1996–2003 and the 2004–2011 panel.

Estimating job value Under the assumptions of my search model, each worker employed at a firm in my sample has the following job value:

$$V_{it} = ln(w_{it}) + ln(a^{€}_{\boldsymbol{J}(i,t)}) + \epsilon^{€}_{ij}, \tag{5}$$

where I observe worker i's wage in year t, $ln(w_{it})$, in the data and estimate the non-wage value of his current firm $ln(a^{€}_{\boldsymbol{J}(i,t)})$ with my search model. I do not observe the realization of $\epsilon^{€}_{ij}$, but I can obtain an estimate of its distribution from my search model.[50] I estimate the job value of each person-year observation in the 1996–2003 and 2004–2011 panel (columns 2 and 4 of Table 1) using the corresponding search model estimates. I obtain confidence intervals for the job value distribution estimates by bootstrapping the search model and the AKM decomposition 50 times as detailed in Online Appendix D.

The distribution of job value The first row of Table 4 shows the variance of job value among person-year observations in the 1996–2003 and the 2004–2011 panel. We see that inequality in job value among workers, when measured through the variance,

[50] I know the distribution of $\epsilon^{€}_{ij}$ among offered jobs, which is $\epsilon^{€}_{ij} \sim N(0, \frac{1}{\gamma^2_{JK}})$ and can thus use this distribution in the variance decomposition. By doing so, I ignore the fact that the distribution of $\epsilon^{€}_{ij}$ among accepted job-offers is truncated from below, and thus has smaller variance for workers either hired through an employer-to-employer transition, or for those hired otherwise that have received an outside job-offer in the meantime. I ignore this because I cannot observe outside job-offers. My estimates of the variance of $\epsilon^{€}_{ij}$ among accepted job-offers thus represents an upper bound. In Online Appendix H I derive a lower bound on the variance of $\epsilon^{€}_{ij}$. All my results are qualitatively identical and quantitatively similar when using the lower bound for the variance of $\epsilon^{€}_{ij}$.

Table 4: JOB VALUE VARIANCE 1996–2003 AND 2004–2011

	1996–2003 (1)	2004–2011 (2)	(2)-(1)
$Var(V_{it})$	0.828 [0.711 0.951]	0.949 [0.762 1.140]	0.121 [-0.098 0.346]
$Var(ln(w_{it}))$	0.195 [0.194 0.196]	0.197 [0.196 0.198]	0.002 [0.001 0.003]
$Var(ln(a^{€}_{\boldsymbol{J}(i,t)}) + \epsilon^{€}_{ij})$	0.531 [0.425 0.643]	0.614 [0.447 0.780]	0.082 [-0.113 0.283]
$2Cov(ln(w_{it}), ln(a^{€}_{\boldsymbol{J}(i,t)}) + \epsilon^{€}_{ij})$	0.101 [0.087 0.120]	0.138 [0.117 0.162]	0.037 [0.009 0.061]
$2Cov(\alpha_i, ln(a^{€}_{\boldsymbol{J}(i,t)}))$	0.108 [0.098 0.123]	0.119 [0.103 0.136]	0.011 [-0.010 0.031]
$2Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a^{€}_{\boldsymbol{J}(i,t)}))$	-0.015 [-0.023 -0.007]	0.014 [0.006 0.022]	0.029 [0.017 0.040]

Notes: Variance of job value and covariances of job value components 1996–2003 and 2004–2011. 90% confidence intervals in parantheses, bootstrapped as described in Online Appendix D. The variance of firm non-wage value is after subtracting variance from estimation error using equation 3. Full variance-covariance matrices in Tables A.6 and A.7.

increased by 0.121 from 1996–2003 to 2004–2011.[51] This corresponds to an increase of 15 percent. While the bootstrapped confidence intervals in parantheses in Table 4 tell us that this increase is not significant at usual confidence levels, still a large majority of 80 percent of all bootstrap draws show an increase in job value variance.

To understand what drives the increase in job value inequality, note that

$$Var(V_{it}) = Var(ln(w_{it})) + Var(ln(a^{€}_{\boldsymbol{J}(i,t)}) + \epsilon^{€}_{ij}) + 2Cov(ln(w_{it}), ln(a^{€}_{\boldsymbol{J}(i,t)}) + \epsilon^{€}_{ij}), \tag{6}$$

that is, the variance in job value can be decomposed into wage variance, variance in non-wage value, and the covariance between wage and non-wage value. Rows 2-4 of Table 4 show how these components contribute to total job value variance. We see that 20 to 25 percent of job value variance stems from variance in wage, around 65 percent from variance in non-wage value, and 10 to 15 percent from the covariance between wage and

[51] I focus on the variance of job value as inequality metric. The reason is that other measures of inequality (e.g., Gini index, Theil index) would depend on the location of non-wage value, which I cannot identify.

non-wage value. The variance in wage is almost the same in both periods, reflecting the very stable wage structure in Austria between 1996 and 2011. The variance in non-wage value increased from 1996–2003 to 2004–2011, but this increase is very imprecisely estimated.

A substantial part of the increase in job value variance is attributable to the statistically significant increase in covariance between wage and non-wage value, as shown in the 4th row of Table 4. The covariance between wage and non-wage value is positive in both periods, reflecting sorting of workers with high wage to firms offering high non-wage value. The increase in the covariance thus shows that this sorting got stronger over time.

Additional insights into the increase in job value inequality can be gained by examining the covariance between non-wage value and the AKM components of wage, that is,

$$\begin{aligned} Cov(ln(w_{it}), ln(a^{€}_{\boldsymbol{J}(i,t)}) + \epsilon^{€}_{ij}) = Cov(\alpha_i, ln(a^{€}_{\boldsymbol{J}(i,t)})) + Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a^{€}_{\boldsymbol{J}(i,t)})) \\ + Cov(\boldsymbol{X}'_{it}\beta, ln(a^{€}_{\boldsymbol{J}(i,t)})), \end{aligned} \tag{7}$$

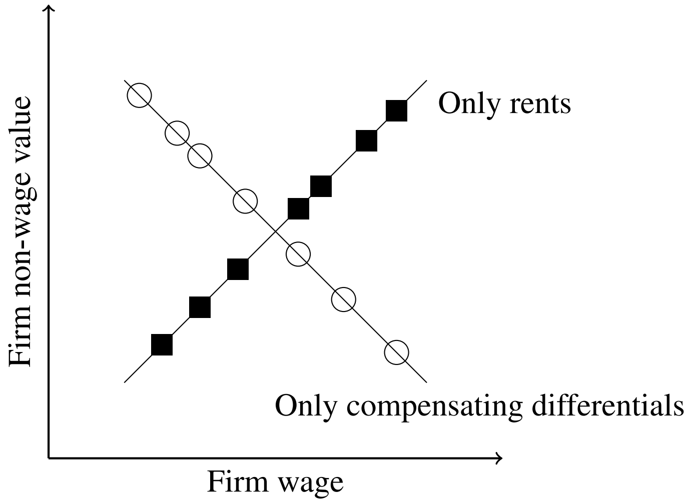
where $Cov(\alpha_i, ln(a^{€}_{\boldsymbol{J}(i,t)}))$ measures how workers with different wage capacity are sorted to firms with respect to firms' non-wage value. The row $Cov(\alpha_i, ln(a^{€}_{\boldsymbol{J}(i,t)}))$ of Table 4 shows that workers with higher wage capacity are sorted to firms offering higher non-wage value in both periods. While this sorting explains about 12 percent of overall job value variance, it only increased slightly between 1996–2003 and 2004–2011.

$Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a^{€}_{\boldsymbol{J}(i,t)}))$ measures how firm non-wage value covaries with firm wage premium. Figure 2a shows how the relationship between $\psi_{\boldsymbol{J}(i,t)}$ and $ln(a^{€}_{\boldsymbol{J}(i,t)})$ can be interpreted as evidence for compensating differentials and rents. Intuitively, if there is no variation in rents that firms offer and firms fully compensate through wage for the quality of their non-wage characteristics, firm wage and non-wage value will be perfectly negatively correlated. If there are no compensating differentials and dispersion of wage and non-wage value is purely due to firms offering rents, firm wage and non-wage value will be perfectly positively correlated. The covariance of firm wage and non-wage value thus reflects the sum of these two effects.

A negative value of $Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a^{€}_{\boldsymbol{J}(i,t)}))$ implies compensating differentials have

Figure 2: RELATION OF FIRM WAGE & FIRM NON-WAGE VALUE

(a) Theoretical: Compensating Differentials and Rents



(b) Empricial: Compensating Differentials and Rents 1996–2003 and 2004–2011

Notes: Figure a shows the theoretical relationship between firm wage and firm non-wage value in two limit cases; 1., when there is full compensation between firm wage and firm non-wage value and thus no rent dispersion, and 2., when there is no compensation between firm wage and firm non-wage value and thus all of firm wage and firm non-wage value is rents. Figure b shows a scatterplot of the actual distribution of firm wage and firm non-wage value in 1996–2003 and 2004–2011. The lines in Figure b represent an OLS regression of firm non-wage value on firm wage, with firms weighted by the number of people-years they represent.

an attenuating effect on job value inequality. A positive value of $Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a^{€}_{\boldsymbol{J}(i,t)}))$ implies that job value inequality is exacerbated by firm-level rents. As shown in Figure 2b and the last row of Table 4, there is a striking and statistically significant difference in $Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a^{€}_{\boldsymbol{J}(i,t)}))$ between 1996–2003 and 2004–2011. It is negative in 1996–2003, and positive in 2004–2011. Thus, compensating differentials had a substantial inequality attenuating effect in 1996–2003, but this effect vanished and was dominated by increased dispersion in firm-level rents in 2004–2011.[52]

Relation to the Literature The evidence presented in this section echoes several findings from the literature on wage differentials between industries and firms, and on compensating wage differentials. Pierce (2001), Maestas et al. (2018), Marinescu et al. (2021), Dube et al. (2022) and Sockin (2022) show that various non-wage characteristics are better for workers earning higher wages, which is consistent with the positive correlation between the person wage effect and the non-wage value I find. Krueger and Summers (1988), find that industry wage premia cannot be explained as compensating differentials for non-wage characteristics, which is consistent with the positive correlation between firm wage and firm non-wage value I find for 2004–2011.[53] Hall and Mueller (2018) estimate the non-wage values of jobs offered to unemployed job seekers. They find a correlation of $-.17$ between the wage and non-wage value of jobs, close to the correlation of $-.09$ between firm wage and non-wage value I find for 1996–2003.

Taber and Vejlin (2020) find that the standard deviation of workers' total flow utility exceeds the standard deviation of wage by 50 percent, while I find that the standard deviation of job value exceeds the standard deviation of wage by 105 to 120 percent. I differ from Taber and Vejlin (2020) in that I study total job value (including expectations about future flow job value), while Taber and Vejlin (2020) study workers' flow utility. Indeed, after removing wage growth expectations my estimates are more similar to Taber and Vejlin (2020) with a standard deviation of job value that exceeds the standard deviation of wage by only 75 to 90 percent (Table A.11).[54]

Hamermesh (1999) shows that, over time, non-wage values can evolve differentially along the wage distribution because of income effects, that is, workers use their productivity gain over time differentially to buy higher wage or higher non-wage value. This channel is not at work in my study because the Austrian wage structure remained al-

[52]Changes to the components of job value variance not reported in Table 4 only have a minor impact on the evolution of job value inequality between 1996–2003 and 2004–2011. The full variance-covariance matrices of job value components can be found in Tables A.6 and A.7.

[53]Katz et al. (1989) find a slightly stronger positive correlation between industry pay premia and the quality of non-wage characteristics.

[54]Table A.8 compares the parameters I identify with those identified by Hall and Mueller (2018) and Taber and Vejlin (2020).

most constant 1996–2011. I add to the findings of Hamermesh (1999) by showing that inequality in non-wage compensation can also change over time even in the absence of changes in the wage structure.

5 Robustness

The validity of this study also depends on the extent to which mechanisms not captured by my search model can account for the observed pattern of mobility between firms. I will now present evidence addressing concerns that my results are driven by expectations about future wage growth, preference heterogeneity, or my assumption on the process generating employment offers.

Wage growth In the baseline specification of my model expected future wage growth is an element of firm non-wage value. This is intentional, as the aim of the article is to provide a measure that can be directly compared to inequality measures based on workers' *current* wage. Nevertheless, it is of interest to evaluate whether my inequality results might be driven by workers expecting to earn higher wages in the future at particular firms.

In Online Appendix I I re-estimate the model excluding expected wage growth over the coming 4, 6, and 8 years. The results show that wage growth indeed reflects a substantial part of non wage value. For example, the standard deviation of firm non-wage job value is reduced by about one third if expected wage growth over the next 8 years is excluded. Nevertheless, all results regarding inequality of job value compared to wage inequality remain qualitatively the same, and the evolution over time is also quantitatively similar if wage growth expectations are excluded.

Preference heterogeneity and match-specific non-wage value Preference heterogeneity, meaning that different workers *perceive* the non-wage value of a firm differently, and match-specific non-wage value, meaning that different workers are *offered* a different non-wage value at a firm, have the same implications for my model. I will thus now only refer to preference heterogeneity, noting that the discussion and the provided evidence also apply for match-specific non-wage value.

Preference heterogeneity over firms' non-wage characteristics is allowed for in my model by the idiosyncratic component of worker utility. My model does, however, not account for potential systematic preference heterogeneity between groups of workers. If there is systematic preference heterogeneity over firms' non-wage value between groups that are compared, assuming common preferences when identifying firms' non-wage value may lead to biased results. To see this, suppose that low-wage workers prefer working in low-wage industries, while high-wage workers equally strongly prefer working in high-wage industries. Estimating my model with these preferences would result in firms' non-wage value being some weighted average of high-wage and low-wage workers' preferences. This would potentially lead me to infer differences in non-wage values between high-wage workers and low-wage workers, while both actually perceive the same non-wage value at their firms.

If preference heterogeneity between high and low-wage workers were important, we should observe different mobility patterns of high-wage workers compared to low-wage workers. As a result, my model should, when it is estimated using employer-to-employer transitions of workers with wages *above* the median, estimate different non-wage values than when it is estimated using employer-to-employer transitions of workers with wages *below* the median.[55] However, this not the case, leading me to conclude that systematic preference heterogeneity does not have an important impact on my results (Table A.10).

Offer generating process An arguably strong assumption of my model is that all firms direct an identical share of job offers to non-employed workers, which implies that employed workers receive offers from the same distribution as non-employed workers. This assumption allows me to estimate the distribution of offers across firms that employed workers face from where non-employed workers get hired. An alternative assumption on how firms direct offers is that every job is first offered to an employed

[55]To test this, I would ideally estimate firms' non-wage values separately using the sample of high and low-wage workers and compare them. This is not possible, however, because different firms are strongly connected in the sample of high and low-wage workers (recall that firms' non-wage values are only identified within the strongly connected set). I can, however, estimate my model using transitions of low-wage workers between firms strongly connected by transitions of low-wage workers, and check whether I obtain similar non-wage values when adding transitions of high-wage workers between these firms to the sample.

worker, and if and only if the employed worker rejects the offer is the job offered to a non-employed worker. If offers are generated following this process, I can estimate the offer distribution employed workers face from the number of workers a firm hires from both employment and non-employment.

I obtain non-wage values very similar to my baseline estimates when estimating the model under this alternative offer generating assumption (Table A.9). To further confirm that my results are not driven by the assumption on the offer generating process, I estimate the model under the naive, and deliberately unrealistic, assumption that all firms are of equal size and make equally many offers. Even holding firm size and the number of offers constant across firms does not qualitatively change any of my results regarding job value inequality. I thus conclude that the assumption on the offer generating process does not drive my results.

6 Conclusion

The aim of this article is to estimate total non-wage job value, and show how the distribution of non-wage value among workers affects labor market inequality. I develop a labor market search model in which workers value both wage and non-wage value of jobs. I estimate the model using a large sample of full-time workers in Austria for the periods 1996–2003 and 2004–2011.

The key finding is that job value dispersion increased over time, in spite of a stable wage structure. The main reason is that compensating wage differentials, attenuating job value inequality, lost importance, while rents, exacerbating job value inequality, became more important. A natural follow-up question would be to study the underlying drivers of this change on the labor demand and the labor supply side. Potential explanations include technological progress leading to a decline in the cost of non-wage value provision for firms paying compensating differentials (Rosen, 1986), and an overall decrease in labor supply elasticity leading to increased rent dispersion (Manning, 2021).

The parsimonious model I develop allows for a tractable mapping of non-wage value estimates to descriptive evidence on wage differentials and worker flows. The flipside is that my model is quite stylized and does not incorporate features like persistent preference heterogeneity over firms' non-wage values. While I provide evidence that these caveats are unlikely to alter my conclusions regarding job value dispersion and inequality, it may be desirable to enrich the model for future studies.[56] Fruitful avenues could be the study of non-wage value differences from job changes associated with events such as child birth or involuntary job loss.

APPENDIX

A Identification

A.1 Theorem 1

The key is finding the values of (w_j, w_k) so that

$$Pr(S_i = j|w_j, w_k, a_j, a_k) = 0.5,$$

which means we must find the wage at which half of workers choose firm j, and the other half chooses firm k. We do not observe all worker choices. Specifically, we only observe when a worker at j receives an offer from firm k, and accepts the offer, but not when a worker at j receives an offer from firm k and rejects the offer. However, by Assumption 1 we observe the number of workers firm j and k hire from non-employment. Because non-employed workers accept all job offers and as search is random, we also know the intensity with which firms j and firm k make offers to employed workers. Combining this with the number of workers at firm j and k, which is also known by Assumption 1, we know how many offers firm j makes to workers at firm k relative to the number of offers firm k makes to workers at firm j. Denoting $EE_{jk}|w_j, w_k$ the number of workers making an employer-to-employer transition from firm j to firm k with wage offer (w_j, w_k) and k the offer intensity and g the firm size, the goal is to find

[56] An interesting approach would be to combine the model with search frictions with features of Lamadon et al. (2022), who model the labor market without search frictions but with persistent preference heterogeneity over firms' non-wage value.

(w_j, w_k) such that

$$\frac{EE_{jk}|w_j, w_k}{EE_{kj}|w_j, w_k} = \frac{f_k g_j}{f_j g_k},$$

which is all observed by Assumption 1.

For this value of (w_j, w_k) it must then hold that

$$Pr(\epsilon_{ik} - \epsilon_{ij} \leq V(w_j, a_j) - V(w_k, a_k)) = 0.5.$$

But as by Assumption 2 $\epsilon_{ik} - \epsilon_{ji}$ has median zero this implies that

$$V(w_j, a_j) = V(w_k, a_k), \tag{A.1}$$

and we know by Assumption 3 that at least one such combination of (w_j, w_k) exists.

Lets define a standard expenditure function $e()$ where $w_j = e(a_j, V(w_j, a_j))$ is the wage needed to achieve utility $V(w_j, a_j)$ given a_j, and $w_k = e(a_k, V(w_k, a_k))$ the wage needed to achieve utility $V(w_k, a_k)$ given a_k. Using Equation A.1 we can write $w_k = e(a_k, V(w_j, a_j))$. Then, by the definition of compensating variation,

$$w_k - w_j = e(a_k, V(w_j, a_j)) - e(a_j, V(w_j, a_j)),$$

is the monetary amount needed to compensate the worker for the difference in non-wage value between a_j and a_k.

In other words, the compensating differential for $a_j - a_k$ equals $w_k - w_j$ and is thus identified in monetary terms. QED.

A.2 Theorem 2

Note that we can observe from the data

$$
\begin{aligned}
Pr(S_i = j|w_j, w_k, a_j, a_k) &= Pr(V(w_j, a_j) + \epsilon_{ij} > V(w_k, a_k) + \epsilon_{ik}) \\
&= \frac{EE_{kj}|w_j, w_k}{rf_j g_k} \\
&= \frac{EE_{kj}|w_j, w_k}{rf_j g_k} * \Bigg(\underbrace{\frac{EE_{kj}|w_j, w_k}{rf_j g_k} + \frac{EE_{jk}|w_j, w_k}{rf_k g_j}}_{=1}\Bigg)^{-1} \\
&= \frac{EE_{kj}|w_j, w_k}{f_j g_k} * \left(\frac{EE_{kj}|w_j, w_k}{f_j g_k} + \frac{EE_{jk}|w_j, w_k}{f_k g_j}\right)^{-1},
\end{aligned}
\tag{A.2}
$$

where the second line reflects the share of workers at firm k who chose firm j when facing the binary choice between firm k and firm j. To see this, let's denote r a constant such that $f_k * r$ is the *number* of offers made by firm k to each worker. While r is unobserved, it cancels out after some algebra allowing us to obtain the expression in the last line which is solely in terms of elements observed in the data by Assumption 1.

Using Assumption 5 we can write this as

$$
\begin{aligned}
Pr(S_i = j|w_j, w_k, a_j, a_k) =& Pr(\gamma ln(w_j) + ln(a_j) + \epsilon_{ij} > \gamma ln(w_k) + ln(a_k) + \epsilon_{ik}) \\
&= Pr(\epsilon_{ik} - \epsilon_{ij} \leq \gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)),
\end{aligned}
$$

which is the cumulative distribution function of $(\epsilon_{ik} - \epsilon_{ij})$ evaluated at the point $(\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k))$. By Assumption 4 γ is strictly positive and by Assumption 6 either w_j or w_k varies along the full positive real line. Thus the evaluation point $(\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k))$ varies along the full real line. By varying the point of evaluation one can identify the full distribution of $(\epsilon_{ik} - \epsilon_{ij})$. QED.

A.3 Theorem 3

Using Assumption 7 we can write

$$\begin{aligned} Pr(S_i = j|w_j, w_k, a_j, a_k) =& Pr(\epsilon_{ik} - \epsilon_{ij} \leq \gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)) \\ =& \Phi[\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)], \end{aligned}$$

where Φ denotes the standard normal cumulative distribution function and the last step uses Assumption 7 and the fact that $\epsilon_{ik} - \epsilon_{ij} \sim N(0, 1)$.

Assume for now that we observe *all* worker choices after receiving a job offer, i.e., not only when workers accept the job offer from the outside firm (and make an employer-to-employer transition), but also when they decline the job offer from the outside firm and choose to stay with their current firm. Using the notation that $\Omega = Pr(S_i = j|w_j, w_k, a_j, a_k)$, the joint likelihood of all observed worker choices at a wage pair w_k, w_j will then be

$$\begin{aligned} \mathcal{L} = & \underbrace{\Phi[\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)]^{rf_jg_k\Omega}}_{\text{Accepted offers from firm } j \text{ to firm } k\text{'s workers}} \\ & * \underbrace{\Phi[\gamma(ln(w_k) - ln(w_j)) + ln(a_k) - ln(a_j)]^{rf_jg_k(1-\Omega)}}_{\text{Rejected offers from firm } j \text{ to firm } k\text{'s workers}} \\ & * \underbrace{\Phi[\gamma(ln(w_k) - ln(w_j)) + ln(a_k) - ln(a_j)]^{rf_kg_j(1-\Omega)}}_{\text{Accepted offers from firm } k \text{ to firm } j\text{'s workers}} \\ & * \underbrace{\Phi[\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)]^{rf_kg_j\Omega}}_{\text{Rejected offers from firm } k \text{ to firm } j\text{'s workers}}, \end{aligned} \tag{A.3}$$

which simplifies to the following log-likelihood

$$\begin{aligned} ln(\mathcal{L}) = r(f_jg_k + f_kg_j)\Big(& \Omega ln(\Phi[\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)]) \\ & + (1-\Omega) ln(\Phi[\gamma(ln(w_k) - ln(w_j)) + ln(a_k) - ln(a_j)]\Big). \end{aligned} \tag{A.4}$$

This would allow consistent estimation of γ and $ln(a)$ under standard regularity conditions the Maximum Likelihood Estimator, but is not feasible given that we only observe job offers leading to an employer-to-employer transition.

However, note that if we plug all employer-to-employer transitions into the likelihood function in Theorem 3 we obtain

$$\mathcal{L} = \underbrace{\Phi[\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)]^{rf_jg_k\Omega * \frac{1}{f_j}\frac{1}{g_k}}}_{\text{Accepted offers from firm } j \text{ to firm } k\text{'s workers}}$$
$$* \underbrace{\Phi[\gamma(ln(w_k) - ln(w_j)) + ln(a_k) - ln(a_j)]^{rf_kg_j(1-\Omega) * \frac{1}{f_k}\frac{1}{g_j}}}_{\text{Accepted offers from firm } k \text{ to firm } j\text{'s workers}}, \tag{A.5}$$

which simplifies to the following log-likelihood

$$ln(\mathcal{L}) = r\Bigg(\Omega ln(\Phi[\gamma(ln(w_j) - ln(w_k)) + ln(a_j) - ln(a_k)])$$
$$+(1-\Omega)ln(\Phi[\gamma(ln(w_k) - ln(w_j)) + ln(a_k) - ln(a_j)]\Bigg). \tag{A.6}$$

Note that Equation A.6 is equivalent to A.4 up to the multiplicative constant $(f_jg_k + f_kg_j)$. Therefore, Equation A.6 allows for consistent estimation of γ and $ln(a)$.[57] QED.

B Jackknife Correction

The jackknife correction by Hahn and Newey (2004) The jackknife correction by Hahn and Newey (2004) builds on the insight from Neyman and Scott (1948) that in non-linear panel data models with fixed length T the fixed-effects estimator γ will not converge in probability to its true value γ_0 but to $\gamma^T = \gamma_0 + \frac{B}{T} + \frac{D}{T^2} + O(\frac{1}{T^3})$, where B and D are parameters that do not depend on T. Therefore, each jackknife estimator based on a subsample excluding observations of the tth period will converge to γ^{T-1}. Hence the jackknife correction estimator γ_{JK} will converge in probability to

$$\gamma_{JK} = T\gamma^T - (T-1)\gamma^{T-1}$$
$$= \gamma_0 + \left(\frac{1}{T} - \frac{1}{T-1}\right)D + O(\frac{1}{T^2}) = \gamma_0 + O(\frac{1}{T^2}),$$

and will thus have bias reduced to order $\frac{1}{T^2}$.

[57] While the likelihood function is for a particular pair of w_j, w_k, identification of γ comes from variation in w_j, w_k as shown in Theorem 2.

Implementing the jackknife correction in my framework I face two challenges when applying the method by Hahn and Newey (2004) to my framework. First, the "length" of the panel T, corresponding to the number of observations contributing to the identification of each fixed effect, varies across firms according to the number of employer-to-employer transitions that are observed for the respective firm. Second, the standard jackknife approach of excluding some observations is not applicable to my framework, because my estimator is only identified for strongly connected firms, and removing some employer-to-employer transitions would break some of the necessary links between firms.

To deal with the first problem I calculate the average number of firm-to-firm transitions used to identify each firm non-wage value. This corresponds to $\frac{2n(S)}{n(J)}$, where 2 accounts for the fact that each employer-to-employer transition entails two firm non-wage values. I then randomly assign all $n(J)$ transitions to $t = 1, ..., \frac{2}{n(J)}$ groups, and estimate the jackknife estimator based on a subsample excluding observations of the tth group.

To deal with the second problem, I assign a weight of 0.01 to excluded observations instead of dropping them from the sample. This allows me to preserve the strongly connected structure of the network, while at the same time ensuring that excluded observations are unimportant when maximizing the likelihood function.

Monte Carlo evidence The idea of the Monte Carlo (MC) simulation is to simulate datasets of employer-to-employer transitions between firms with known underlying γ_0, and to then apply the estimator in Theorem 3 and the jackknife correction and evaluate the bias in the estimate of γ. To that end I generate 50 MC replications using the following steps:

1. Generate 1,000 firms with non-wage value $ln(a^{€})$ randomly drawn from $N(0, \sigma^2_{ln(a^{€})})$, where $\sigma^2_{ln(a^{€})}$ corresponds to the variance of firm non-wage value in the 2004–2011 panel (Table 2).

2. Generate job offers made between these 1,000 firms such that each firm makes offers to 4 employees at 12 other firms. Generate the wage differential for each job offer as a random draw from $N(\overline{\Delta ln(w)}, \sigma^2_{\Delta ln(w)})$, where $\overline{\Delta ln(w)}$ is the mean

wage differential and $\sigma^2_{\Delta ln(w)}$ the variance of the wage differential of employer-to-employer transitions in the 2004–2011 panel. Generate an idiosyncratic utility shock $\epsilon^{€}$ for each job offer as a random draw from $N(0, \frac{0.5}{\gamma_0^2})$, where I use γ_{JK} estimated in the 2004–2011 panel for γ_0.

3. Generate the sample of employer-to-employer transitions as the subsample of the job offers where $\Delta ln(w) + ln(a_j^{€}) - ln(a_k^{€}) + \epsilon_j^{€} - \epsilon_k^{€} > 0$ (i.e., the accepted job offers), where j denotes the firm making the offer and k denotes the firm receiving the offer.

I then apply the estimator in Theorem 3 and the bias correction described above. Table A.1 shows in column (1) descriptive statistics on the MC samples, and in column (2) and (3) the corresponding statistics in the two empirical panels. The panel *Descriptives* shows that, except for the number of firms, the MC samples closely match characteristics of the two empirical panels.[58]

Table A.1: JACKKNIFE CORRECTION

	Monte Carlo (1)	1996–2003 (2)	2004–2011 (3)
Descriptives			
Number of Firms	940	4,544	5,944
$Var((ln(a^{€}))$	0.39	0.43	0.49
Mean transitions per firm	12.26	12.84	12.50
Mean other firms connected	12.22	9.97	11.08
Mean transition wage difference	0.04	0.06	0.05
Variance transition wage difference	0.05	0.05	0.05
Estimates			
γ_0	2.05		
$\hat{\gamma}$	2.35	2.76	2.65
$\hat{\gamma}_{JK}$	1.94	2.19	2.05

Notes: This table shows descriptive statistics and estimates for the 50 Monte Carlo simulation samples (column (1)) and the 1996–2003 panel (column (2)) and the 2004–2011 panel (column (3)). γ_0 is missing for column (1)–(2) because it is only known in the Monte Carlo simulation samples.

The panel *Estimates* shows that the average γ^T exceeds γ_0 by 15 percent, while γ_{JK}

[58] I limit the number of firms in the MC simulations to 1,000 for computational reasons.

deviates from γ_0 by only 5 percent. We see that γ is shrunken more substantially in the empirical panels than in the MC simulation. This might be attributable to features of the empirical panels that are not matched by the MC simulation, like the number of firms or the unequal distribution of the number of transitions across firms in the empirical panels. Overall, the MC simulation shows that incidental parameter bias is substantially reduced, and that the extent of bias remaining after the correction is low.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Acemoglu, Daron and David Autor, "Skills, tasks and technologies: Implications for employment and earnings," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward, "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

Arcidiacono, Peter and Paul B Ellickson, "Practical methods for estimation of dynamic discrete choice models," *Annu. Rev. Econ.*, 2011, *3* (1), 363–394.

Azar, José A, Steven T Berry, and Ioana Marinescu, "Estimating labor market power," Technical Report, National Bureau of Economic Research 2022.

Becker, Dan, "Non-wage job characteristics and the case of the missing margin," *Available at SSRN 1805761*, 2011.

Berger, David W, Kyle F Herkenhoff, Andreas R Kostøl, and Simon Mongey, "An Anatomy of Monopsony: Search Frictions, Amenities and Bargaining in Concentrated Markets," Technical Report, National Bureau of Economic Research 2023.

Bonhomme, Stephane and Gregory Jolivet, "The pervasive absence of compensating differentials," *Journal of Applied Econometrics*, 2009, *24* (5), 763–795.

Brown, Charles, "Equalizing differences in the labor market," *The Quarterly Journal of Economics*, 1980, *94* (1), 113–134.

Büchi, Simon, "Lohnsteuerdaten," *University of Zurich*, 2008.

Bundeskanzleramt, "Beamtinnen und Beamte, Vertragsbedienstete – Pragmatisierungsstopp," Technical Report, Bundeskanzleramt der Republik Østerreich 2021. https://www.oeffentlicherdienst.gv.at/fakten/bundespersonal/bedienstete/index.html, Accessed on February 6th 2023.

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

— **and Dale T Mortensen**, "Wage differentials, employer size, and unemployment," *International Economic Review*, 1998, pp. 257–273.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

— **, Jörg Heining, and Patrick Kline**, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 2013, *128* (3), 967–1015.

— **, Raj Chetty, and Andrea Weber**, "Cash-on-hand and competing models of intertemporal behavior: New evidence from the labor market," *The Quarterly journal of economics*, 2007, *122* (4), 1511–1560.

Dube, Arindrajit, Suresh Naidu, and Adam D Reich, "Power and Dignity in the Low-Wage Labor Market: Theory and Evidence from Wal-Mart Workers," Working Paper 30441, National Bureau of Economic Research September 2022.

Fink, Martina, Esther Segalla, Andrea Weber, and Christine Zulehner, "Extracting firm information from administrative records: The ASSD firm panel," Technical Report, NRN Working Paper, NRN: The Austrian Center for Labor Economics and the Welfare State 2010.

French, Eric and Christopher Taber, "Identification of models of the labor market," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 537–617.

Glassner, Vera and Julia Hofmann, "Austria: from gradual change to an unknown future," *Collective bargaining in Europe: towards an endgame*, 2019, p. 33.

Greene, William, "Panel data models for discrete choice," in Nelson B.H. Baltagi, ed., *The Oxford Handbook of panel data*, Oxford: Oxford University Press, 2015, chapter 6, p. 171–201.

Gruetter, Max and Rafael Lalive, "The importance of firms in wage determination," *Labour Economics*, 2009, *16* (2), 149–160.

Hahn, Jinyong and Whitney Newey, "Jackknife and analytical bias reduction for nonlinear panel models," *Econometrica*, 2004, *72* (4), 1295–1319.

Hall, Robert E and Andreas I Mueller, "Wage dispersion and search behavior: The importance of nonwage job values," *Journal of Political Economy*, 2018, *126* (4), 1594–1637.

Hamermesh, Daniel S, "Changing inequality in markets for workplace amenities," *The Quarterly Journal of Economics*, 1999, *114* (4), 1085–1123.

Heckman, James J and Bo E Honore, "The empirical content of the Roy model," *Econometrica: Journal of the Econometric Society*, 1990, pp. 1121–1149.

Hwang, Hae-shin, Dale T Mortensen, and W Robert Reed, "Hedonic wages and labor market search," *Journal of Labor Economics*, 1998, *16* (4), 815–847.

__ , W Robert Reed, and Carlton Hubbard, "Compensating wage differentials and unobserved productivity," *Journal of Political Economy*, 1992, *100* (4), 835–858.

Jarosch, Gregor, "Searching for job security and the consequences of job loss," Technical Report, National Bureau of Economic Research 2021.

Katz, Lawrence F, Lawrence H Summers, Robert E Hall, Charles L Schultze, and Robert H Topel, "Industry rents: Evidence and implications," *Brookings Papers on Economic Activity. Microeconomics*, 1989, *1989*, 209–290.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-out estimation of variance components," *Econometrica*, 2020, *88* (5), 1859–1898.

Krueger, Alan B and Lawrence H Summers, "Efficiency wages and the inter-industry wage structure," *Econometrica: Journal of the Econometric Society*, 1988, pp. 259–293.

Lalive, Rafael and Josef Zweimüller, "How does parental leave affect fertility and return to work? Evidence from two natural experiments," *The Quarterly Journal of Economics*, 2009, *124* (3), 1363–1402.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," Technical Report 1 2022.

Leoni, Thomas, Wolfgang Pollan et al., "Lohnentwicklung und Lohnunterschiede in der Industrie seit 2000," *WIFO Monatsberichte, WIFO October*, 2011.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till von Wachter, and Jeffrey B Wenger, "The Value of Working Conditions in the United States and Implications for the Structure of Wages," Working Paper 25204, National Bureau of Economic Research October 2018.

Manning, Alan, "Monopsony in labor markets: A review," *ILR Review*, 2021, *74* (1), 3–26.

Marinescu, Ioana, Yue Qiu, and Aaron Sojourner, "Wage Inequality and Labor Rights Violations," Working Paper 28475, National Bureau of Economic Research February 2021.

Mas, Alexandre and Amanda Pallais, "Valuing Alternative Work Arrangements," *American Economic Review*, December 2017, *107* (12), 3722–59.

Matzkin, Rosa L, "Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models," *Econometrica: Journal of the Econometric Society*, 1992, pp. 239–270.

Nekoei, Arash and Andrea Weber, "Does extending unemployment benefits improve job quality?," *American Economic Review*, 2017, *107* (2), 527–61.

Neyman, Jerzy and Elizabeth L Scott, "Consistent estimates based on partially consistent observations," *Econometrica: Journal of the Econometric Society*, 1948, pp. 1–32.

Pierce, Brooks, "Compensation inequality," *The Quarterly Journal of Economics*, 2001, *116* (4), 1493–1525.

Robinson, Joan, *The Economics of Imperfect Competition*, London: Macmillan, 1933.

Rosen, Sherwin, "The theory of equalizing differences," *Handbook of labor economics*, 1986, *1*, 641–692.

Roussille, Nina and Benjamin Scuderi, "Bidding for talent: Equilibrium wage dispersion on a high-wage online job board," Technical Report, Working Paper 2022.

Sockin, Jason, "Show Me the Amenity: Are Higher-Paying Firms Better All Around?," 2022.

Sorkin, Isaac, "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 2018, *133* (3), 1331–1393.

_, "Ranking Firms Using Revealed Preference: Online Appendix," 2018. https://drive.google.com/file/d/1J0M4Rfr_jA8a2Sq7XFgsbQfEtbZgfEQ_/view, Accessed on February 6th 2023.

Stiglbauer, Alfred, Florian Stahl, Rudolf Winter-Ebmer, and Josef Zweimüller, "Job creation and job destruction in a regulated labor market: The case of Austria," *Empirica*, 2003, *30* (2), 127–148.

Sullivan, Paul and Ted To, "Search and nonwage job characteristics," *Journal of Human Resources*, 2014, *49* (2), 472–507.

Taber, Christopher and Rune Vejlin, "Estimation of a Roy/search/compensating differential model of the labor market," *Econometrica*, 2020, *88* (3), 1031–1069.

Wiswall, Matthew and Basit Zafar, "Preference for the Workplace, Investment in Human Capital, and Gender," *The Quarterly Journal of Economics*, 08 2018, *133* (1), 457–507.

Zweimüller, Josef, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf, and Simon Büchi, "Austrian social security database," *Available at SSRN 1399350*, 2009.