Year : 2023

Essays in Labor Economics

Lehmann Tobias

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT D'ÉCONOMIE

ESSAYS IN LABOR ECONOMICS

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteur ès Sciences Économiques,
mention « Économie politique »

par

Tobias Patrick LEHMANN

Directeur de thèse
Prof. Rafael Lalive

Co-directrice de thèse
Prof. Camille Terrier

Jury

Prof. Paul André, président
Prof. Rustamdjan Hakimov, expert interne
Prof. Philipp Kircher, expert externe
Prof. Nicole Maestas, experte externe

LAUSANNE
2022

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT D'ÉCONOMIE

ESSAYS IN LABOR ECONOMICS

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteur ès Sciences Économiques,
mention « Économie politique »

par

Tobias Patrick LEHMANN

Directeur de thèse
Prof. Rafael Lalive

Co-directrice de thèse
Prof. Camille Terrier

Jury

Prof. Paul André, président
Prof. Rustamdjan Hakimov, expert interne
Prof. Philipp Kircher, expert externe
Prof. Nicole Maestas, experte externe

LAUSANNE
2022

UNIL | Université de Lausanne
HEC Lausanne
Le Décanat
Bâtiment Internef
CH-1015 Lausanne

IMPRIMATUR

Sans se prononcer sur les opinions de l'auteur, la Faculté des Hautes Etudes Commerciales de l'Université de Lausanne autorise l'impression de la thèse de Monsieur Tobias Patrick LEHMANN, titulaire d'un bachelor en Business Administration de l'Université de Berne, titulaire d'un master en Economie de l'Université de Zurich, en vue de l'obtention du grade de docteur ès Sciences économiques, mention « économie politique ».

La thèse est intitulée :

ESSAYS IN LABOR ECONOMICS

Lausanne, le 09 décembre 2022

La Doyenne

Marianne SCHMID MAST

Thesis Committee

Prof. Rafael LALIVE
Thesis supervisor
University of Lausanne

Prof. Camille TERRIER
Thesis co-supervisor
Queen Mary University London

Prof. Rustamdjan HAKIMOV
Internal member of the doctoral committee
University of Lausanne

Prof. Philipp KIRCHER
External member of the doctoral committee
Cornell University

Prof. Nicole MAESTAS
External member of the doctoral committee
Harvard University

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Tobias Patrick LEHMANN

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: R. Lalive Date: October 5, 2022

Prof. Rafael LALIVE
Thesis supervisor

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Tobias Patrick LEHMANN

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: Date: 03/10/2022

Prof. Camille TERRIER
Thesis co-supervisor

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Tobias Patrick LEHMANN

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: Date: 03/10/2022

Prof. Rustamdjan HAKIMOV
Internal member of the doctoral committee

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Tobias Patrick LEHMANN

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: Philipp Kircher Date: 5.10.2022

Prof. Philipp KIRCHER
External member of the doctoral committee

University of Lausanne
Faculty of Business and Economics

PhD in Economics
Subject area Political Economy

I hereby certify that I have examined the doctoral thesis of

Tobias Patrick LEHMANN

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Nicole Maestas

Signature: ______________________________ Date: ________10-4-22______

Prof. Nicole MAESTAS
External member of the doctoral committee

Essays in Labor Economics

Tobias Lehmann

2022

Acknowledgements

I owe gratitude to many people who made this dissertation possible, but I am particularly thankful to my advisors Camille Terrier and Rafael Lalive. Rafael provided me with an inspiring environment from the first day of my PhD. He invited me to join his reading group, and a little later to work on a research project with him and Camille. I learned incredibly much from Camille and Rafael through this project, from how to approach a research question to how to write up an article in an exciting yet scientifically rigorous way. At the same time, their doors were always open to provide me with inspiring feedback and advice on my own research projects. What I appreciated the most, however, was their personal support by being kind, empathic, and encouraging at all times. Camille and Rafael provided me with the environment that every PhD student wishes for.

I would like to express my deepest gratitude to Josh Angrist and David Seim for hosting me at MIT and Stockholm University. I learned much from them and the inspiring environment they allowed me to enjoy.

I would also like to thank the members of my thesis committee – Rustam Hakimov, Philipp Kircher, and Nicole Maestas – for their insightful comments and challenging questions.

Since starting the PhD program in 2007, some exceptionally bright people have crossed my path. I have learned a lot from discussions with Fabrizio Colella, Mitch Downey, John Horton, Arash Nekoei, David Strömberg, and Josef Zweimüller. Thank you for sharing your time and knowledge.

I would also like to thank my mother Ursula, my father Robert, and last and most important my wonderful fiancée Anna. This dissertation would not have been possible without the love and warmth that you so unconditionally provide.

CONTENTS

Introduction

Many of the outcomes in our lives do not only depend on our own effort and performance, but also on who we undertake them with. For example, how well an employee fares, that is, how much he earns, how his career progresses, how much he enjoys his job, does not only depend on his skills and how much effort he puts into work, but also on his employer. On the other hand, even the best entrepreneur will probably not succeed if she fails to hire the right people. An even stronger example is the search for a partner. Here whom we match up with not only explains outcomes, but is itself the outcome of interest.

At first sight the labor market and the (heterosexual) marriage market show little resemblance. A closer look, however, reveals a great deal of overlap. Both markets are characterized by heterogeneous agents on two sides searching for a matching partner on the other side. In the labor market we have workers with differing skills on one side and firms with differing productivity on the other, while the heterosexual dating market is characterized by men and women with different traits and physical attributes. The search for matching partners, as well as the matches made, can be the source of concern in these markets. One reason is that not everyone can get the partner he or she prefers the most due to capacity constraints: A firm can only hire a limited number of workers, and a woman can only marry one man. The ambition of this thesis is to provide empirical evidence enlarging our understanding of: how agents in two-sided markets search for and find partners; what problems and inefficiencies occur in this process; and how this may result in inequalities in outcomes.

The first chapter of this thesis studies inequality in the labor market. Current research aimed at understanding inequality in the labor market almost exclusively focuses on wages, while a recent literature suggests that non-pecuniary aspects of jobs are equally important for workers. Examples for non-pecuniary aspects are whether one can work from home, or whether the content of the work is meaningful (Maestas et al., 2018). I demonstrate how we can extract information about the non-pecuniary value of a job from observing how workers switch jobs between firms. In the context of Austria, I show that workers with high wages also tend to hold jobs that offer a high non-pecuniary value. Thus, when only looking at wages, the degree of inequality in the labor market is underestimated. In addition, I also provide evidence that inequality in non-pecuniary aspects of jobs has become more pronounced over time. The first chapter thus shows that the firm a job seeker matches with not only has important implications for his wage, but also for how well off he is with respect to other job characteristics.

While the focus of the first chapter is on the distribution of outcomes in a matching market, the second chapter, jointly written with Camille Terrier and Rafael Lalive, studies problems that might occur in the process leading up to a match. In the context of a heterosexual online dating market, the second chapter shows how the common endeavor of men to match with the most popular women can harm all men in this market. The problem is that women only have limited time available, and therefore do not manage to consider all men that would be interested. As a result, some men who would have had a good chance of matching with the woman never get to be considered by her. Based on the PageRank algorithm used by Google, we design a novel measure of how likely it is that a woman will like a man and vice versa. We show that this measure can be useful in helping actors in matching markets better target whom they approach, thus making the matching process more efficient for everyone.

After considering inequality in matching market outcomes and inefficiencies in the process leading up to a match, the third chapter of this thesis studies how agents decide which potential matching partner they contact. For the sake of simplicity, most workhorse models of the labor market assume that within his occupation or industry a job seeker just randomly sends applications to firms seeking to hire individuals in his occupation or industry. A recent literature, however, advocates models in which workers and firms deliberately decide whom they get in contact with. This literature shows that how agents behave in the search and application process can have important implications for our understanding of aggregate outcomes, including, for example, the duration of job-seeker unemployment (Hornstein et al., 2011). In the third chapter I show that job seekers are much more likely to apply for jobs that are also considered interesting by other job seekers. Likewise, firms are more likely to contact job seekers that other firms like. Interestingly, I find that job seekers often refrain from applying to jobs where they have very low chances of being invited for an interview, suggesting that job seekers not only care about how good a job is when deciding where to apply, but also about their probability of ultimately getting offered the job.

This thesis shows that it is important to have a detailed understanding of how agents in two-sided markets search for matching partners, that such an understanding can help in mitigating problems occurring in the matching process, and to understand resulting inequalities in outcomes. While the empirical applications in this thesis are limited to the labor market and the market for online dating, the concepts and methodologies developed can be readily applied to many other contexts in which heterogeneous agents on two sides of a market seek to match with one or several agents on the other side.

Chapter 1

Non-Wage Job Values and Implications for Inequality

Non-Wage Job Values and Implications for Inequality *

Tobias Lehmann

October 2022

Abstract

I study inequality in job values, both in terms of wages and non-wage values, in Austria over the period 1996 to 2011. Identification of non-wage job values is based on patterns of worker flows between firms and wage differentials. Intuitively, firms with high non-wage value attract workers without paying a wage premium. Looking at the distribution of job value among workers, I find a positive correlation between wage and non-wage value. Inequality in job value is thus greater than wage inequality. Job value inequality increases between 1996 and 2011, although wage inequality remains constant. This is due to a change in the relationship between the part of wage that is systematically attributable to a firm, the firm wage premium, and the non-wage value that firms offer. Between 1996 and 2003, firms' wage premium and their non-wage value are negatively correlated, reflecting compensating differentials attenuating job value inequality. In the 2004 to 2011 period, however, this correlation becomes positive. Compensating differentials disappear because providing non-wage value becomes cheaper over time for firms initially offering low non-wage value. The disappearance of compensating differentials comes with an increase in the dispersion of job value offered by firms. Using a model of monopsonistic competition, I provide evidence that this is caused by two developments over time: first, workers respond less to firms' job value offers, reflecting a decline in the elasticity of labor supply. Second, labor supplied to firms offering low value increases disproportionately because of labor immigration.

Keywords: Inequality, Amenities, Worker heterogeneity, Firm heterogeneity, On-the-job search, Wage dispersion, Matched employer–employee data

JEL Classification Numbers: E24, J31, J32

*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 also thank Jakob Beuschlein, Adrien Bilal, Matthias Doepke, Jonathan Cohen, Mitch Downey, Lea Fricke, Rustamdjan Hakimov, John Horton, Simon Jäger, Gregor Jarosch, Philipp Kircher, Nicole Maestas, Andreas Mueller, Arash Nekoei, Laura Pilossoph, Heather Sarsons, Julian Schärrer, David Seim, Isaac Sorkin, David Strömberg, Pinar Yildirim, and seminar participants at University of Lausanne, MIT labor lunch, Stockholm University, as well as conference audience at EALE, IZA summer school, SKILS, SSES, and YSEM for helpful comments.

1 Introduction

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.[1]

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) and Dube et al. (2022) show that non-wage characteristics tend to be worse in low-wage jobs, therefore exacerbating inequality in job value compared to wage inequality.[2] 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 remains an open question how labor market inequality is affected if *all* non-wage characteristics of jobs are taken into account.[3] Knowing the value of *all* non-wage characteristics of jobs, however, is necessary for statements about inequality in workers' overall well-being.

In this article I address this question by estimating the total non-wage value each worker has at his 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 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

[1] *"The ultimate desideratum is a grand measure of inequality in the returns to work that embodies all monetary and nonpecuniary returns."* (Hamermesh, 1999, p. 1086)

[2] 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. Dube et al. (2022) focus on a set of characteristics related to workplace dignity, and Marinescu et al. (2021) on labor rights violations.

[3] For example, high ranked firms in the *Glassdoor Best Places To Work In 2021* ranking are often associated with *transparent senior leadership* and *mission-driven company culture*, suggesting that such intangible characteristics are important for workers too (Glassdoor, 2021).

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.[4] 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. My main sample focuses on male full-time workers. I split the sample into two consecutive 8-year intervals to study developments over time. The 1996–2003 sample covers 800,000 workers at 4,500 employers, and the 2004–2011 sample covers 960,000 workers at 5,900 employers.

I measure voluntary employer-to-employer transitions, which are those that do not follow a layoff, or firm-level dynamics such as firm mergers and takeovers.[5] 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 even accept a wage decrease.[6] 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 and Mortensen (1998). Workers search for job offers, which they receive at Poisson rate. Employers' job offers consist of a wage, and an employer-specific non-wage value. In addition, workers have an idiosyncratic valuation for each employer. When receiving an offer from an outside employer, workers compare it to the offer of their current employer, and transition to the outside employer if it offers them greater value than their current employer. I assume that the value of a job for a worker is an additive combination of the log-wage, the employer-specific non-wage value, and the worker-employer idiosyncratic value.[7]

My model gives rise to a simple probit-style likelihood function, where every likelihood contribution represents a job-to-job transition between two employers.[8] I account for differing employer sizes and the intensity with which employers make job offers to each other's em-

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

[5] I exclude layoffs to the extent they are observed registered by the unemployment agency, and apply the procedure by Sorkin (2018) to account for unobserved layoffs at contracting firms.

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

[7] The underlying assumption is that workers' valuation for employers' 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.

[8] 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 employer.

ployees by appropriately weighting each likelihood contribution.[9] I allow for heterogeneity between workers in two ways: First, I let the intensity with which workers receive offers from different employers depend on the worker's current employer.[10] Second, I allow the non-wage value workers are offered by an employer to be heterogeneous through a worker-employer idiosyncratic value component. I estimate three parameters with my model: The first is each employer's non-wage value.[11] 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. The third parameter is the variance of the employer-worker idiosyncratic non-wage value.

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 wage information from my data, which gives me an estimate of 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. In both periods, 1996–2003 and 2004–2011, 43 percent of job value variance is explained by wage, and 57 percent by non-wage value.[12]

I find that between 1996–2003 and 2004–2011, job value variance increases by 8 percent. Job value variance can increase for three reasons: variance of wage, variance of non-wage, and their covariance. I find that neither the variance of wage nor the variance of non-wage value did increase much. Thus, the main driver of the increase in job value variance is an increase in covariance between wage and non-wage value. To understand the sources of this increase, I decompose wage following Abowd et al. (1999) into a part systematically attributable to worker quality, and a 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 1996–2003 the covariance between firm wage premium and firm non-wage value is negative, whereas it is positive in 2004–2011.[13]

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

[9] While I directly observe employer size in the data, I follow Bonhomme and Jolivet (2009) and Sorkin (2018) and estimate the intensity with which employers 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.

[10] With this, I allow for sorting of workers across employers.

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

[12] This is close to Taber and Vejlin (2020) finding that 49 percent of job value variance is explained by wage, and 51 percent by non-wage value.

[13] 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.

sures the importance of compensating differentials relative to firm-level rents (Robinson, 1933; Rosen, 1986). Intuitively, if 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. My results show that compensating differentials attenuated job value inequality 1996–2003. By 2004–2011, however, they have disappeared and dispersion of firm-level rents has increased, leading to an increase in job value inequality.

What fundamental developments can explain these patterns? I interpret the findings in the framework of a simple monopsonistic competition model (Manning, 2013). In this framework, firms first decide which total value they offer to workers, and second, how to best divide it into wage and non-wage value (Lang and Majumdar, 2004). Thus, I can separately address the question of why rent dispersion increased, and why compensating differentials disappeared. I test multiple potential explanations for the increase in rent dispersion from 1996–2003 to 2004–2011. I find that a decline in the elasticity of labor supply, caused by an increase in the idiosyncrasy of workers' preferences over employers, explains part of the increase in rent dispersion among firms (Card et al., 2018; Lamadon et al., 2021). I then provide evidence that an immigration-induced increase in labor supply for firms offering low value also accounts for part of the increase in rent dispersion (Borjas, 2014).

I show that the disappearing of compensating differentials must be explained by firm-specific (or industry-specific) changes in the marginal cost of non-wage value provision (Rosen, 1986).[14] I derive an estimate of firms' marginal cost of non-wage value provision in 1996–2003 and 2004–2011. I do so by combining my estimates of firms' non-wage value and firms' wage premium with the assumption that firms equalize the marginal cost of providing job value through wage and non-wage value. I find that the cost of non-wage value provision declined most in the construction and the real estate service industry, where firms tend to compensate workers for low non-wage value with a wage premium.

I conclude the paper by discussing the robustness of my results. My model of the labor market allows for tractable identification of non-wage values. The flip side is that it is quite stylized and omits some mechanisms discussed in the literature, including systematic forms of preference heterogeneity, labor market learning, or firm-specific human capital. I provide evidence that these mechanisms are unlikely to have an important effect on my results. Another potential limitation of my framework is related to a data requirement of my estimator: For employers'

[14] This solely relies on the assumption that employers are cost minimizing when deciding how to divide the value they offer to workers between wage and non-wage value, which is plausible even for public sector employers, and with union bargaining.

non-wage value in my model to be identified, a sufficient number of workers moving from and to an employer is required. Small employers often do not satisfy this requirement, meaning that I cannot identify their non-wage value. I show evidence that my sample nevertheless reflects well the overall structure and dynamics of the labor market.

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, 2018; Taber and Vejlin, 2020; Jarosch, 2021). Most closely related are Sorkin (2018) and Taber and Vejlin (2020), who also rely on worker flows between firms to identify total job values. Relative to Sorkin, I incorporate wage differentials in the estimation of my model, which allows me to separately identify the contribution of wage and non-wage value to job value.[15] 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.[16] In contrast, my model directly uses patterns observed in the data and provides a transparent estimation framework. Taber and Vejlin (2020) find that non-wage value accounts for half of workers' flow utility, but do not provide any evidence on how non-wage value varies along the wage distribution.

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.[17] Subsequent work has shown that search frictions (Hwang et al., 1998; Bonhomme and Jolivet, 2009) as well as idiosyncratic preferences of workers over firms (Lamadon et al., 2021; Manning, 2021) can explain this result.[18] My model incorporates both, search frictions and idiosyncratic preferences of workers over firms. I add to the literature by showing in a simple model of monopsonistic competition how they can both 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.

[15] The job value identified by Sorkin are in utility units with an unknown scale and thus cannot be separated into wage and non-wage value.

[16] The richness of the model by Taber and Vejlin (2020) is driven by their attempt 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.

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

[18] 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.

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).[19] 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% 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 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 (Figure A.1).[20]

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.[21] The social security data contain detailed daily information on worker labor market status (e.g., employed, unemployed, retired). Each employment spell is linked to an employer identifier and information on the employer's industry and location.[22]

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 employers. They contain workers' uncensored gross labor earnings,[23] and since the year 2002 an indicator whether an individual is working full-time or part-time. Before 2002, over 97% of working men were full-time employed (Figure A.2).[24] 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.[25]

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

[20]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.

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

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

[23]Including bonus payments.

[24]Only around 70% of women were employed full-time between 1996 and 2002 (Figure A.3).

[25]Table A.1 shows the distribution of full-time workers' weekly hours across industries based on the *Mikrozen-*

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.[26] Individuals in my panel satisfy the following three conditions: (1) The person 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.[27] 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.[28] To ensure my model is well-identified, I only consider employers that have overall at least five employer-to-employer transitions with other employers in the strongly connected set.[29]

sus survey. Industry-level averages of weekly working hours range from 39.8 hours in utilities to 44.4 hours in hotel and restaurant.

[26]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).

[27]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 Sections 3.3 and Appendix E.2).

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

[29]This restriction is motivated by the so-called *incidental parameter bias* (Greene, 2015, pp. 188–192), which is relevant for my model because I identify a large number of fixed effects in a non-linear model (cf. Section 3.3). 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. The resulting strongly connected set contains more than 10 times as many observations (transitions) than subjects (firm) in both periods, 1996–2003 and 2004–2011. It has been shown for the panel fixed effects probit estimator (which is similar to my estimator) that incidental parameter bias is small when there are at least 10 observations per subject (Hahn and Newey, 2004, Table 3 and 4; Greene, 2002, Table 2; Heckman, 1981, Table 4.1).

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
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 (adj.)	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 Appendix G.2)

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.

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 Appendix D, showing that my sample of strongly connected employers well reflects the overall structure of the Austrian labor market, and that dynamics are very similar among workers not considered in my sample. Thus while the sample restrictions should be kept in mind when reading the paper, this provides evidence that my results are nevertheless likely to be meaningful for the Austrian labor market overall.

Employer-to-Employer Transitions Figure 1 shows how I identify employer-to-employer transitions. First, 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.[30]

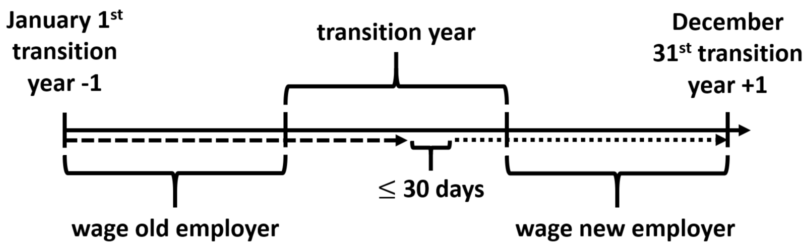
My model is built around the idea that employer-to-employer transitions are the outcome of a worker's choice between a job offer from his old employer and a job offer from his new employer. I therefore exclude all transitions that most likely are not the result of such worker decisions. Specifically, I exclude all transitions that follow a layoff recorded in the social security data.[31] I also exclude all transitions that follow firm-level dynamics such as firm renaming, takeovers, mergers, spin-offs, or firm closures, which I identify following Fink et al. (2010).[32]

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

[31]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.

[32]I identify employer-level dynamics from collective actions of groups of workers, as recorded in the social security data. For example, an employer takeover is identified if an employer-identifier disappears from the records and if at least two thirds of workers work for the same employer in the following quarter. See Appendix B for

Figure 1: EMPLOYER-TO-EMPLOYER TRANSITIONS



Notes: This figure illustrates how I identify employer-to-employer transitions and associated wage differentials. A transition in year t is considered an employer-to-employer transition if the following criteria are satisfied: (1) Less than 30 days between two employment spells, and no unemployment spell in between. (2) The worker works the full calendar year before the year of the transition for the old employer. (3) The worker works the full calendar year after the year of the transition for the new employer.

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 (2018) proposes a probabilistic approach to correct for these transitions. The underlying idea is that these transitions are most likely to happen at contracting firms. To see this approach consider Figure 2, which shows employer-to-employer and employer-to-nonemployment separation probabilities as a function of the annual employer growth. I calculate the average employer-to-employer separation rate at expanding employers, which I use as an estimate for the expected separation rate from voluntary employer-to-employer transitions. When an employer 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 employers with $(1 - \frac{\text{excess}}{\text{excess+expected}})$.[33,34]

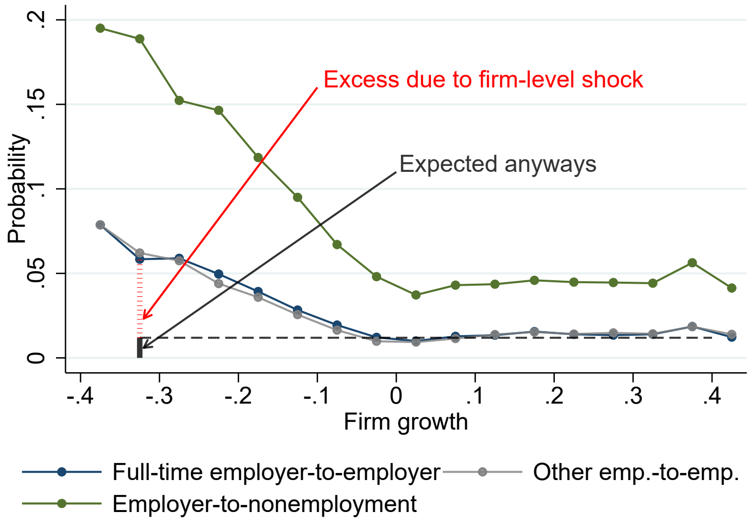
Panel D. of Table 1 shows descriptive statistics on employer-to-employer transitions. 58,349 transitions occur between firms in my sample for 1996–2003 and 74,271 transitions between 2004–2011. In both periods, employer-to-employer transitions come on average with a log wage increase of about 0.05, and 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.

details.

[33]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.

[34]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. This gives qualitatively identical and quantitatively similar results.

Figure 2: EMPLOYMENT GROWTH AND TRANSITION PROBABILITIES



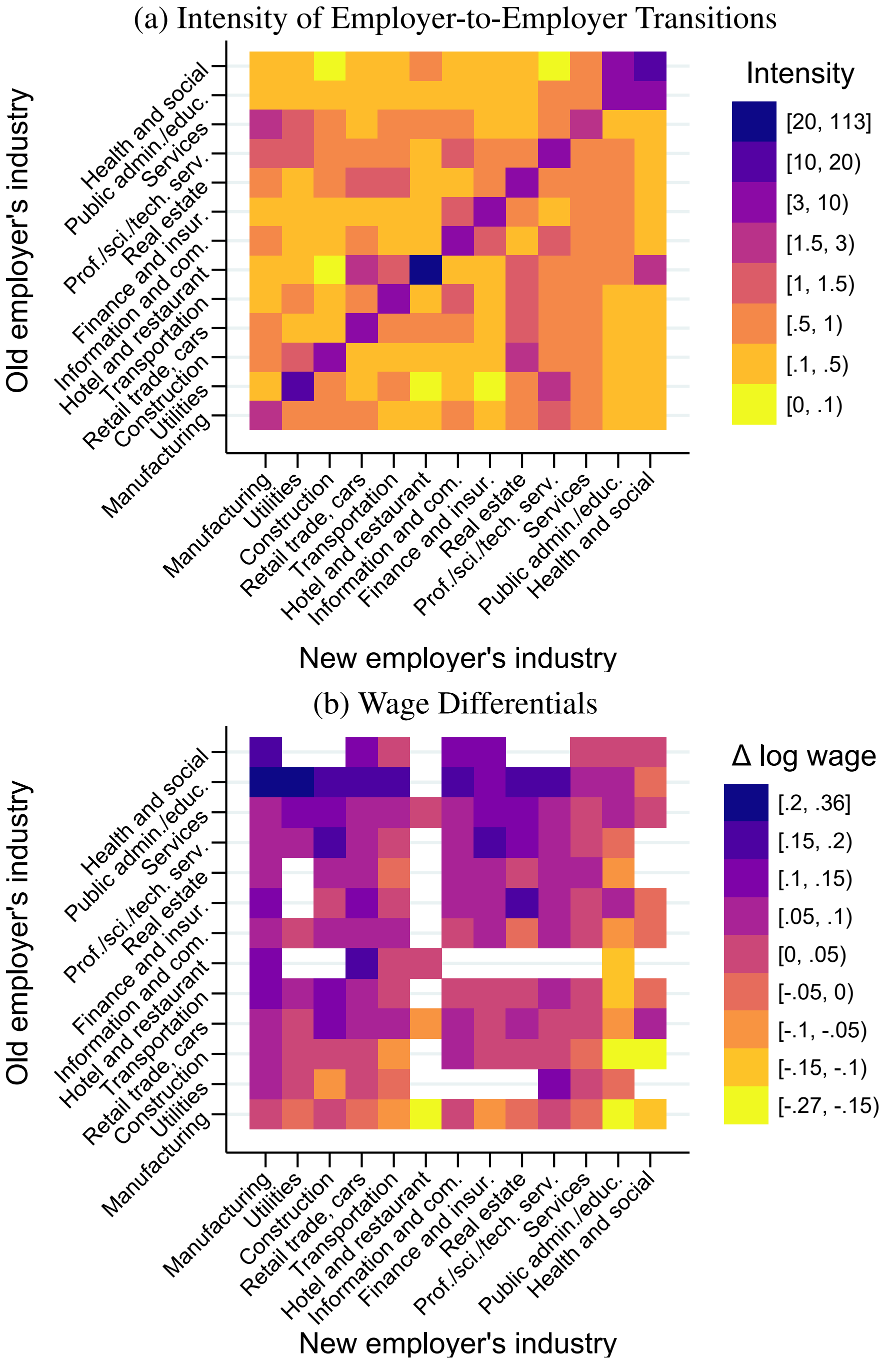
Notes: This figure shows the probability (per year) a worker in column 3 of Table 1 makes a transition, by 0.05 employer growth rate bin. *Full-time employer-to-employer* corresponds to the employer-to-employer transitions defined in this section. *Other employer-to-employer* corresponds to all transitions in which the worker starts at the new employer within 30 days, but do otherwise not satisfy the conditions detailed in this section. *Employer-to-nonemployment* corresponds to employment spells ending in year $t+1$ for which the worker does not join a new employer within 30 days. Share excess transitions $\frac{\text{excess}}{\text{excess+expected}}$. Figure A.7 shows the corresponding figure for the 1996–2003 sample.

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. The same reasoning applies for 1996–2003, and the corresponding industry-level descriptive statistics are shown in Appendix C.

Figure 3a 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.[35] 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

[35]Each cell corresponding to row-industry j and column-industry k equals $\frac{\text{transitions}_{jk}}{\sum_{s\in J}\sum_{l\in J}\text{transitions}_{sl}} * \left(\frac{\sum_{l\in J}\text{transitions}_{jl}}{\sum_{s\in J}\sum_{l\in J}\text{transitions}_{sl}} * \frac{\sum_{s\in J}\text{transitions}_{sk}}{\sum_{s\in J}\sum_{l\in J}\text{transitions}_{sl}}\right)^{-1}$, 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 3: 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 employer in the row-industry in the new employer in the column-industry. Missing cells in figure b contain fewer than 10 observations. Both figures are based on transitions between employers in the strongly connected 2004–2011 sample (column 4 in Table 1). See Figure A.8 for transitions in the 1996–2003 sample, and Figure A.11 for employer-to-employer transitions of all workers over the period 2004–2011.

random mobility counterfactual, and values below 1 intensities below. The large variation in intensities depicted in Figure 3a 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 3a.

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, they 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.[36]

Figure 3b 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.[37] 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 the intuition behind the identification of non-wage values in my model, which I will explain in the following section.

[36]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).

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

3 Identification of Non-Wage Values

In the following, I construct an on-the-job search model in the vein of Burdett and Mortensen (1998). The model is partial equilibrium, meaning that I take firm behavior as exogenously given. Employers 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. In the framework of this model, I interpret voluntary employer-to-employer transition as the result of a binary choice over two job offers, one from the employer that the worker joins and one from the employer that the worker leaves. This then allows me to identify employers' non-wage values from the extent to which worker choice can or cannot be explained by wage differentials. I focus on a discussion of the model's structure and intuition in this section. Additional information on the model, including workers' value functions, can be found in Appendix E.

3.1 Primitives

Employers Each employer $j \in J$ is fully characterized by the tuple $\langle \psi_j, a_j, g_j, \boldsymbol{f_j} \rangle$. ψ_j denotes the log-wage premium, which I assume following Abowd et al. (1999) (henceforth AKM) that the employer pays to every worker equally. a_j denotes the non-wage value employer j offers to all its workers equally. One can think of $a_j = a(\mathbf{m}_j)$, where $\mathbf{m}_j$ is an arbitrary-dimensional vector containing characteristics besides the present wage that are valuable to a worker when working at employer j, and that are converted to a non-wage value through the function $a()$.[38] g_j denotes the size of the employer, that is, the number of employees of employer j. $\boldsymbol{f_j} = [f_{j1}, ..., f_{j,k \neq j}, f_{j,J}]$ denotes the vector of employer j's offer intensities, that is, the intensity with which employer j makes employment offers to workers at other employers.

Workers Employed workers are characterized by the pair $\langle \alpha_{it}, j \rangle$. α_{it} denotes, following AKM, worker i's skills, labor market experience, and other factors for which the worker is compensated equally by all employers. j denotes worker i's current employer. I assume a worker's value from working at employer j is a linear combination of his log wage w_{ij}, the log of his current employer's non-wage value a_j, and the worker's idiosyncratic valuation for

[38] Besides amenities that provide flow-utility to the worker, a_j also contains expectations about future wage and non-wage value, and, due to the absence of a random search assumption, future job opportunities. 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 besides the contemporaneous wage. I show in Appendix I that expectations about future wage growth explain some part of a_j. However, the results and conclusions of this article are unchanged if expectations about wage growth are excluded from a_j.

working at employer j ϵ_{ij}:[39]

$$V_{ij} = \gamma ln(w_{ij}) + ln(a_j) + \epsilon_{ij} \tag{1}$$

This log-additive 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 about the same fraction of wage for various amenities. I normalize $\gamma = 1$, which implies that V_{ij}, $ln(a_j)$ and ϵ_{it} are in log-wage equivalent units.[40]

3.2 Search, Offers and Employer-to-Employer Transitions

Employed workers search for job offers from other employers, which they receive at an exogenous rate. A job offer consists of a pair $\langle ln(w_{ij}), a_j \rangle$, where the wage part of the offer is composed as follows:

$$ln(w_{ij}) = \alpha_{it} + \tilde{\psi}_j + \eta_{ij} \tag{2}$$

the wage firm j offers to worker i consists of a worker-specific part α_{it}, a firm-specific wage premium offer $\tilde{\psi}_j$ and a idiosyncratic part η_{ij}. I assume that η_{ij} is a random draw from a symmetric mean-zero distribution.[41]

When a worker employed at employer j receives a job offer from an outside employer k, he draws a new job offer from his current employer j and makes a binary choice over the two offers:

$$\begin{aligned} P(V_{ik} > V_{ij}) &= P(ln(w_{ik}) + ln(a_k) + \epsilon_{ik} > ln(w_{ij}) + ln(a_j) + \epsilon_{ij}) \\ &= \Phi(ln(w_{ik}) - ln(w_{ij}) + ln(a_k) - ln(a_j)) \end{aligned} \tag{3}$$

where Φ denotes the cumulative distribution function of a normal distribution with mean zero and variance $2\sigma^2$ and the last equality follows from assuming that $\epsilon_{is} \sim i.i.d.\ \mathrm{N}(0, \sigma^2)$, so $(\epsilon_{ij} - \epsilon_{ik}) \sim i.i.d.\ \mathrm{N}(0, 2\sigma^2)$.

With an ideal dataset in which the analyst observes all offers from outside employers and all binary choices of employed workers, the structure put on the environment so far would be sufficient to estimate the $ln(a_j)$ of each employer and σ by simply maximizing the joint-likelihood of all binary choices. With the administrative data I have available, however, I only observe job offers that employed workers accept, which are the employer-to-employer

[39] Throughout the paper the term "utility" refers to job value, that is, the value of the value function, and *not* the flow utility.

[40] γ converts log-wage to job value units. By setting $\gamma = 1$, I set the scale of job value to equal the log-wage scale.

[41] I assume that $\tilde{\psi}_j = \psi_j - \mathbb{E}_k[\eta_{ij}|\text{accepted offers}]$, i.e., that the wage premium employers offer equal the AKM wage premium ψ_j minus the expected value of the wage residual among all offers made by employer j that are accepted. So by offering $\tilde{\psi}_j$ the employer ends up paying an average wage premium of ψ_j.

transitions discussed in the previous section. In particular, I do not observe when workers receive a job offer from an outside employer and choose to stay with the current employer. Hence, in order to render this model estimable with my data, I need some measure of the number of offers employers make to employees of other employers. I follow Bonhomme and Jolivet (2009) and Sorkin (2018) in assuming that non-employed workers search from the same offer distribution as employed workers, and that non-employed workers do not reject job offers. Therefore, I can recover the intensity with which employers make job offers from the number of non-employed workers they hire. In addition, I assume the following for the pattern with which employers make offers to each others' workers:

> Assumption 1: $\frac{f_{jk}}{f_j^{NE}} = \frac{f_{kj}}{f_k^{NE}}$; There exists some measure of the intensity with which employers make offers f^{NE}, and the probability a worker at employer k receives an offer from employer j, relative to the total number of offers made by employer j, equals the probability a worker at employer j receives an offer from employer k, relative to the total number of offers made by employer k.

Intuitively, this assumption states that if an employer makes offers to another employer's workers with a higher intensity than vice versa, then it must also be that this employer makes offers with an overall higher intensity. While this assumption is in line with random search, that is, that all workers receive job offers from a particular employer with equal probability, it is less restrictive in that it allows for directed search, that is, that the probability of receiving a job offer from a particular employer depends on a worker's current employer. For example, in my model f_{jk} is allowed to be higher if employers k and j rely on the same type of labor input (with respect to education, skills, experience), which is intuitively as well as empirically plausible (see, for example, Nimczik, 2020).

This model provides enough structure to identify employers' non-wage values from observed employer-to-employer transitions. In the following, I discuss how I estimate the model.

Proposition 1. 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 employers in J generated under the model above. The joint likelihood of all S transitions is:

$$\mathcal{L} = \prod_{s=1}^{S} \Phi[(ln(w_{(i,t),j}) - ln(w_{(i,t),k})) + ln(a_j) - ln(a_k)]^{\frac{1}{f_j^{NE}} \frac{1}{g_k}}$$

Proof. See Appendix E.2. □

Proposition 1 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 Proposition 1, it is instructive to consider the case when all employers make equally many offers, so f_j^{NE} is constant, and all employers are of equal size, so g_j is constant. In this case, Proposition 1 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 employers j and k the number of workers at employer j that receives an employment offer from employer k, but rejects the offer, is equivalent to the number of workers at employer k that receives an employment offer from employer j and accepts, and vice versa.

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

A very similar intuition applies for the likelihood weight $\frac{1}{g_k}$, which is the inverse of the number of employees at the employer the worker *leaves*. Consider two employers 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 employer than the smaller. By downweighting the likelihood contribution of every employer-to-employer transition away from the larger employer by one half, the estimator accounts for the difference in size between these employers.

Due to the simple form of the likelihood function in Proposition 1, it is relatively easy to pin down the variation identifying employer non-wage value. First, employer non-wage value is driven by the net flow of workers (after accounting for employer size and offer intensity) between employers, where employer j's non-wage value is higher relative to employer k's non-wage value if there is more worker flow from employer k to employer j. This is the same source of variation that Sorkin (2018) exploits. The novelty of my model is that I use wage differentials between employers for identification, where, assuming constant worker flows, employer j's non-wage value is higher relative to employer k's non-wage value if wages are *lower* at employer j relative to employer k.[42] Another novelty of my model is that I allow the probability with which a worker receives a job offer from a particular employer to depend on the worker's current employer.[43] This implies that my model generates sorting of workers to employers without modeling comparative advantages or systematic preferences of workers over employers, thus allowing for high tractability.[44]

3.3 Estimation

I estimate the model separately using the 1996–2003 panel and the 2004–2011 panel. For strongly connected employers, the likelihood function in Proposition 1 is continuous and twice differentiable. This implies that I can use standard maximum likelihood routines in estimation.[45]

Search Model Estimates Table 2 shows the distribution of the two model parameters I use in estimation, for the 1996–2003 and the 2004–2011 panels. I measure the employer size parameter g by the number of people-years per employer. 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.[46]

Employers' non-wage values $ln(a)$ in Proposition 1 are identified relative to a base employer's non-wage value, which I select to be the employer with the most employer-to-employer transitions. Table 2 summarizes the estimates of employers' non-wage values. As each employer's non-wage value is only identified relative to a base employer, I standardize the distribution of employers' non-wage values to have mean zero. To avoid having the dispersion of

[42] Because there is random variation in firms' wage offer through η, which affects workers' probability of accepting a job offer, I can separately identify firm non-wage value and firm wage, even though the only systematic variation of non-wage value and wage (net of the person-specific component) is on the firm-level (see Appendix E.2).

[43] I allow this probability to vary for non-employed workers as well.

[44] Sorting through comparative advantages is typically obtained by modeling production complementarities between worker and employer types (Rogerson et al., 2005; Wright et al., 2021). Another approach is to model persistent preferences of workers over employers, locations and industries (Lamadon et al., 2021).

[45] 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 every likelihood contribution in Proposition 1 by $(1 - \rho_{kt})$, where ρ_{kt} represents the share of employer-to-employer transitions at employer k in year t that are in excess of the expected number of employer-to-employer transitions.

[46] By the definition of employer-to-employer transitions as detailed in Section 2, for the 2004–2011 panel only person-year observations from 2003 to 2010 are at risk of being hired by another employer because they need to work one full calendar year for the new employer 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 some employer j, only full-time workers hired from other employers 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 employer offer intensity on the market for full-time full year workers, f_j, is 2004 to 2011. The same reasoning applies for the 1996–2003 panel.

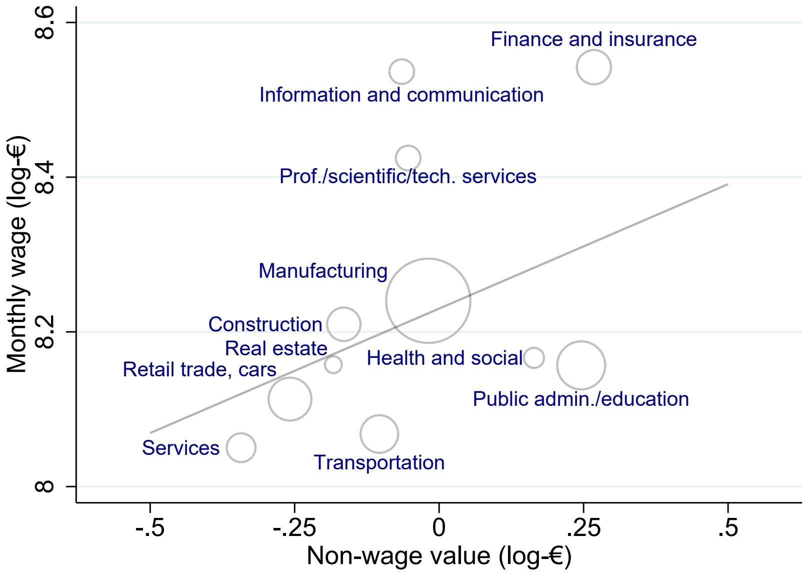
Table 2: SEARCH MODEL ESTIMATES

	Mean	Var	Min	Max
	1996 – 2003			
Parameters				
Firm size (people-years)	799	2136^2	2	56744
Hires from non-employment	60	134^2	1	4918
Estimates				
Firm non-wage value (log €)	0	.36	-3.44	2.38
Shrinked firm non-wage value (log €)	0	.20	-3.04	1.67
Corr(Firm non-wage, Shrinked firm non-wage)	.94			
Number of transitions	58,349			
	2004 – 2011			
Parameters				
Firm size (people-years)	727	1830^2	5	53569
Hires from non-employment	59	132^2	1	4140
Estimates				
Firm non-wage value (log €)	0	.39	-3.45	3.44
Shrinked firm non-wage value (log €)	0	.21	-2.32	1.93
Corr(Firm non-wage, Shrinked firm non-wage)	.94			
Number of transitions	74,271			

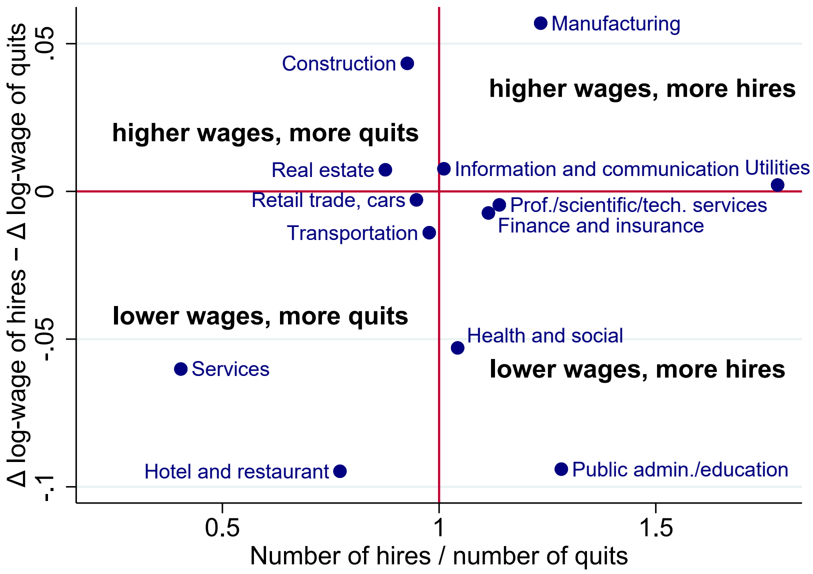
Notes: The panel *Search Model* shows the model parameters and estimates from estimating the model in Proposition 1 on the sample in Table 1, columns 2 and 4. Firm size is measured from 1995 to 2002 (2003 to 2010). Hires from non-employment are measured from 1996 to 2003 (2004 to 2011). Firms' non-wage values are only identified relative to each other and thus standardized to have mean zero. The mean and variance of firm non-wage value and shrinked firm non-wage value are with firms weighted by the number of person-year observations they represent. Shrinkage uses an empirical bayes with industry by federal state averages as prior distribution (see Appendix F for details). I rely on shrinked non-wage values throughout the paper.

Figure 4: HIRES, QUITS, WAGE DIFFERENTIALS AND NON-WAGE VALUES

(a) Non-Wage Values



(b) Employer-to-Employer Transitions and Wages



Notes: Figure a shows non-wage values and log-wages by industry, with circle size relative to the number of person-year observations in the corresponding industry. The gray line plots the regression line run at the industry level, with industries weighted by their number of person-year observations. Two industries are not shown in figure a: Utilities (coordinates: (.85,8.47)) and Hotel and restaurant (-.55,7.79). Figure b shows, on the x-axis, the number of employer-to-employer hires divided by the number of employer-to-employer quits (based on columns 3 and 4 of Table A.4), and on the y-axis: Log-wage increase of employer-to-employer hires minus log-wage increase of employer-to-employer quits (corrected for time/experience effects, based on Table A.5). Figures are based on the 2004–2011 sample (column 4 in Table 1). See Figure A.9 for the 1996–2003 sample.

employers' non-wage values driven by sampling variation, I shrink the estimates of employers' non-wage values towards the respective industry by federal state average using an empirical bayes approach (see Appendix F for details). Table 2 shows that while this reduces variation, in particular in the tails of the distribution, the shrinked non-wage values remain highly correlated with the non-shrinked values. I rely on shrinked non-wage values throughout the paper. Overall, Table 2 does not point to any substantial difference in parameter values and estimates between the 1996–2003 and the 2004–2011 panel.

Figure 4 shows, for the 2004–2011 panel, that my estimates of employers' non-wage values (Figure 4a) intuitively map to summary statistics on hires and quits (Figure 4b).[47] For example, we see in Figure 4a that employers in *public administration/education* offer high non-wage value. Figure 4b shows that employers in *public administration/education* hire more workers than they loose (x-axis), despite paying lower wages (y-axis). An example of a low non-wage value industry is *construction*, where we see in Figure 4b that its employers loose more workers than they hire, despite paying higher wages.

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.[48] The relation to my search model is $\alpha_{it} = \alpha_i + X'_{it}\beta$, that is, the two terms on the right hand side capture the wage an individual is paid by every employer equally. $\psi_{\boldsymbol{J}(i,t)}$ is the wage premium paid by employer j to every worker. $\boldsymbol{J}(i,t)$ indicates the workplace for 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.[49]

Table 3 summarizes the variation in worker and employer 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

[47] A similar picture is obtained for the 1996–2003 sample (Figure A.9).

[48] 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.

[49] (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). Appendix G.2 provides details on the estiation 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.7 and A.8 for a full decomposition of wage variance.

firms.[50] 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 wages (see Card et al. (2013) for a detailed discussion of this assumption).[51] In my search model, however, workers are the more likely to move to an outside employer if the residual component of the wage offer made by the outside employer is higher. Nevertheless, I show in Appendix G.3 that under a condition on firm offer intensity, the identification assumptions of the AKM model are nested in my search model. Intuitively, the reason is that the AKM model identifies employer wage premia from *all* transitions between employers, including those with an interrupting non-employment spell, while my search model only uses *voluntary and direct* transitions between employers for identification.

[50]Comparing the estimates to recent estimates by Bonhomme et al. (2020) and Kline et al. (2020), I find a larger variance of the worker wage effect, while the variance of the firm effect and the covariance between worker and firm effect is smaller, suggesting that differences between firms are less important for wages in my sample (see Table A.6).

[51]I show in 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 estimate the job value of each worker in my sample, and analyze its distribution in the 1996–2003 and the 2004–2011 panel. Guided by a simple model of a monopsonistic labor market, I will then provide evidence on changes in labor market fundamentals driving the observed evolution of job value over time.

4.1 The Distribution of Job Value 1996–2003 and 2004–2011

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

$$V_{it} = ln(w_{it}) + ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{it} \tag{5}$$

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, $ln(a_{\boldsymbol{J}(i,t)})$, in my search model. I do not observe the realization of ϵ_{it}, but I can obtain an estimate of its distribution from my search model.[52] 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.

The Distribution of Job Value Table 4 shows, in the first row, 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, increased by 7.6 percent from 1996–2003 to 2004–2011.[53]

To understand the drivers of this increase in inequality, note that

$$Var(V_{it}) = Var(ln(w_{it})) + Var(ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{it}) + 2Cov(ln(w_{it}), ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{it}) \tag{6}$$

[52]I know the distribution of ϵ_{it} across offered jobs, which is $\epsilon_{it} \sim N(0, \hat{\sigma}^2)$ and can thus use this distribution in the variance decomposition. By doing so, I ignore the fact that the distribution of ϵ_{it} among accepted job-offers is truncated from below, and has thus smaller variance, for workers either hired through an employer-to-employer transition, or workers 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 ϵ_{it} among accepted job-offers should thus be seen as an upper bound. In Appendix H I derive a lower bound on the variance of ϵ_{it}. The only result that is affected by this is the share of job value variance that is due to non-wage value, which decreases to 54 percent in both periods.

[53]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.

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

	1996–2003 (1)	2004–2011 (2)	(2)-(1)
$Var(V_{ij})$	0.524	0.564	0.040
$Var(ln(w_{ij}))$	0.195	0.197	0.002
$Var(ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{ij})$	0.265	0.277	0.012
$2Cov(w_{ij}, ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{ij})$	0.064	0.090	0.026
$2Cov(\alpha_i, ln(a_{\boldsymbol{J}(i,t)}))$	0.075	0.082	0.007
$2Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a_{\boldsymbol{J}(i,t)}))$	-0.015	0.006	0.021

Notes: This table reports the variance of job value, and covariances of job value components in the 1996–2003 sample and in the 2004–2011 sample. The variance-covariance matrix of all job value components is reported in Appendix Table A.7 and A.8.

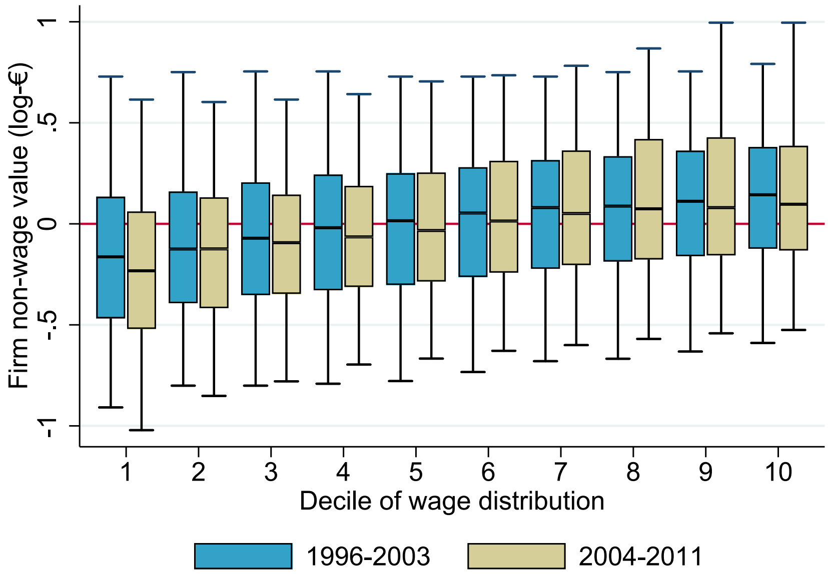
that is, the variance in job value can be decomposed into wage variance, variance in non-wage value, and the covariance between wage and non-wage value. Rows 2-4 of Table 4 show how these components contribute to total job value variance. We see that in both periods of the panel around 35 percent of job value variance stems from variance in wage, around 50 percent from variance in non-wage value, and the rest from the covariance between wage and 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 slightly from 1996–2003 to 2004–2011, contributing about one third to the increase in total job value variance between the two periods.[54]

The other two thirds of the increase in job value variance are attributable to the 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. Graphically, this can be seen in Figure 5, which shows the distribution of non-wage value with workers grouped by decile of the wage distribution. We see a particularly strong downward shift in the non-wage value distribution for workers in the lowest wage decile from 1996–2003 to 2004–2011. At the same time, the distribution of non-wage value for workers with above-median wage shifted slightly upwards. Both together explain the increase in covariance between wage and non-wage value between

[54] Due to limitations on computational resources I only calculate point estimates and no standard errors for the covariances. While bootstrapping of standard errors would in principle be possible, it is not state of the literature to report standard errors for such covariances (e.g., Card et al., 2013, 2016; Song et al., 2019).

1996–2003 and 2004–2011.

Figure 5: NON-WAGE VALUES BY WAGE DECILE



Notes: This figure shows non-wage values of workers' firms, by decile of the wage distribution. The box ranges from the 1st to the 3rd quartile of the firm non-wage value distribution in the respective decile. The whiskers range from the 5th to the 95th percentile of the firm non-wage value distribution in the respective decile.

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,

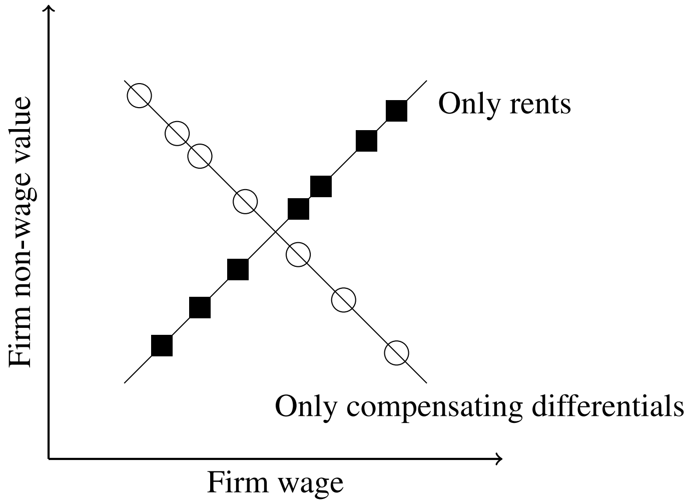
$$Cov(ln(w_{it}), ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{it}) = Cov(\alpha_i, ln(a_{\boldsymbol{J}(i,t)})) + Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a_{\boldsymbol{J}(i,t)})) + Cov(\boldsymbol{X}'_{it}\beta, ln(a_{\boldsymbol{J}(i,t)})) \tag{7}$$

where $Cov(\alpha_i, ln(a_{\boldsymbol{J}(i,t)}))$ measures the extent to which workers with different wage capacity are sorted among firms with respect to the non-wage value they offer. 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 10 percent of overall job value inequality, 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 6a 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.

Figure 6: 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.

A negative value of $Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a_{\boldsymbol{J}(i,t)}))$ implies compensating differentials have 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 6b and the last row of Table 4, there is a striking difference between $Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a_{\boldsymbol{J}(i,t)}))$ in 1996–2003 and 2004–2011. While it is substantially negative in 1996–2003, it is slightly positive in 2004–2011.[55] Thus, compensating differentials had a substantial inequality attenuating effect in 1996–2003, but this effect vanished and is dominated by increased dispersion in firms' job value offers by 2004–2011. This explains more than half of the overall increase in job value inequality between 1996–2003 and 2004–2011.[56]

4.2 Why Did Job Value Inequality Increase?

I will now discuss potential explanations for the increase in job value inequality caused by changes in firms' wage and non-wage value offer. These explanations should account for the following two empirical results: First, for the increase in $Var(ln(a_j)+\psi_j)$, that is, the increase in dispersion of value offered by firms. Second, for the increase in $Cov(ln(a_j), \psi_j)$, reflecting the disappearance of compensating differentials. Lang and Majumdar (2004) show that these two can be considered separately. Intuitively, the firm's problem consists of a stage where it chooses which value to offer, and a stage where it best allocates value between wage and non-wage value (see Appendix J).

Increase in Firm Value Dispersion Because of search frictions and idiosyncratic preferences of workers over firms, firms in my search model face a labor supply that is upwards sloping in the firm value they offer.[57] My search model thus represents a standard monopsony framework as depicted in Figure7 a.[58] Figure 7a shows two firms: one with high marginal revenue product of labor (MRPL), and one with low MRPL, both facing the same labor supply curve. Figure 7a shows that the high MRPL will maximize profits by offering a higher firm value than the low MRPL firm. Thus, dispersion of firm value can arise due to differences in MRPL across firms.

As we can see from Figure 7a, the increase in dispersion of firm value offer I find in Austria can be explained by changes in the slope or location of either, the labor demand or labor supply curves. In particular, it can also be explained by a decrease in labor supply elasticity, as illustrated in Figure 7b by the increased slope of the labor supply curve. Intuitively, the

[55]Figure A.4 describes the underlying change in the relationship between employer-to-employer transitions and employer wage premium. From 1996–2003 to 2004–2011, the relationship between a firm's number of hires relative to its quits and the wage gain of its hires relative to that of its quits became considerably stronger, indicating that firms offering higher wage were considered more attractive by workers in 2004–2011 than in 1996–2003.

[56]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 Table A.7 and A.8.

[57]Firm value = firm wage + firm non-wage value.

[58]This figure is inspired by Manning (2021, Figure 1).

high MRPL firm has a greater incentive to increase its value offer in response to a decrease in labor supply elasticity as it has greater opportunity costs of losing workers and scaling down production (formal derivation in Appendix J.2). I will now evaluate whether there is evidence for a decrease in labor supply elasticity between 1996–2003 and 2004–2011.

The elasticity of labor supply is decreasing in the degree of search frictions in the labor market (Burdett and Mortensen, 1998).[59] I estimate the intensity with which workers receive outside job offers and find no evidence for an increase in search frictions (Table A.10). Another potential reason for a decrease in labor supply elasticity is that labor markets become less segregated, that is, it becomes more likely that workers at high-value firms receive offers from low-value firms and vice versa (Berger et al., 2022).[60] I find no evidence that labor markets became less segregated between 1996–2003 and 2004–2011 (Table A.11).

The third reason why the labor supply elasticity could decrease is an increase in the idiosyncrasy of workers' preferences over firms (Card et al., 2018; Lamadon et al., 2021; Manning, 2021).[61] My search model provides me with an estimate of the variance of workers' idiosyncratic preferences over firms. The variance of workers' idiosyncratic preferences over firms increased by 8 percent from 1996–2003 to 2004–2011. Can thus a decline in labor supply elasticity fully account for the increase in job value dispersion I find? Figure 7 shows that for this to be plausible, there should be two patterns observable in the data: first, firms offering high value in 1996–2003 (high MRPL firms) should have seen a relative decline in employment from 1996–2003 to 2003–2011. Second, large firms (high MRPL firms) should have seen a relative increase in the job value they offer, relative to smaller firms. I observe a decline in employment at high value firms between 1996–2003 and 2004–2011, but do not find that large firms increased their value offer (Table A.12).

Thus, while a decline in labor supply elasticity can account for part of the increase in value dispersion among firms, there must also have been changes in labor supply or labor demand in a way that increases job value dispersion. I can only provide suggestive evidence on this.[62] I use the share of workers with foreign nationality in a firm or a local labor market as a proxy

[59]To see this, consider the case where there are no search frictions, that is, workers can instantaneously choose over the full set of firms in the labor market. This will result in perfectly elastic labor supply (absent idiosyncratic preferences). On the other hand, if there are infinite search frictions, thus workers never receive job offers from outside firms, labor supply to the firm becomes perfectly inelastic (see Appendix J.3).

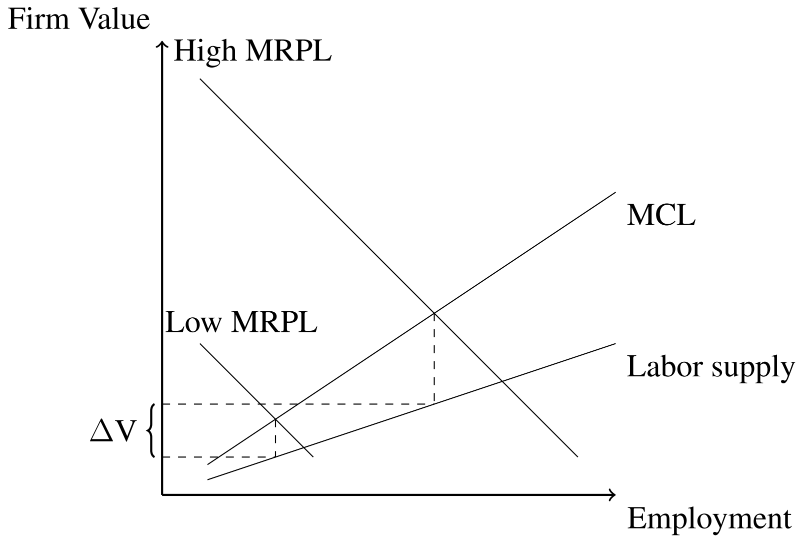
[60]Consider the case where a worker faces the binary choice between a firm A and an outside firm. The more different the outside firm's job value offer, the less firm A's labor supply depends on its own job value offer, because the marginal probability gain in the worker's binary choice from an increase in its own job value offer is lower (see Appendix J.4).

[61]Consider the case where a worker faces the binary choice between a firm A and an outside firm. Increasing idiosyncratic preferences implies that the same increase in job value offer by firm A will lead to a smaller gain in probability the worker will choose firm A over the outside firm.

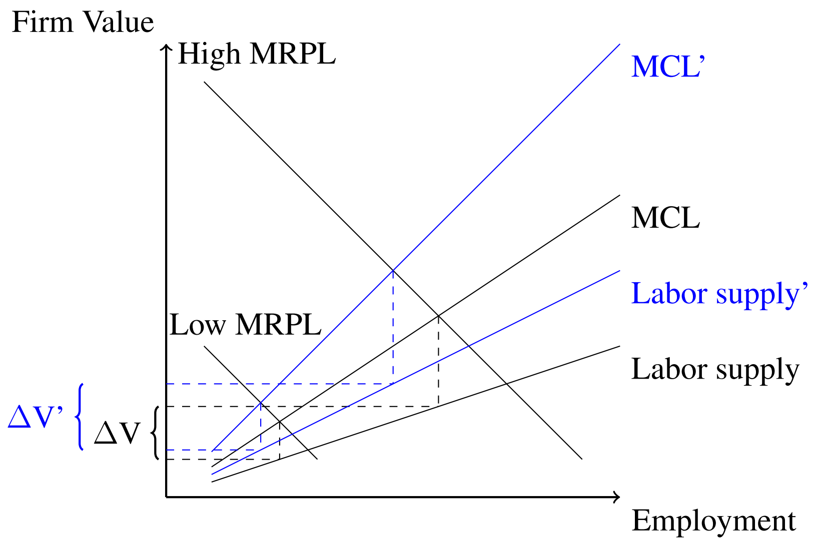
[62]I do not have any information available related to firms' MRPL function.

Figure 7: LABOR SUPPLY AND DEMAND UNDER MONOPSONY

(a) High and Low MRPL Firm



(b) Increase in Elasticity of Labor Supply



Notes: $V = \phi + ln(a)$; MRPL, marginal revenue product of labor; MCL, marginal cost of labor. Figure a illustrates the firm values offered by a firm with high MRPL and a firm with low MRPL, in a model with finite elasticity of labor supply. Figure b illustrates the effect of a decrease in labor supply elasticity on the profit-maximizing firm value offer of the high MRPL and the low MRPL firm.

for an, at least in part, exogenous labor supply shifter. I find that the share of workers with foreign nationality increased more between 1996–2003 and 2004–2011 in firms offering low value in 1996–2003, and that firm value decreased more between 1996–2003 and 2004–2011 in firms where the share of workers with foreign nationality increased more between 1996–2003 and 2004–2011 (Table A.13). This suggests that changes to firm-specific labor supplies also contribute to explaining the increase in dispersion of employer value between 1996–2003 and

2004–2011.

Disappearance of Compensating Differentials The increase in $Cov(ln(a_j), \psi_j)$ I find, reflecting the disappearance of compensating differentials, can be explained by changes in firms' marginal cost of non-wage value provision (Rosen, 1986). For example, compensating differentials could disappear if the marginal cost of non-wage value provision declined in firms that compensate for low non-wage value by paying high wage premia. I cannot directly test this, as I do not know what the cost of non-wage value provision is for firms. However, I show in Appendix J.6 that I can infer firms' marginal cost of non-wage value provision by assuming that firms allocate the value they provide between wage and non-wage value in a cost-minimizing way. The intuition is that in the optimum, a firm's marginal cost of providing utility to workers through non-wage value equals its marginal cost of providing utility through wage. Thus, I can infer each firm's marginal cost of non-wage value provision from the wages it pays. I find that the marginal cost of non-wage value provision declined most between 1996–2003 and 2004–2011 in the construction and the real estate services industry, where firms compensate workers for low non-wage value through a wage premium (Figure A.5). This supports the explanation that firms that compensated for low non-wage values through high wage premia in 1996–2003 increased their non-wage value offer by 2004–2011, because it got cheaper for them to do so.

4.3 Relation to Literature

The evidence presented in this section echoes several findings from the literature on wage differentials between industry and firms, and on compensating wage differentials. Pierce (2001) and Maestas et al. (2018) 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 close to zero correlation between firm wage and firm non-wage value I find for 2004–2011.[63] 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 $-.12$ between firm wage and non-wage value I find for 1996–2003.[64] Taber and Vejlin (2020) find that 51 percent of the variance in workers'

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

[64] Table A.9 compares the parameters I identify with those identified by Hall and Mueller (2018) and Taber and Vejlin (2020).

flow utility is explained by non-wage values, where I find that 57 percent of workers' job value variance is explained by non-wage values.[65]

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 almost constant 1996–2011. I add to the findings of Hamermesh (1999) by showing that inequality in non-wage compensation can also change over time because of increased search frictions, changes in labor demand and supply, and changes in the cost of non-wage value provision for firms.

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 employers. I will now present evidence addressing concerns that my results are driven by preference heterogeneity, firm-specific skills, labor market learning, or my assumption on the process generating employment offers.

Offer Generating Process An arguably strong assumption of my model is that all firms direct an identical share of employment offers to non-employed workers, which implies that employed workers on average 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 worker, and if and only if the employed worker rejects the offer, the job is 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.

Estimating the model under this alternative assumption on the offer generating process, I obtain non-wage values very similar to my baseline estimates (Table A.14). 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

[65]For both, 1996–2003 and 2004–2011 I find that 57 percent of job value variance is explained by non-wage value, which I calculate as $\frac{Cov(ln(a_j)+\epsilon_{it},V_{it})}{Var(V_{it})}$. This can be compared to the estimate by Taber and Vejlin (2020) to the extent that non-wage value in my model is driven by instantaneous non-wage value flows, and not by expectations about future wage and non-wage value flows.

employers does not 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.

Preference Heterogeneity and Match-Specific Amenities Preference heterogeneity, that is, different workers *perceive* the value of the set of amenities they are offered by a particular firm differently, and match-specific amenities, that is, different workers *are actually* offered a different set of amenities by a particular firm, have the same implications for my model. I will thus in the following discussion only refer to preference heterogeneity, noting that the discussion and the provided evidence directly applies to match-specific amenities as well.

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 then 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 is 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, identify different non-wage values than when it is estimated using employer-to-employer transitions of workers with wages *below* the median.[66] However, this not the case. I conclude that systematic preference heterogeneity does not have an important impact on my results (Table A.15).

Labor Market Learning Another alternative explanation for mobility patterns between employers is that transitions are the result of employers learning about worker quality, rather than

[66]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.

the arrival of an offer and a worker's choice.[67] My framework accounts for some forms of labor market learning: To the extent that labor market learning is the same across firms, it is accounted for by workers' idiosyncratic non-wage value draw. Labor market learning that leads to a layoff is accounted for in my model if the layoff either leads to an unemployment spell, or to a reduction in a firm's number of employees.

Nevertheless, it is still possible that labor market learning partly drives employer-to-employer mobility in my sample. Labor marker learning has been shown to be quite quick (Lange, 2007). Thus, if learning were important, we should observe different mobility patterns among young workers, where employers learn a lot about worker quality, as opposed to among old workers, where employers no longer learn much about worker quality. I test for this by splitting the sample of workers at the median worker age, and compare model estimates obtained with young and old workers. I find that non-wage values are highly correlated, and thus conclude that labor market learning is unlikely to affect my results (Table A.15).

Firm-Specific Human Capital A potential concern is that workers acquire firm-specific human capital over time, leading them to earn an idiosyncratic compensation premium at their current firm, which they are not offered by outside firms. Firm-specific human capital would thus violate the assumption of my model that firms offer the same wage and non-wage value to both, current and outside workers. If firm-specific human capital were to drive my model estimates of non-wage values, then the probability a worker accepts a job offer from an outside firm should decline when the worker acquires human capital, that is, with increasing tenure. In particular, the decline should be stronger for firms that I estimate to offer high non-wage value. Figure A.6 presents a test of this prediction at the industry level. I find no evidence that firm-specific human capital is related to my estimate of non-wage values.[68]

Overall, the robustness checks show that my assumption on the offer generating process does not drive my results. I also find no evidence that preference heterogeneity, match-specific amenities, asymmetric labor market learning, or firm-specific human capital have a relevant impact on my results.

[67] Sorkin (2018, 1385–1386) provides a thorough discussion of asymmetric learning and its implications for employer-to-employer mobility. Examples of markets where learning is important include academia (assistant professor tenure track) or law firms (the best will be promoted to partner, the others leave).

[68] The only industry with a markedly distinct pattern in the probability a worker makes a firm-to-firm transition is the *Services*-industry. However, the decline as a function of tenure is *steeper* than in the other industries, while firms in *Services* offer *low* non-wage value.

6 Conclusion

The aim of this article is to estimate non-wage values of jobs, and show how the distribution of non-wage values 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. This finding is likely to be relevant for other developed countries, as many of its potential driving forces, such as industry-specific changes in the cost of non-wage value provision, are more likely to be a global phenomenon than to be specific to the Austrian labor market. At minimum, my findings show that non-wage value of jobs should be considered when monitoring inequality in the labor market, and when designing policies aimed at mitigating it.

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 flip-side is that my model does not incorporate features like systematic forms of preference heterogeneity over firms' non-wage values, or asymmetric learning in the labor market. While I provide evidence that these caveats are unlikely to alter my conclusions regarding job value dispersion and inequality, it might be desirable to enrich the model to incorporate some of these features in future studies.[69] Fruitful avenues could be the study of non-wage value differences from employer switches around events such as child birth or involuntary job loss.

[69] An interesting approach would be to combine the model with search frictions with features of Lamadon et al. (2021), who model the labor market without search frictions but with persistent preference heterogeneity over employers' amenities.

A Additional Tables and Figures

Table A.1: Industry-Level Variation in Full-Time Workers' Weekly Hours

	1996 – 2003					2004–2011				
	Mean	Sd	Median	P10	P90	Mean	Sd	Median	P10	P90
Industry										
Manufacturing	39.9	4.3	39.0	38.0	40.0	41.6	5.4	40.0	38.5	50.0
Utilities	39.8	3.1	40.0	38.0	40.0	42.5	6.2	40.0	38.5	50.0
Construction	40.3	4.4	40.0	38.0	40.0	42.0	5.7	40.0	38.5	50.0
Retail trade, cars	41.1	6.1	40.0	38.0	46.0	42.2	6.2	40.0	38.5	50.0
Transportation	41.4	5.8	40.0	38.0	45.0	43.8	7.4	40.0	40.0	55.0
Hotel and restaurant	44.4	9.9	40.0	38.0	60.0	44.1	8.8	40.0	40.0	55.0
Information and communication	41.4	6.0	40.0	38.0	50.0	43.5	6.3	40.0	38.5	50.0
Finance and insurance	40.0	4.7	39.0	38.0	40.0	43.0	6.3	40.0	38.5	50.0
Real estate	41.2	5.6	40.0	38.0	44.0	42.8	7.4	40.0	38.5	50.0
Prof./scientific/tech. services	42.9	7.9	40.0	38.0	55.0	43.6	7.0	40.0	38.5	53.0
Services	40.9	5.4	40.0	38.0	40.0	42.2	6.3	40.0	38.5	50.0
Public admin./education	40.4	3.3	40.0	38.0	40.0	42.9	6.6	40.0	40.0	50.0
Health and social	41.3	5.7	40.0	38.0	45.0	42.8	7.6	40.0	38.5	50.0
Observations			691,247					393,278		

Notes: This table reports summary statistics on weekly working hours of full-time workers by industry, estimated using data from the Austrian Mikrozensus. Summary statistics are calculated using inverse probability weights provided by Statistics Austria. I classify a worker as full-time worker if he is not self-employed and reports working at least 36 hours in a normal work week. A major reform of the Mikrozensus in 2004, including a change in definition of employment status limits comparability of working hours before 2004 and after 2004 (Lehmann, 2019).

Table A.2: EMPLOYMENT SPELLS ENDING 1996–2003 & 2004–2011

	1996 – 2003		**2004–2011**	
	Count	**Annual hazard**	**Count**	**Annual hazard**
	(1)	(2)	(3)	(4)
Employment spells ending	555,088	0.1493	658,737	0.1486
Thereof due to				
Firm rename	81,010	0.0218	124,129	0.0280
Firm takeover	354	0.0001	189	0.0000
Firm spin-off	72	0.0000	40	0.0000
Firm closure	12,440	0.0033	11,341	0.0026
Spells ending excluding those due to firm dynamics	461,212	0.1240	523,038	0.1180
Thereof				
Employer-to-Nonemployment transitions	201,510	0.0542	247,551	0.0558
All Employer-to-Employer transitions	181,458	0.0488	196,437	0.0443
Thereof				
new job < 1 calendar year or not full time	100,503	0.0270	96,652	0.0218
Employer-to-Employer transitions (full-time & full year)	80,955	0.0218	99,785	0.0225
Thereof				
Involving firm not in strongly connected set	22,606		25,514	
Employer-to-Employer transitions in SC set	58,349	0.0157	74,271	0.0168
Employer-to-Employer transitions in SC set (weighted)	26,931	0.0072	39,426	0.0089

Note: This table shows how I obtain employer-to-emplyer transitions from all employment spells ending at firms in the strongly connected sample (column(2) and (4) of Table 1). Annual Hazard as (number of transitions)/(number of person-year observations). Definition of firm rename, takeover, spin-off, and closure in Appendix B. *Employer-to-Nonemployment*: Employment spells ending with a layoff or with at least 30 days of non-employment after the spell ends. *Employer-to-Employer (full-time & full year)*: All employer-to-employer transitions satisfying the definition in Section 2 *Employer-to-Employer in SC set (weighted)*: All employer-to-employer transitions between firms in the strongly connected set, after reweighting transitions at contracting firms using the procedure described in Section 2.

Table A.3: BY INDUSTRY –
EMPLOYER-TO-EMPLOYER TRANSITION RATES
AT EXPANDING EMPLOYERS

	1996–2003	2004–2011
Manufacturing	0.009	0.009
Utilities	0.005	0.007
Construction	0.011	0.012
Retail trade, cars	0.012	0.012
Transportation	0.014	0.014
Hotel and restaurant	0.009	0.009
Information and communication	0.021	0.018
Finance and insurance	0.012	0.012
Real estate	0.012	0.012
Prof./scientific/tech. services	0.014	0.016
Services	0.017	0.031
Public admin./education	0.006	0.008
Health and social	0.012	0.010

Notes: This table reports the annual probability a worker in the sample makes a employer-to-employer transition as defined in section 2, at firm-years with employment growth ≥ 0.

Table A.4: BY INDUSTRY – NUMBER OF HIRES AND QUITS 2004 – 2011

	Unweighted		Layoff weighted	
	Hires (1)	Quits (2)	Hires (3)	Quits (4)
Manufacturing	24,195	18,007	12,659	10,256
Utilities	2,090	1,087	1,034	580
Construction	4,795	5,837	2,365	2,553
Retail trade, cars	7,737	8,617	4,193	4,425
Transportation	6,512	6,201	3,049	3,120
Hotel and restaurant	247	369	137	178
Information and communication	5,707	4,971	2,106	2,083
Finance and insurance	5,560	5,157	3,050	2,739
Real estate	913	2,103	601	686
Prof./scientific/tech. services	4,614	6,103	2,529	2,220
Services	4,830	9,884	2,600	6,434
Public admin./education	5,523	4,120	4,025	3,140
Health and social	1,163	1,463	840	806

Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for the 2004–2011 sample. Columns 1 and 2: Number of hires and quits by industry. Columns 3 and 4: Number of hires and quits by industry, after downweighting quits from contracting firms according to procedure explained in Section 2.

Table A.5: BY INDUSTRY – WAGES AND WAGE DIFFERENTIALS

	1996–2003 (1)		**2004–2011** (2)	
Median monthly wages by industry (2012 €)				
Manufacturing	3,365		3,569	
Utilities	3,432		4,103	
Construction	3,083		3,194	
Retail trade, cars	3,122		3,262	
Transportation	2,728		2,837	
Hotel and restaurant	2,187		2,240	
Information and communication	4,914		4,561	
Finance and insurance	4,506		4,900	
Real estate	3,233		3,450	
Prof./scientific/tech. services	3,783		4,244	
Services	2,909		2,938	
Public admin./education	2,780		3,081	
Health and social	2,927		3,163	
Δ *log-wage of employer-to-employer transitions*	Hires	Quits	Hires	Quits
Manufacturing	0.086	0.010	0.069	0.005
Utilities	0.011	0.022	0.035	0.029
Construction	0.037	0.015	0.054	0.021
Retail trade, cars	0.053	0.062	0.054	0.058
Transportation	0.004	0.044	0.023	0.045
Hotel and restaurant	-0.014	0.084	-0.011	0.079
Information and communication	0.121	0.106	0.059	0.048
Finance and insurance	0.089	0.081	0.075	0.072
Real estate	0.048	0.027	0.055	0.049
Prof./scientific/tech. services	0.066	0.094	0.069	0.086
Services	0.041	0.097	0.031	0.078
Public admin./education	-0.037	0.086	-0.020	0.083
Health and social	0.014	0.080	0.010	0.055

Note: This table reports wages and wage differentials by industry, using the sample of strongly connected firms (columns 2 and 4 of Table 1). The panel Δ *log-wage of employer-to-employer transitions* takes into account that wages at the old employer are observed in year t, and at the new employer in year $t+2$ by subtracting time and experience effects from the wage at the new employer using the estimates from my AKM-regression (see Appendix G.2). In the lower panel, transitions are weighted by their probability of being an excess separation as defined in the text in Section 2.

Table A.6: COMPARISON OF WAGE VARIANCE DECOMPOSITION WITH KLINE ET AL. (2020) AND BONHOMME ET AL. (2020)

	Own 1996–2003 (1)	Own 2004–2011 (2)	K. et al. (2020) (3)	B. et al (2020) (4)
var of log-wage†	0.190	0.193	0.184	0.182
share firm effect	0.073	0.065	0.130	0.129
share person effect	0.794	0.805	0.608	
share sorting	0.057	0.057	0.160	0.130
Corr(firm,person)	0.114	0.120	0.262	0.340

Notes: This table reports results from decomposing wage variance. The variance of log-wage is the variance net of time and experience effects, that is, var(log-wage net of time and experience) $= var(\alpha) + var(\psi) + 2 * cov(\alpha, \psi) + var(r)$. The person share is not reported by Bonhomme et al. (2020).
† After removing time/experience effects.

Table A.7: COVARIANCES OF JOB VALUE COMPONENTS 1996–2003

		Job value	Wage	Non-wage	Wage				Non-wage	
					Person	Employer	$\boldsymbol{X}'_{it}\beta$	r_{it}	Employer	Idio.
	Job value	0.524								
	Wage	0.227	0.195							
	Non-wage	0.297	0.032	0.265						
Wage	Person	0.193	0.156	0.037	0.154					
Wage	Employer	0.012	0.020	-0.008	0.006	0.014				
Wage	$\boldsymbol{X}'_{it}\beta$	0.007	0.005	0.002	-0.003	0	0.008			
Wage	r_{it}	0.015	0.015	0	0	0	0	0.015		
Non-wage	Employer	0.231	0.032	0.199	0.037	-0.008	0.002	0	0.199	
Non-wage	Idiosyncratic	0.066	0	0.066	0	0	0	0	0	0.066

Notes: This table reports covariances of job-value components in the 1996–2003 sample. The covariances are estimated using all person-year observations from Table 1 column 2, and the estimates on wage and non-wage value components from Section 3.3.

Table A.8: Covariances of Job Value Components 2004–2011

		Job value	Wage	Non-wage	Wage: Person	Wage: Employer	Wage: $X'_{it}\beta$	Wage: r_{it}	Non-wage: Employer	Non-wage: Idio.
	Job value	0.564								
	Wage	0.242	0.197							
	Non-wage	0.322	0.045	0.277						
Wage	Person	0.202	0.161	0.041	0.157					
Wage	Employer	0.021	0.018	0.003	0.006	0.013				
Wage	$X'_{it}\beta$	0.005	0.004	0.001	-0.001	0	0.006			
Wage	r_{it}	0.014	0.014	0	0	0	0	0.014		
Non-wage	Employer	0.250	0.045	0.205	0.041	0.003	0.001	0	0.205	
Non-wage	Idiosyncratic	0.071	0	0.071	0	0	0	0	0	0.071

Notes: This table reports covariances of job-value components in the 2004–2011 sample. The covariances are estimated using all person-year observations from Table 1 column 4, and the estimates on wage and non-wage value components from Section 3.3.

Table A.9: Comparison of Own Estimates with Hall and Mueller (2018) and Taber and Vejlin (2020)

	Own 1996–2003 (1)	Own 2004–2011 (2)	HM (2018) (3)	TV (2020) (4)
Var of log-wage	0.20	0.20	0.24	0.12
Share person	0.88	0.87	0.76	
Var of non-wage value	0.26	0.28	0.12	
Corr(non-wage value, wage)†	-0.12	0.05	-0.17	
Var of job value	0.52	0.56		0.25
Share non-wage value	0.57	0.57		0.51

Notes: This table reports results from decomposing job value variance. Column 3 uses values reported in Table 2 in Hall and Mueller (2018), applying the following calculation (using the notation of Hall and Mueller (2018)): *var of log-wage* $= \sigma_y^2 + \sigma_x^2$; *share person* $= \frac{\sigma_x^2}{(\sigma_y^2+\sigma_x^2)}$; *var of non-wage value* $= \sigma_\eta^2 + \kappa^2 * \sigma_y^2$; *Corr(non-wage value, wage)* $= \frac{\kappa*\sigma_y^2}{\sqrt{(\sigma_\eta^2+\kappa^2*\sigma_y^2)}*\sigma_y}$. I calculate the values in Column 1 and 2 using the estimates reported in Table A.7 and A.8, where *share person* $= \frac{\text{var(person)}}{\text{var(wage)-2*cov(person,firm)}}$, and *Corr(non-wage value, wage)* $= \frac{\text{cov(non-wage value, wage employer)}}{\sqrt{\text{var(non-wage value)*var(wage employer)}}}$. Column is based on Table 6 and 7 in Taber and Vejlin (2020), including the residual variation in wage estimated by Taber and Vejlin (2020) (.02) for consistency with my estimates.

† After removing personal-specific components and wage residual.

Table A.10: LABOR MARKET FRICTIONS

	Offer intensity (1)	Annual employment hazard (2)	Annual employer-to-employer transition hazard (3)
1996–2003 Panel	0.074	0.124	0.022
2004–2011 Panel	0.080	0.118	0.023

Note: This table reports evidence on the evolution of labor market frictions between 1996–2003 and 2004–2011. Offer intensity measured as total number of offers by firms in sample (estimated from hires from non-employment), divided by the number of people-years in sample. Columns (2) and (3) show annual hazard rates from Table A.2.

Table A.11: LABOR MARKET SEGREGATION

	Difference in firm value offer							
	Accepted offers				Estimated offers			
	rescaled | Mean |	rescaled SD	| Mean |	SD	rescaled | Mean |	rescaled SD	| Mean |	SD
1996–2003	0.825	0.964	0.373	0.436	0.946	1.207	0.428	0.546
2004–2011	0.866	1.048	0.415	0.502	0.966	1.239	0.463	0.594

Note: This table reports summary statistics on the degree of labor market segregation in 1996–2003 and in 2004–2011. *Accepted Offers* shows the distribution of the absolute value of the firm value difference between the firm making the offer and the firm receiving the offer, among all employer-to-employer transitions (Panel D of Table 1). The columns *Estimated offers* show the distribution of the absolute value of the firm value difference between the firm making the offer and the firm receiving the offer, among all offers made, estimated as described in Appendix K. The prefix *rescaled* indicates that the value is divided by the standard deviation of total firm value, calculated as the standard deviation of firm value among firms, weighted by the number of people-years observations, in the respective sample period. The conclusion that segregation did not contribute to a decline in labor supply elasticity is based on the lack of a substantial increase in the measures of value difference in offers that are corrected for overall value dispersion in the respective panel.

Table A.12: TEST OF PATTERNS IMPLIED BY A DECREASE IN LABOR SUPPLY ELASTICITY

	Ln(employees)		Δ ln(employees) 94-03 to 03-12		Δ Value 94-03 to 03-12	
	(1) 96-03	(2) 03-11	(3) employers	(4) cells	(5) employers	(6) cells
Firm value 1996-2003	0.927***		-0.088***	-0.010		
	(0.045)		(0.020)	(0.214)		
Firm value 2004-2011		0.603***				
		(0.038)				
Ln(employees) 1996-2003					-0.029***	-0.011
					(0.006)	(0.021)
Person-years	4,513,833	5,480,901	7,239,585	9,994,734	7,239,585	9,994,734
Industry-Federal state cells				82		82
Firms	4,544	5,944	2,495	7 993	2,495	7 993
Std indep. var.	0.452	0.479	0.447	0.259	1.367	0.997

Notes: This table reports coefficients from bivariate regressions of the variable in the column header on the variable in the row. Column 1: Regression of the log of the average number of yearly employees 1996-2003 on the firm value 1996-2003. Column 2 same as column 1 for 2004-2011. Columns 3 and 4: Regression of the change in log-average number of yearly employees 1996-2003 to 2004-2011 on the firm value 1996-2003. Columns 5 and 6: Regression of the change in firm value 1996-2003 to 2004-2011 on the log-average number of employees 1996-2003. The observational unit in columns 3 and 5 is the firm, limiting the sample to firms belonging to the sample in the 1996–2003 and the 2004–2011 period. Columns 4 and 6 are based on a repeated cross section with firms aggregated on the federal state by industry level. Both regressions are weighted using the sum of the number of people-years in both periods as analytical weights.
* p<0.05, ** p<0.01, *** p<0.001

Table A.13: FIRM VALUE AND Δ SHARE FOREIGN NATIONALS

	Δ share foreigners 1996-2003 to 2004-2011		Δ firm value 1996-2003 to 2004-2011	
	(1)	(2)	(3)	(4)
Firm value 1996-2003	-0.015***	-0.026*		
	(0.002)	(0.014)		
Δ share foreigners 1996-2003 to 2004-2011			-0.552***	-1.563**
			(0.188)	(0.611)
Person-years	7,239,585	9,994,734	7,239,585	9,994,734
Industry-Federal state cells		82		82
Firms	2,495	7,993	2,495	7,993
Share foreigners 1996–2003	0.101	0.106	0.101	0.106
Share foreigners 2004–2011	0.132	0.140	0.132	0.140

Notes: This table reports regression results on the relationship between the change in workforce with foreign nationality and firm value. Columns 1 and 2 show results from regressing the change in the share of the workforce with foreign nationality on the firm value 1996–2003. Columns 3 and 4 show results from regressing the change in firm value 1996–2003 to 2004–2011 on the change in the share of the workforce with foreign nationality. Nationality is measured at labor market entry. The observational unit in columns 1 and 3 is the firm, limiting the sample to firms belonging to the sample in the 1996–2003 and the 2004–2011 period. Columns 2 and 4 are based on a repeated cross section with firms aggregated on the federal state by industry level All regressions are weighted using the sum of the number of people-years in both periods as analytical weights.
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table A.14: RESULTS FROM ESTIMATING THE MODEL UNDER ALTERNATIVE OFFER INTENSITIES

	1996 – 2003			**2004–2011**		
	Non-emp. hires	**All hires**	**Constant**	**Non-emp. hires**	**All hires**	**Constant**
A. Summary stats on offers						
Mean # of offers per firm	60	119		59	123	
Std # of offers per firm	134	236		132	263	
Corr(Non-employment hires, all hires)	0.94			0.87		
B. Model results						
Correlation firm wage, firm non-wage	-0.12	-0.13	-0.12	0.05	0.03	0.03
Correlation person wage, firm non-wage	0.21	0.19	0.18	0.22	0.21	0.20

Note: Panel A of this table reports summary statistics on offers made by firms to other firms' employees, when offers are estimated under two different assumptions on the process generating them. *Non-emp. Hires* refers to the baseline approach using all hires from non-employment to estimate firms' offer intensity. *All Hires* refers to the alternative approach of using all workers hired in the corresponding sample period to estimate firms' offer intensity. *Constant* keeps offer distribution and employer size constant across employers. Panel B reports the correlation of job value components obtained when estimating the model using the offer distribution indicated in the corresponding column.

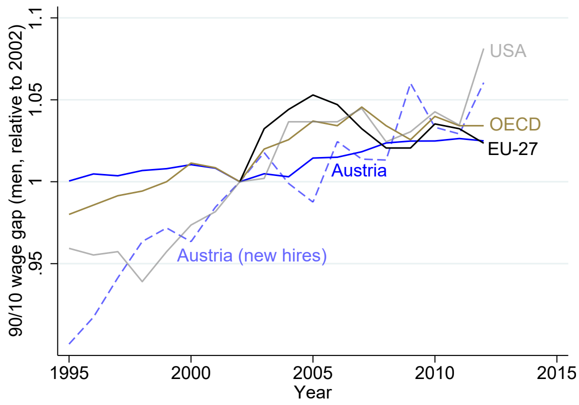
Table A.15: ROBUSTNESS – PREFERENCE HETEROGENEITY AND ASYMMETRIC LEARNING

	Estimate	Strongly connected	Transitions (Restricted)	Transitions (All)
Panel A: Preference Heterogeneity – Low-Wage vs. High-Wage Workers				
1996–2003				
Using low-wage workers' SC set				
Corr. low-wage's $ln(a)$, all's $ln(a)$	0.81	1,303	11,986	21,548
Corr. high-wage's $ln(a)$, all's $ln(a)$	0.89	1,792	25,697	36,135
Offer distributions				
Corr. low-wage's, high-wage's offer dist.	0.48			
2004–2011				
Using low-wage workers' SC set				
Corr. low-wage's $ln(a)$, all's $ln(a)$	0.87	2,092	18,775	34,660
Using high-wage workers' SC set				
Corr. high-wage's $ln(a)$, all's $ln(a)$	0.90	2,449	33,422	45,986
Offer distributions				
Corr. low-wage's, high-wage's offer dist.	0.63			
Panel B: Asymmetric Learning – Young vs. Old Workers				
1996–2003				
Using young workers' SC set				
Corr. young's $ln(a)$, all's $ln(a)$	0.95	2,908	30,316	46,417
Offer distributions				
Corr. young's, old's offer dist.	0.83			
2004–2011				
Using young workers' SC set				
Corr. young's $ln(a)$, all's $ln(a)$	0.95	4,012	42,841	60,344
Offer distributions				
Corr. young's, old's offer dist.	0.83			

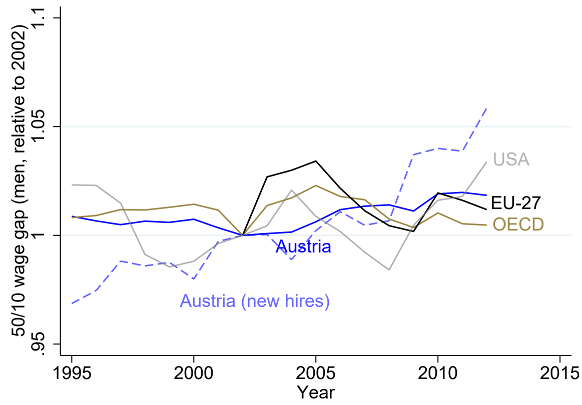
Note: This table reports the correlation between firms' non-wage values from the model estimated using the restricted sample of workers and from the model estimated using the full sample of workers (weighted by the number of person-year observations), on the subsample of firms strongly connected by employer-to-employer transitions of workers from restricted sample. The statistic reported on *Offer distributions* is the correlation between the number of offers all firms in sample (columns 2 and 4 of Table 1) make to the two subgroups of Panel A and B of this table. The samples are split by median age/wage. Median age 1996–2003: 38.04; Median age 2004–2011: 40.30; Median monthly wage (2012 €) 1996–2003: 3048.13; Median monthly wage (2012 €) 2004–2011: 3195.62. Example of how to read the table: The 1st row *Corr. low-wage's $ln(a)$, all's $ln(a)$* shows that the firm non-wage value estimates estimated using transitions of low-wage workers, and the set of firms strongly connected by at least 5 transitions of low-wage workers, is .81 correlated with non-wage values estimated on the set of firms strongly connected by at least 5 transitions of low-wage workers, but using transitions of low-wage & high-wage workers. The 3rd row *Corr. low-wage's, high-wage's offer dist.* shows that the distribution from which high-wage workers are estimated to receive offers is .048 correlated with the distribution from which low-wage workers are estimated to receive offers (estimated using firms' hires from non-employment).

Figure A.1: 90-10 AND 50-10 WAGE GAP 1995–2012

(a) 90-10 Gap



(b) 50-10 Gap

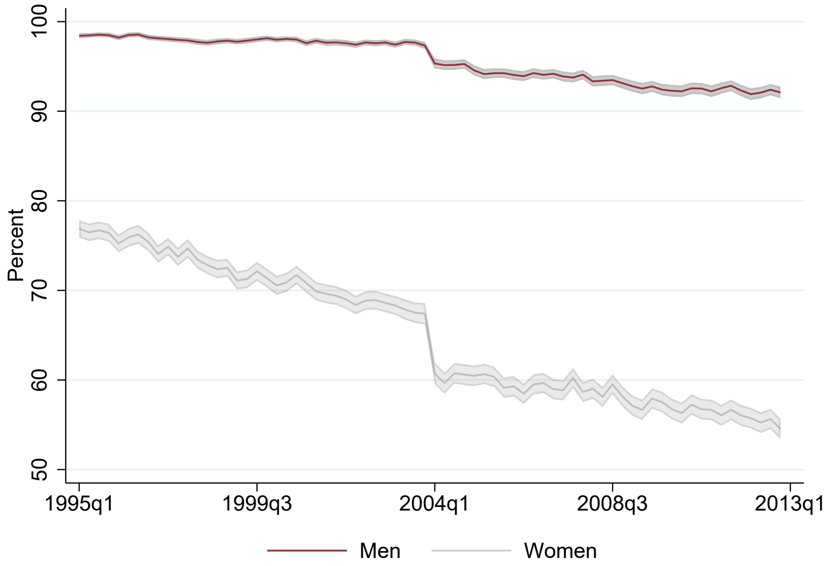


Notes: Figure a shows the gap between the 90th and the 10th percentile of the wage distribution in a given year for full-time working men. Figure b shows the gap between the 50th and the 10th percentile of the wage distribution in a given year for full-time working men. The dashed line reports the gap among workers that were hired in the previous year. The gap is reported relative to the gap in year 2002.
Source: Austria: Own calculations; USA, EU-27 and OECD: OECD (2013).

Figure A.2: SHARE MEN EMPLOYED FULL-TIME 1995–2012

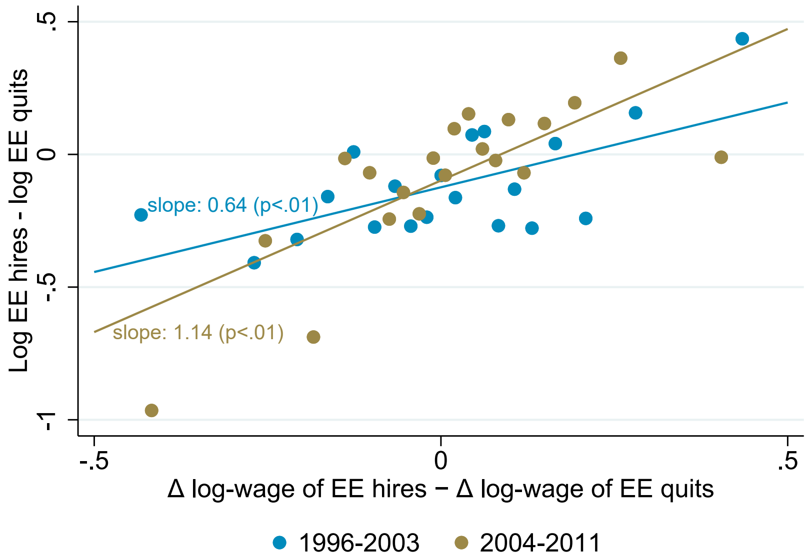
Notes: This figure shows the evolution of the share of all male employees that is working full-time. It is based on data from the Austrian Mikrozensus (Austrian labor force survey). I classify a worker as full-time employed if he reports working at least 36 hours in a normal work week. The discontinuity in year 2004 is due to a reform of the Mikrozensus, which included a change in the definition of employment status (Lehmann, 2019).

Figure A.3: SHARE MEN AND WOMEN EMPLOYED FULL-TIME 1995–2012



Notes: This figure shows the evolution of the share of all employees that is working full-time separately for men and women. It is based on data from the Austrian Mikrozensus (Austrian labor force survey). I classify a worker as full-time employed if he reports working at least 36 hours in a normal work week. The discontinuity in year 2004 is due to a reform of the Mikrozensus, which included a change in the definition of employment status (Lehmann, 2019).

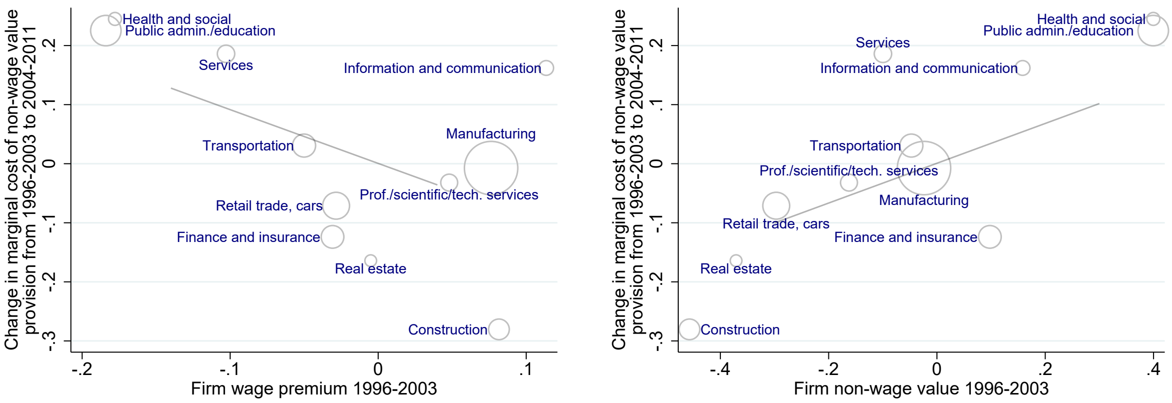
Figure A.4: Descriptive Evidence on Relationship Between Employer Attractiveness and Employer Wage Premium from 1996–2003 to 2004–2011



Notes: This figure reports the relationship between the log of $\frac{\text{employer-to-employer hires}}{\text{employer-to-employer quits}}$ and the log of $\frac{\text{average wage increase of employer-to-employer hires}}{\text{average wage increase of employer-to-employer quits}}$. Firms for 1996–2003 (column 2 of Table 1) and 2004–2011 (column 4 of Table 1) are separately grouped into 20 firm-size (measured by the number of people-years) weighted bins by Δ log-wage of EE hires $-\Delta$ log-wage of EE quits. The regression lines represent the slope of an OLS regression, with firms weighted by the number of person-year observations they represent in the corresponding sample period. See Figure A.12 for version considering all workers and firms (columns 1 and 3 of Table 1.)

Figure A.5: CHANGE IN MARGINAL COST OF NON-WAGE VALUE PROVISION AND FIRM WAGE AND NON-WAGE VALUE

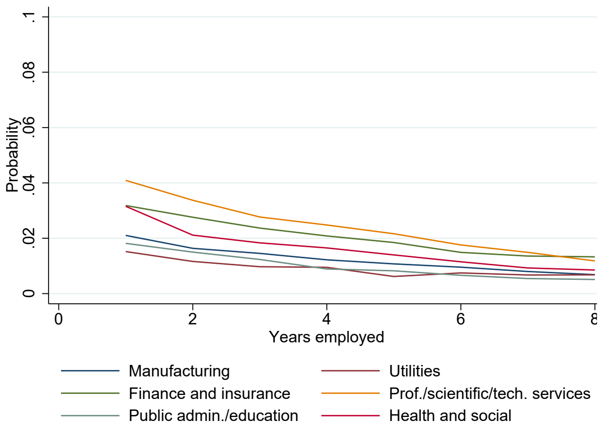
(a) Firm wage (b) Firm non-wage value



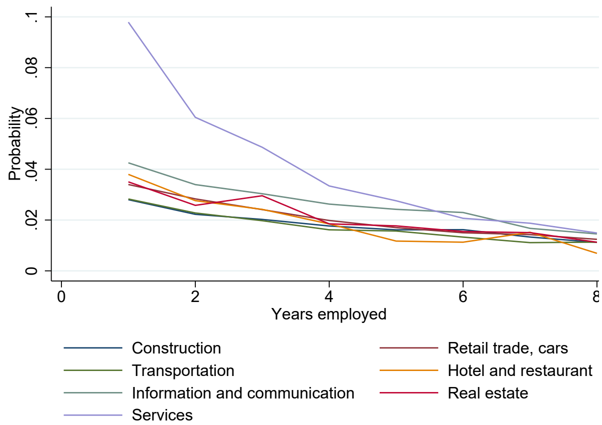
Notes: These figures show the difference between the industry-level average marginal cost of non-wage value provision for firms in the 2004–2011 panel and in the 1996–2003 panel, as a function of the firm wage premium and firm non-wage value in the 1996–2003 panel. The figure shows industry-level averages, which are calculated with firms weighted by the number of person-year observations they represent. The regression line represents a linear regression run at the industry-level, with industries weighted by the total number of person-year observations they represent in the 1996–2003 and the 2004–2011 panel. Marginal cost of non-wage value provision are derived as explained in Appendix J.6.

Figure A.6: Annual Probability of Employer-to-Employer Transition by Tenure and Industry, 1996–2011

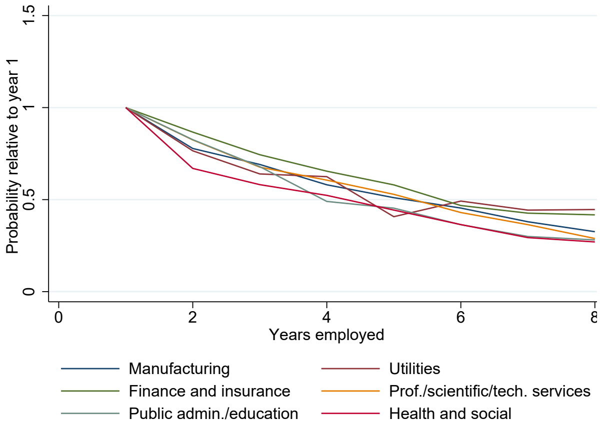
(a) High Non-Wage Industries, absolute



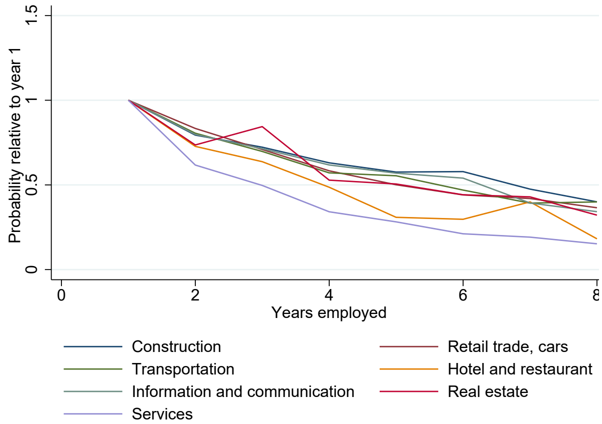
(b) Low Non-Wage Industries, absolute



(c) High Non-Wage Industries, relative



(d) Low Non-Wage Industries, relative



Notes: These figures plot the annual probability a worker in a given industry makes an employer-to-employer transition satisfying the criteria defined in Section 2. Figures a and b show the absolute probabilities, and figures c and d show the probabilities relative to the first year of tenure. Figures a and c show the six industries where firms on average offer the highest non-wage value, and figures b and d show the seven industries where firms in average offer the lowest non-wage value.

B Identification of Employer-Level Dynamics

The following exposition follows Fink et al. (2010), with some adjustment tailored to my aim to identify all employer-level dynamics that do not follow a worker's binary choice.

I start by creating a quarterly panel of person-employer employment. I then identify the following employer level dynamics by applying the following criteria:

Rename: I classify an employer rename from A to B if

- Employer-identifier A exists in quarter t but not in quarter $t + 1$
- Employer-identifier B does not exist in quarter t and exists in quarter $t + 1$
- At least two third of individuals employed at employer A in quarter t are employed at employer B in quarter $t + 1$
- Employer A has at least 3 employees in quarter t

Takeover: I classify an employer takeover from A to B if

- Employer-identifier A exists in quarter t but not in quarter $t + 1$
- Employer-identifier B exists in quarter t and exists in quarter $t + 1$
- At least two third of individuals employed at employer A in quarter t are employed at employer B in quarter $t + 1$
- Employer A has at least 3 employees in quater t

Spin-off: I classify an employer spin-off from A to B if

- Employer-identifier A exists in quarter t and exists in quarter $t + 1$
- Employer-identifier B does not exist in quarter t and exists in quarter $t + 1$
- At least 10 percent of employees and at least three employees working at employer A in quarter t work at employer B in quarter $t + 1$

Closure: I classify an employer closure of employer A if

- Employer-identifier A exists in quarter t but not in quarter $t + 1$
- there is no rename or takeover

I merge employer-identifiers in case of a rename. If there is a takeover of employer A by employer B in quarter t I drop all transitions between employer A and employer B in quarter t and $t - 1$. If there is a spin-off from employer A to employer B in quarter t I drop all transitions between employer A and employer B in quarter t and $t + 1$. If there is an employer closure at employer A in quarter t I drop all transitions away from employer A in quarter t and $t - 1$.[70] Table A.2 shows the number of transitions caused by the respective employer-level dynamics.

[70]By including adjacent quarters I account for the fact that employer-level transitions might not affect all workers at the exactly same point in time.

C Tables and Figures for 1996–2003 for Section 2 and Section 3.3

Table A.16: BY INDUSTRY – NUMBER OF HIRES AND QUITS
1996 – 2003

	Unweighted		Layoff weighted	
	Hires	Quits	Hires	Quits
	(1)	(2)	(3)	(4)
Manufacturing	17,031	16,293	8,366	7,958
Utilities	1,201	919	444	300
Construction	4,762	4,207	1,563	1,629
Retail trade, cars	6,446	7,006	3,076	3,276
Transportation	3,774	4,029	1,846	2,094
Hotel and restaurant	299	309	122	154
Information and communication	3,494	2,523	1,824	1,454
Finance and insurance	6,042	6,141	2,470	2,424
Real estate	1,512	1,631	672	786
Prof./scientific/tech. services	3,651	4,460	1,587	1,766
Services	2,996	4,113	1,531	2,246
Public admin./education	5,204	4,152	2,538	2,027
Health and social	1,522	2,403	741	725

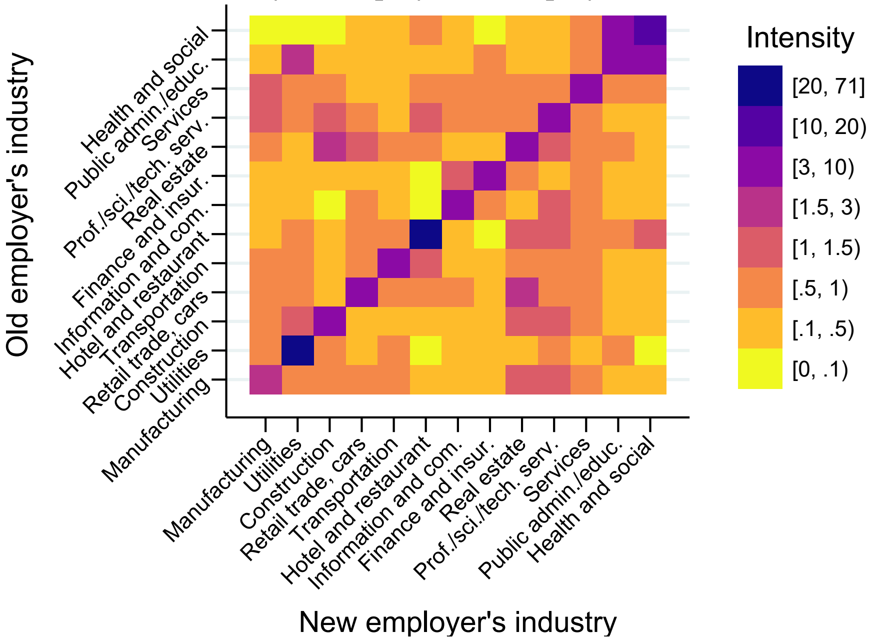
Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for the 1996–2003 sample. Columns 1 and 2: Number of hires and quits by industry. Columns 3 and 4: Number of hires and quits by industry, after downweighting quits from contracting firms according to procedure explained in Section 2.

Figure A.7: EMPLOYMENT GROWTH AND TRANSITION PROBABILITIES 1996–2003

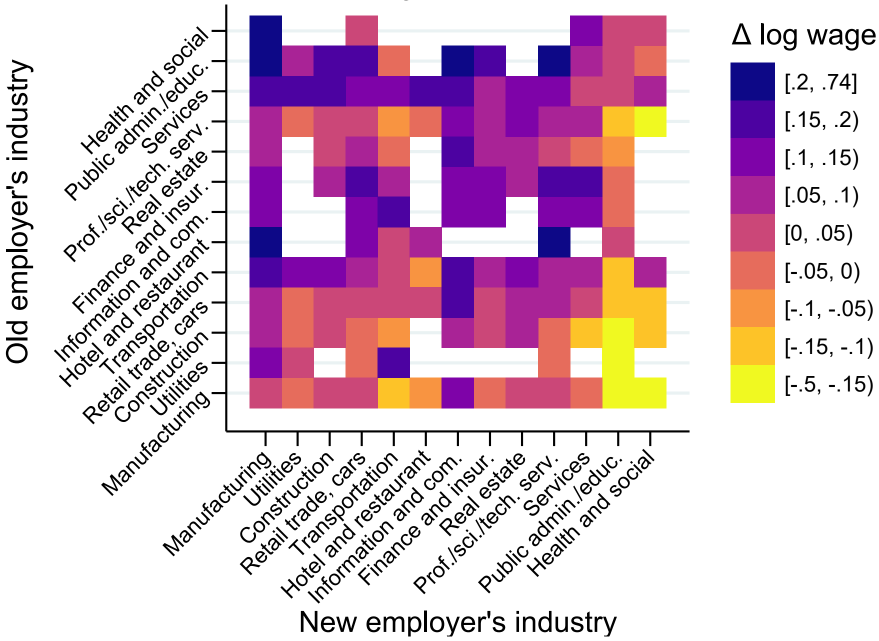
Notes: This figure shows the probability (per year) a worker in column 1 of Table 1 makes a transition, by 0.05 employer growth rate bin. Full-time employer-to-employer corresponds to the employer-to-employer transitions as defined in this section. Other employer-to-employer correspond to all transitions in which the worker starts at the new employer within 30 days, but do otherwise not satisfy the conditions detailed in this section. Employer-to-nonemployment are employment spells ending in year $t+1$ for which the worker does not join a new employer within 30 days. Share excess transitions as $\frac{\text{excess}}{\text{excess+expected}}$. Corresponding figure for 2004–2011 in Figure 2.

Figure A.8: EMPLOYER-TO-EMPLOYER TRANSITIONS AND WAGE DIFFERENTIALS 1996–2003

(a) Intensity of Employer-to-Employer Transitions



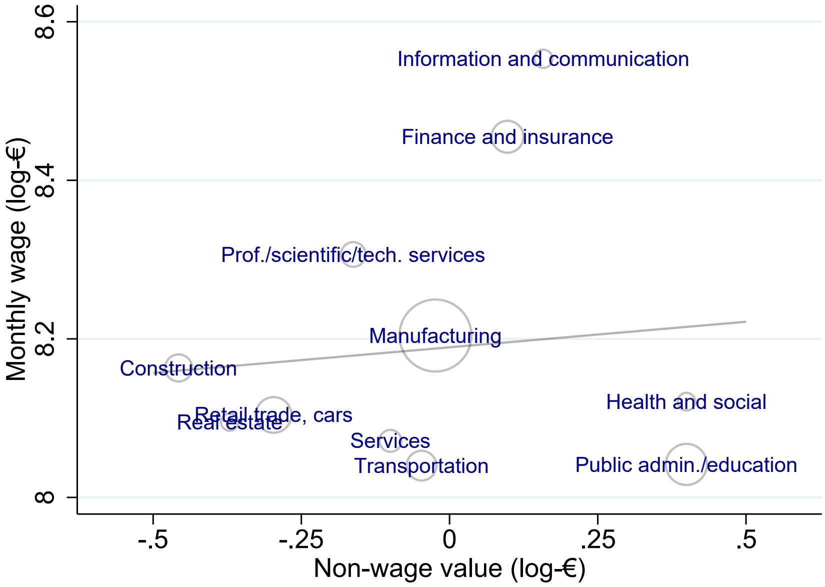
(b) Wage Differentials



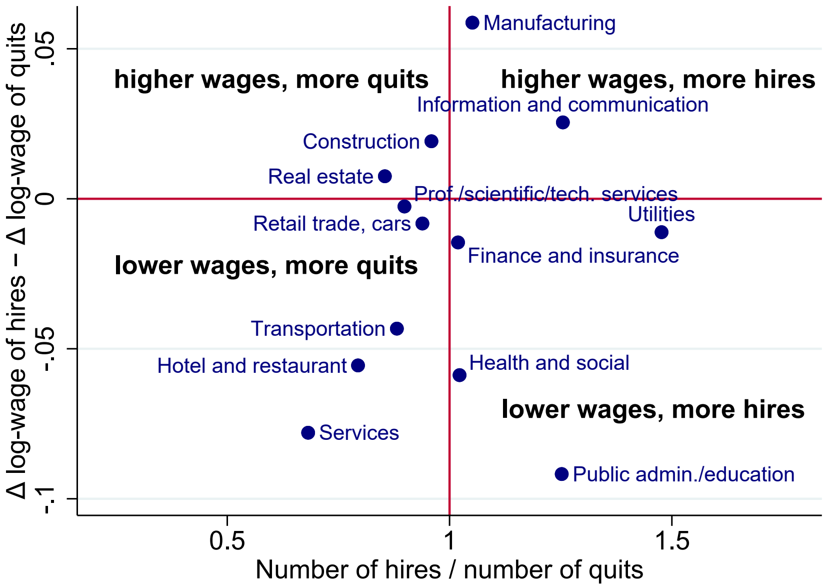
Notes: Figure a shows the intensity of employer-to-employer transitions between industries 1996–2003. If mobility was random, the intensity would be equal to 1 for each cell. Intensities above 1 indicates that there are more transitions from the row-industry to the column-industry than expected under random mobility. See text in Section 2 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 employer in the row-industry in the new employer in the column-industry. Missing cells in figure b contain fewer than 10 observations. Both figures are based on transitions between employers in the strongly connected 1996–2003 sample (column 2 of Table 1). See Figure A.11 for employer-to-employer transitions of all workers.

Figure A.9: Hires, Quits, Wage Differentials and Non-Wage Values 1996–2003

(a) Non-Wage Values



(b) Employer-to-Employer Transitions and Wages



Notes: Figure a shows non-wage values and log-wages by industry, with circle size relative to the number of person-year observations in the corresponding industry. The gray line plots the regression line run at the industry level, with industries weighted by their number of person-year observations. Two industries are not shown in figure a: Utilities (coordinates: (.47,8.41)) and Hotel and restaurant (-.65,7.82). Figure b shows, on the x-axis, the number of employer-to-employer hires divided by the number of employer-to-employer quits (based on columns 3 and 4 of Table A.16), and on the y-axis: Log-wage increase of employer-to-employer hires minus log-wage increase of employer-to-employer quits (corrected for time/experience effects, based on Table A.5). Figures are based on the 1996–2003 sample (column 4 of Table 1). See Figure 4 for the 2004–2011 sample.

D Geralizability to Whole Austrian Labor Market

There are three restrictions I impose that distinct my sample from the whole population of workers in Austria, and that might raise concerns related to external validity:

1. My sample is limited to firms in the strongly connected set
2. My sample is limited to men
3. My sample is limited to full-time workers

In following, I will address 1.-3., showing that neither of the restrictions is likely to limit my sample in a way that affects external validity of my results with respect to the entire Austrian labor market.

D.1 Restriction 1: Only Strongly Connected Employers

In following I discuss descriptive statistics comparing the composition and dynamics in my sample of strongly connected employers (columns 2 and 4 of Table 1) with the composition and dynamics in the sample considering all employers (columns 1 and 3 of Table 1).

From Table 1 we see a comparison of my samples of strongly connected employers with all full-time working men. The largest differences are that men in my sample earn higher wages, and are more likely to work in manufacturing jobs. While these differences are substantial, two things are important to note. First, the difference in composition between my sample and all workers is about constant between 1996–2003 and 2004–2011. Second, dynamics of employer-to-employer transitions seem to be very similar in my sample as among all workers, as we can see from Panel D. of Table 1.

My estimator uses three pieces of information to identify non-wage values of firms: (1) the number of employer-to-employer hires of a firm compared to the number of employer-to-employer quits of a firm, (2) the pattern of these employer-to-employer hires and quits, and (3) the wage differentials associated with these employer-to-employer hires and quits. In following, I will thus compare descriptive statistics on the industry-level on (1)-(3) between all firms and my sample of strongly connected firms.

Tables A.17 and A.18 show the number of employer-to-employer hires and quits by industry in the Austrian labor market overall (columns 1 and 2) and in my sample (columns 3 and 4). Up to very few exceptions, industries where firms in my sample hire more workers than they loose workers are also industries that hire more workers than they loose workers if all firms are considered. Figures A.10 a & b and A.11 a & b show that also the pattern of worker flows between industries is similar in the Austrian labor market overall and in my sample of strongly connected firms.

Figures A.10 c & d and A.11 c & d and Tables A.19 and A.20 show wage differentials associated with employer-to-employer transitions. We see that disparities between my sample and all workers in terms of wage differentials are remarkably small.

The main result of this paper, that the inequality-attenuating effect of compensating differentials in the 1996–2003 panel was dominated by firm-level rents in the 2004–2011 panel, is driven by underlying changes in the pattern of worker flows and associated wage differentials. Figure A.4 shows how this pattern changed from 1996–2003 to 2004–2011 in my sample. Figure A.12 shows that this pattern changed in a very similar way in the Austrian labor market

Table A.17: BY INDUSTRY – NUMBER OF HIRES AND QUITS 1996 – 2003

	Overall		Strongly Connected	
	Hires (1)	Quits (2)	Hires (3)	Quits (4)
Manufacturing	20,268	20,590	8,366	7,958
Utilities	1,372	860	444	300
Construction	6,006	7,417	1,563	1,629
Retail trade, cars	15,215	16,339	3,076	3,276
Transportation	6,042	6,357	1,846	2,094
Hotel and restaurant	1,414	1,890	122	154
Information and communication	4,203	3,372	1,824	1,454
Finance and insurance	4,903	4,671	2,470	2,424
Real estate	2,431	2,307	672	786
Prof./scientific/tech. services	5,198	5,754	1,587	1,766
Services	4,706	5,378	1,531	2,246
Public admin./education	6,886	4,118	2,538	2,027
Health and social	1,920	1,788	741	725

Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for 1996–2003. Columns 1 and 2: Number of hires and quits in Austrian labor market overall . Columns 3 and 4: Number of hires and quits by industry between firms in the sample of strongly connected firms. All statistics after downweighting quits from contracting firms according to procedure explained in Section 2.

overall, indicating that the result that compensating differentials in 1996–2003 got dominated by firm-level rents in 2004–2011 is not limited to my sample, but holds for the Austrian labor market overall.

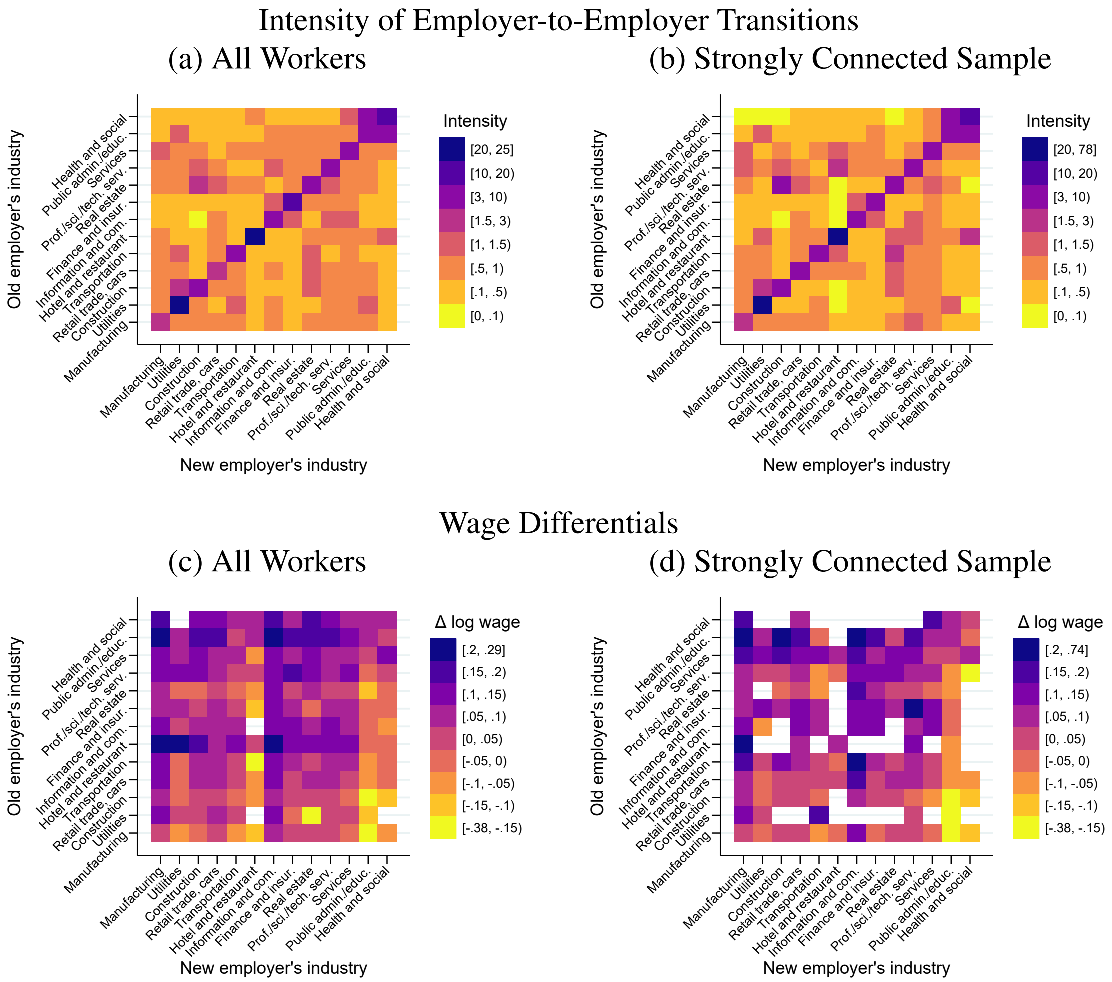
Overall, the descriptive evidence let me conclude that the sample of strongly connected firms does not differ from the entire Austrian labor market in terms of structure and dynamics in a way that would substantially affect external validity of my results.

Table A.18: BY INDUSTRY – NUMBER OF HIRES AND QUITS 2004 – 2011

	Overall		Strongly Connected	
	Hires	Quits	Hires	Quits
	(1)	(2)	(3)	(4)
Manufacturing	23,015	21,175	12,659	10,256
Utilities	2,246	1,281	1,034	580
Construction	7,467	8,482	2,365	2,553
Retail trade, cars	15,987	16,339	4,193	4,425
Transportation	7,027	7,219	3,049	3,120
Hotel and restaurant	1,529	1,974	137	178
Information and communication	4,407	4,298	2,106	2,083
Finance and insurance	5,449	4,864	3,050	2,739
Real estate	2,091	2,004	601	686
Prof./scientific/tech. services	6,771	6,625	2,529	2,220
Services	6,304	11,035	2,600	6,434
Public admin./education	7,997	5,370	4,025	3,140
Health and social	2,050	1,894	840	806

Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for 2004–2011. Columns 1 and 2: Number of hires and quits in Austrian labor market overall . Columns 3 and 4: Number of hires and quits by industry between firms in the sample of strongly connected firms. All statistics after downweighting quits from contracting firms according to procedure explained in Section 2.

Figure A.10: EMPLOYER-TO-EMPLOYER TRANSITIONS AND WAGE DIFFERENTIALS 1996–2003



Notes: Figure a shows the intensity of employer-to-employer transitions between all firms. Figure b shows the intensity of employer-to-employer transitions between firms in the strongly connected sample. If mobility was random, the intensity would be equal to 1 for each cell. Intensities above 1 indicates that there are more transitions from row-industry to column-industry than expected under random mobility. See text in Section 2 for formal definition of intensity. Figure c shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions between all firms with old employer in industry in row and new employer in industry in column. Figure d shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions between firms in the strongly connected sample with old employer in industry in row and new employer in industry in column. Missing cells in figures c and d contain fewer than 10 observations. These statistics are based on employer-to-employer transitions described in Table 1 (Figure a and c: Column 1; Figure b and d: Column 2).

Figure A.11: EMPLOYER-TO-EMPLOYER TRANSITIONS AND WAGE DIFFERENTIALS 2004–2011

Notes: Figure a shows the intensity of employer-to-employer transitions between all firms. Figure b shows the intensity of employer-to-employer transitions between firms in the strongly connected sample. If mobility was random, the intensity would be equal to 1 for each cell. Intensities above 1 indicates that there are more transitions from row-industry to column-industry than expected under random mobility. See text in Section 2 for formal definition of intensity. Figure c shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions between all firms with old employer in industry in row and new employer in industry in column. Figure d shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions between firms in the strongly connected sample with old employer in industry in row and new employer in industry in column. Missing cells in figures c and d contain fewer than 10 observations. These statistics are based on employer-to-employer transitions described in Table 1 (Figure a and c: Column 1; Figure b and 2: Column 2).

Table A.19: BY INDUSTRY – WAGES AND WAGE DIFFERENTIALS 1996 – 2003

	All		Strongly connected	
	(1)		(2)	
Median monthly wages by industry (2012 €)				
Manufacturing	3,365		3,674	
Utilities	3,432		4,023	
Construction	3,083		3,532	
Retail trade, cars	3,122		3,369	
Transportation	2,728		2,896	
Hotel and restaurant	2,187		2,414	
Information and communication	4,914		5,302	
Finance and insurance	4,506		4,855	
Real estate	3,233		3,516	
Prof./scientific/tech. services	3,783		3,972	
Services	2,909		3,015	
Public admin./education	2,780		2,972	
Health and social	2,927		3,004	
Δ *log-wage of employer-to-employer transitions*				
	Hires	Quits	Hires	Quits
Manufacturing	0.086	0.010	0.074	0.015
Utilities	0.011	0.022	0.014	0.026
Construction	0.037	0.015	0.025	0.006
Retail trade, cars	0.053	0.062	0.052	0.060
Transportation	0.004	0.044	0.019	0.062
Hotel and restaurant	-0.014	0.084	0.031	0.086
Information and communication	0.121	0.106	0.144	0.118
Finance and insurance	0.089	0.081	0.089	0.103
Real estate	0.048	0.027	0.038	0.030
Prof./scientific/tech. services	0.066	0.094	0.062	0.064
Services	0.041	0.097	0.031	0.109
Public admin./education	-0.037	0.086	-0.015	0.076
Health and social	0.014	0.080	-0.003	0.056

Note: This table reports wages and wage differentials by industry. Column 1 considers all workers according to column 1 of Table 1. Column 2 restricts the sample to workers at employers strongly connected by employer-to-employer transitions (column 2 in Table 1). The panel Δ *log-wage of employer-to-employer transitions* takes into account that wages at the old employer are observed in year t, and at the new employer in year $t+2$ by subtracting time and experience effects from the wage at the new employer using the estimates from my AKM-regression (see Appendix G.2). In the lower panel, transitions are weighted by their probability of being an excess separation as defined in the text in Section 2.

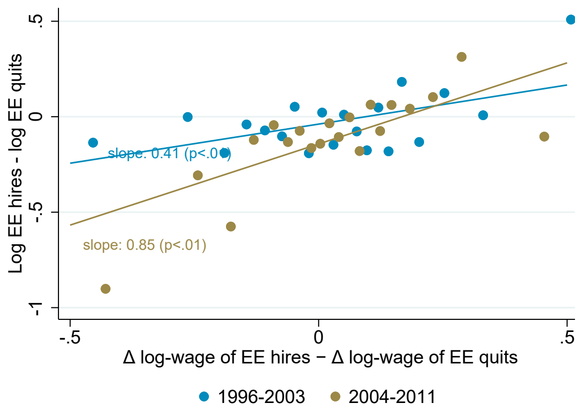
Table A.20: BY INDUSTRY – WAGES AND WAGE DIFFERENTIALS 2004 – 2011

	All		Strongly connected	
	(1)		(2)	
Median monthly wages by industry (2012 €)				
Manufacturing	3,569		3,781	
Utilities	4,103		5,294	
Construction	3,194		3,768	
Retail trade, cars	3,262		3,463	
Transportation	2,837		3,078	
Hotel and restaurant	2,240		2,455	
Information and communication	4,561		4,728	
Finance and insurance	4,900		5,188	
Real estate	3,450		3,405	
Prof./scientific/tech. services	4,244		4,586	
Services	2,938		3,004	
Public admin./education	3,081		3,516	
Health and social	3,163		3,493	
Δ log-wage of employer-to-employer transitions				
	Hires	Quits	Hires	Quits
Manufacturing	0.069	0.005	0.061	0.004
Utilities	0.035	0.029	0.042	0.040
Construction	0.054	0.021	0.056	0.013
Retail trade, cars	0.054	0.058	0.057	0.060
Transportation	0.023	0.045	0.034	0.048
Hotel and restaurant	-0.011	0.079	-0.025	0.070
Information and communication	0.059	0.048	0.054	0.046
Finance and insurance	0.075	0.072	0.083	0.090
Real estate	0.055	0.049	0.042	0.035
Prof./scientific/tech. services	0.069	0.086	0.073	0.078
Services	0.031	0.078	0.018	0.078
Public admin./education	-0.020	0.083	-0.002	0.092
Health and social	0.010	0.055	-0.005	0.048

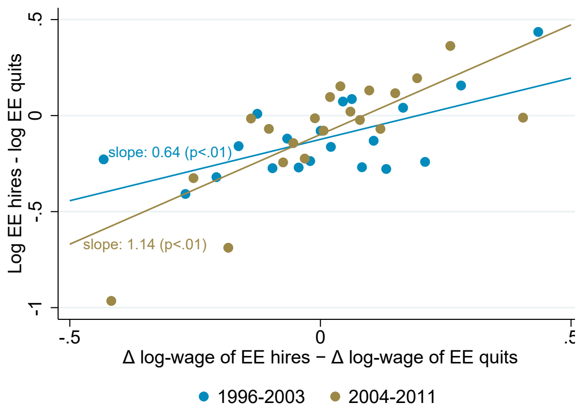
Note: This table reports wages and wage differentials by industry. Column 1 considers all workers according to column 3 of Table 1. Column 2 restricts the sample to workers at employers strongly connected by employer-to-employer transitions (column 4 of Table 1). The panel Δ *log-wage of employer-to-employer transitions* takes into account that wages at the old employer are observed in year t, and at the new employer in year $t+2$ by subtracting time and experience effects from the wage at the new employer using the estimates from my AKM-regression (see Appendix G.2). In the lower panel, transitions are weighted by their probability of being an excess separation as defined in the text in Section 2.

Figure A.12: DESCRIPTIVE EVIDENCE ON RELATIONSHIP BETWEEN EMPLOYER ATTRACTIVENESS AND EMPLOYER WAGE PREMIUM FROM 1996–2003 TO 2004–2011

(a) All Workers



(b) Strongly Connected Sample



Notes: These figures show the relationship between the log of $\frac{\text{employer-to-employer hires}}{\text{employer-to-employer quits}}$ and the log of $\frac{\text{average wage increase of employer-to-employer hires}}{\text{average wage increase of employer-to-employer quits}}$. Firms for 1996–2003 (Figure a: Column 1 of Table 1; Figure b: Column 2 of Table 1) and 2004–2011 (Figure a: Column 3 of Table 1; Figure b: Column 4 of Table 1) separately grouped into 20 firm-size (measured by the number of people-years) weighted bins, grouped by the x-axis variable. The regression lines represent the slope of an OLS regression, with firms weighted by the number of person-year observations they represent in the corresponding sample period.

D.2 Restriction 2: Only Male Workers

I restrict my sample to male workers because a large share of female workers are working part-time (Figure A.3), and I can only identify part-time workers after 2002 in my data. Thus, while I cannot make any statement on full-time working women in the 1996–2003 panel, I can do so for the 2004–2011 panel.[71]

Ideally, I would like to separately estimate my model for men and women, and compare the outcomes. This is not possible, however, because different firms are strongly connected in the sample of women and men (recall that firms' non-wage values are only identified within the strongly connected set). I can, however, estimate my model using transitions of men only (as I do in the baseline specification), and reestimate the model including transitions of women and compare the results. The underlying idea is that if preferences of men and women over firms and their non-wage characteristics were very different, then I should also obtain very different non-wage value estimates using the sample with and without women. I will thus in following evaluate how including women changes the composition of the sample in the 2004–2011 panel, and how this affects non-wage value estimates.

Table A.21 compares my baseline sample of male workers in 2004–2011 with the sample including full-time working women. Differences arise in that women earn less than men, and that women are more likely to work in traditional female-dominated industries such as public administration/education and health and social services.

Panel D. of Table A.26 shows that regarding transitions and associated wage differentials, my baseline sample (column 4) looks similar to the sample including full-time working women (column 2). This is also confirmed when comparing industry-level descriptive statistics on the number of hires and quits (Table A.22), the pattern of transitions (Figure A.13), and the wage differentials associated with employer-to-employer transitions (Table A.23).

I estimate the model described in Section 3 using employer-to-employer transitions of women (column 2 of Table A.21), and conduct the same job value variance decomposition as I do for the baseline sample (Table A.24). I find a close to 10 percent greater overall variance of job value when estimating the sample with women (Table A.25). This is driven by greater estimates of firm non-wage value dispersion, and dispersion of idiosyncratic non-wage value. I find, however, very similar values regarding my two main results: (1) a very similar positive covariance of non-wage value and wage (4th and 5th row of Table A.25); (2) a positive covariance between firm wage and firm non-wage value offer (6th row of Table A.25), confirming the result from the baseline specification on the importance of firm-level rents in 2004–2011.

[71] Ideally, I would separately compare the sample of full-time workers and part-time workers, which is not possible as this would give rise to two different sets of strongly connected firms. For more details see the analogous discussion on including women in Appendix D.2.

Figure A.13: EMPLOYER-TO-EMPLOYER TRANSITIONS AND WAGE DIFFERENTIALS 2004–2011, SAMPLE WITH WOMEN AND MEN ONLY

Notes: Figure a shows the intensity of employer-to-employer transitions between firms in the strongly connected sample using all full-time workers. Figure b shows the intensity of employer-to-employer transitions between firms in the strongly connected sample using male full-time workers only. If mobility was random, the intensity would be equal to 1 for each cell. Intensities above 1 indicates that there are more transitions from row-industry to column-industry than expected under random mobility. See text in Section 2 for formal definition of intensity. Figure c shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions between firms in the strongly connected sample using all full-time workers with old employer in industry in row and new employer in industry in column. Figure d shows average log-wage differences (new log-wage − old log-wage) of employer-to-employer transitions between firms in the strongly connected sample using male full-time workers only with old employer in industry in row and new employer in industry in column. Missing cells in figures c and d contain fewer than 10 observations. These statistics are based on employer-to-employer transitions shown in Table A.21 (Figure a and c: Column 2; Figure b and d: Column 4).

Table A.21: POPULATION AND SAMPLE 2004–2011 WITH AND WITHOUT WOMEN

	With Women		**Men Only**	
	All	**Strongly connected**	**All**	**Strongly connected**
	(1)	(2)	(3)	(4)
A. Sample size				
People-years	15,488,900	8,133,160	9,906,446	5,480,901
People	2,927,762	1,534,497	1,712,585	964,635
Employers	260,429	5,944	182,811	5,944
B. Summary statistics				
Share Female	0.36	0.33	0.00	0.00
Mean age	39.81	39.78	40.21	40.21
Share blue collar	0.35	0.32	0.43	0.39
Median monthly wage (2012 €)	2,948	3,253	3,196	3,481
Mean log monthly wage	8.03	8.15	8.14	8.23
Mean log monthly wage	8.03	8.15	8.14	8.23
Var log monthly wage	0.22	0.20	0.21	0.20
C. Industry shares				
Manufacturing	0.25	0.31	0.31	0.39
Utilities	0.02	0.02	0.02	0.03
Construction	0.07	0.04	0.10	0.06
Retail trade, cars	0.16	0.11	0.15	0.10
Transportation	0.06	0.07	0.07	0.08
Hotel and restaurant	0.03	0.01	0.02	0.00
Information and communication	0.03	0.03	0.03	0.03
Finance and insurance	0.05	0.07	0.05	0.06
Real estate	0.02	0.01	0.02	0.02
Prof./scientific/tech. services	0.05	0.03	0.04	0.03
Services	0.06	0.05	0.05	0.05
Public admin./education	0.15	0.20	0.10	0.13
Health and social	0.05	0.04	0.02	0.02
D. Employer-to-employer transitions				
Transitions	265,904	100,529	178,835	74,271
Share excess separations	0.47	0.45	0.48	0.47
Mean log wage increase	0.10	0.10	0.09	0.10
Mean log wage increase (adjusted)†	0.05	0.05	0.05	0.05
Share wage increase (adj.)	0.59	0.60	0.59	0.60
Share both employers same industry	0.44	0.46	0.43	0.45

Note: Summary statistics on the sample of full-time workers 2004–2011, when restricting the sample to male workers (columns 3 and 4), and without any restriction on workers' sex (columns 1 and 2). Columns 3 and 4 correspond to columns 3 and 4 of Table 1. The industry classification is based on NACE Rev. 2 main sections. I combined section D & E (Utilities), O & P (Public admin./education) and N & S (Services). Not shown: Agriculture, forestry and fishing, Mining, Arts and entertain., 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 Appendix G.2)

Table A.22: BY INDUSTRY – NUMBER OF HIRES AND QUITS 2004–2011, WITH AND WITHOUT WOMEN

	With Women		Men Only	
	Hires (1)	Quits (2)	Hires (3)	Quits (4)
Manufacturing	15,015	12,029	12,659	10,256
Utilities	1,202	686	1,034	580
Construction	2,557	2,708	2,365	2,553
Retail trade, cars	6,188	6,641	4,193	4,425
Transportation	3,774	4,053	3,049	3,120
Hotel and restaurant	237	309	137	178
Information and communication	2,660	2,689	2,106	2,083
Finance and insurance	4,993	4,670	3,050	2,739
Real estate	834	907	601	686
Prof./scientific/tech. services	3,476	3,129	2,529	2,220
Services	3,701	8,238	2,600	6,434
Public admin./education	8,475	7,150	4,025	3,140
Health and social	2,175	2,126	840	806

Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for 2004–2011. Columns 1 and 2: Number of hires and quits in sample without any restriction on workers' sex (column 2 Table A.21). Columns 3 and 4: Number of hires and quits in baseline sample of men (column 4 Table A.21). All statistics after downweighting quits from contracting firms according to procedure explained in Section 2.

Table A.23: BY INDUSTRY – WAGES AND WAGE DIFFERENTIALS 2004–2011 WITH AND WITHOUT WOMEN

	With Women		**Men Only**	
	(1)		(2)	
Median monthly wages by industry (2012 €)				
Manufacturing	3,426		3,582	
Utilities	4,700		4,852	
Construction	3,325		3,342	
Retail trade, cars	2,755		3,039	
Transportation	2,857		2,932	
Hotel and restaurant	2,089		2,236	
Information and communiaction	4,725		5,017	
Finance and insurance	4,435		5,054	
Real estate	3,028		3,185	
Prof./scientific/tech. services	3,857		4,211	
Services	2,668		2,826	
Public admin./education	3,073		3,248	
Health and social	3,009		3,211	
Δ *log-wage of employer-to-employer transitions*	Hires	Quits	Hires	Quits
Manufacturing	0.062	0.006	0.061	0.004
Utilities	0.047	0.041	0.042	0.040
Construction	0.056	0.014	0.056	0.013
Retail trade, cars	0.057	0.058	0.057	0.060
Transportation	0.036	0.047	0.034	0.048
Hotel and restaurant	-0.008	0.079	-0.025	0.070
Information and communiaction	0.058	0.045	0.054	0.046
Finance and insurance	0.080	0.075	0.083	0.090
Real estate	0.043	0.043	0.042	0.035
Prof./scientific/tech. services	0.073	0.078	0.073	0.078
Services	0.020	0.079	0.018	0.078
Public admin./education	0.022	0.069	-0.002	0.092
Health and social	-0.007	0.024	-0.005	0.048

Note: This table reports wages and wage differentials by industry, using the samples of strongly connected firms (columns 2 and 4 of Table A.21). The panel Δ *log-wage of employer-to-employer transitions* takes into account that wages at the old employer are observed in year t, and at the new employer in year $t + 2$ by subtracting time and experience effects from the wage at the new employer using the estimates from my AKM-regression (see Appendix G.2). In the lower panel, transitions are weighted by their probability of being an excess separation as defined in the text in Section 2.

Table A.24: COVARIANCES OF JOB VALUE COMPONENTS – SAMPLE WITH WOMEN

		Job value	Wage	Non-wage	Wage				Non-wage	
					Person	Employer	$\boldsymbol{X}'_{it}\beta$	r_{it}	Employer	Idio.
	Job value	0.616								
	Wage	0.238	0.190							
	Non-wage	0.378	0.048	0.330						
Wage	Person	0.198	0.157	0.040	0.154					
Wage	Employer	0.024	0.019	0.006	0.006	0.013				
Wage	$\boldsymbol{X}'_{it}\beta$	0.002	0.001	0.001	-0.003	-0.001	0.005			
Wage	r_{it}	0.014	0.014	0	0	0	0	0.014		
Non-wage	Employer	0.270	0.048	0.223	0.040	0.006	0.001	0	0.223	
Non-wage	Idiosyncratic	0.107	0	0.107	0	0	0	0	0	0.107

Notes: This table reports covariances of job-value components in the sample of all full-time workers. The covariances are estimated using all person-year observations from Table A.21 column 2. The corresponding covariance matrix for the sample using male full-time workers only is in Table A.8.

Table A.25: JOB VALUE VARIANCE 2004–2011, WITH WOMEN AND WITHOUT WOMEN

	With Women (1)	Men Only (2)
$Var(V_{ij})$	0.616	0.564
$Var(ln(w_{ij}))$	0.190	0.197
$Var(ln(a_j) + \epsilon_{ij})$	0.330	0.277
$2Cov(w_{ij}, ln(a_j) + \epsilon_{ij})$	0.095	0.090
$2Cov(\alpha_i, ln(a_j))$	0.081	0.082
$2Cov(\psi_j, ln(a_j))$	0.012	0.006

Notes: This table reports the variance of job value and covariances of components of job value 2004–2011, in column 1 in the sample of all full-time workers (Table A.21 column 2), and in column 2 in the sample of male full-time workers only (Table A.21 column 4).

D.3 Restriction 3: Only Full-Time Workers

In order to give earnings recorded in administrative data the interpretation of a piece-rate wage, I limit my sample to full-time workers.[72] While I lack information on wages of part-time workers, I can compare the composition of my sample with the sample including part-time workers, and the employer-to-employer transition dynamics in the two samples, which I will do in following.

Table A.26 compares the sample of full-time male workers (columns 2 and 4) with the sample that obtains if women and part-time workers are included (columns 1 and 3). We see that this increases the sample size by more than 50 percent, as well as the number of employer-to-employer transitions (Panel D.).

Table A.27 and Figure A.14 show that patterns of worker flows, when aggregated at the industry level, look very similar in my sample and when including part-time workers and women. This suggests that, indeed, preferences over firms are similiar among all workers as in my sample.[73]

I can estimate firm value offers in the framework of my search model by replacing $ln(w)+a$ by U, which then allows me to directly estimate each employer's total value offer solely using worker flows.[74] Hence, I estimate two U using the likelihood function in Proposition 1, one relying on employer-to-employer transitions of full-time workers in my sample only, and one including employer-to-employer transitions of part-time workers. I otherwise apply the same restrictions for part-time workers' employer-to-employer transitions (see Section 2).

Table A.29 shows model parameters from estimating the model on the two samples for the 1996–2003 and the 2004–2011 period. We see that including part-time workers and women increases the number of transitions between firms by more than 50 percent, and that the firm size parameter and the number of hires from non-employment are about 100 percent greater. Nevertheless, I find that the employer values estimated on the two samples are .9-correlated. This high correlation of total firm value offers suggests that preferences of part-time workers and full-time workers are highly correlated, especially if we consider that sampling variation biases the correlation downwards. I therefore conclude that part-time workers and full-time workers are likely offered similar non-wage values by the employers in my sample.

Sorting of Workers in Sample Vs. Whole Austrian Labor Market Industries explain 31 percent (1996–2003, 25 percent 2004–2011) of the variance of employers' non-wage value.[75] Comparing the distribution of all Austrian workers across industries with the distribution of workers across industries in my sample will thus help understand the extent to which my results are valid for the whole Austrian labor market. Figure A.15 shows how workers in Austria are sorted across industries. The industries in Figure A.15 are ordered by their employers' average non-wage value. Comparing the sorting of workers in my sample to the sorting of workers

[72]Recall that I can identify full-time workers after 2002, and use before 2002 that more than 97 percent of men are working full-time.

[73]This holds if employers offer similar wage premia to part-time workers as to full-time workers.

[74]I can also calculate the average value an employer offers to its employed workers by using the formula $U^{employed} = \psi + a$. Reassuringly, I obtain a correlation of .98 between the directly estimated total value offer and $U^{employed}$ (that the correlation is slightly lower than 1 might be explained by the difference between the offered wage premium ψ and the wage premium workers employed at an employer actually have.)

[75]This is the R^2 of a regression of employers' non-wage value on industry dummies, weighting employers by their number of person-year observations.

Table A.26: STONGLY CONNECTED FIRMS INCLUDING/EXCLUDING PART-TIME WORKERS AND WOMEN 1996–2003 & 2004–2011

	1996 – 2003		2004–2011	
	All	Full-Time Men	All	Full-Time Men
	(1)	(2)	(3)	(4)
A. Sample size				
People-years	7,833,859	4,513,833	9,913,346	5,480,901
People	1,413,665	797,492	1,775,902	964,635
Employers	4,544	4,544	5,944	5,944
B. Summary statistics				
Mean age	38.70	39.07	40.13	40.21
Share blue collar	0.33	0.43	0.28	0.39
Share female	0.40	0.00	0.41	0.00
Share full-time			0.84	1.00
C. Industry shares				
Manufacturing	0.29	0.39	0.28	0.39
Utilities	0.02	0.03	0.02	0.03
Construction	0.04	0.05	0.04	0.06
Retail trade, cars	0.11	0.10	0.12	0.10
Transportation	0.06	0.07	0.06	0.08
Hotel and restaurant	0.01	0.00	0.00	0.00
Information and communication	0.02	0.02	0.03	0.03
Finance and insurance	0.08	0.08	0.07	0.06
Real estate	0.02	0.02	0.01	0.02
Prof./scientific/tech. services	0.04	0.05	0.03	0.03
Services	0.05	0.04	0.05	0.05
Public admin./education	0.22	0.13	0.22	0.13
Health and social	0.05	0.02	0.06	0.02
D. Employer-to-employer transitions				
Transitions	92,902	58,349	117,855	74,271
Share excess separations	0.52	0.54	0.43	0.47
Share both employers same industry	0.48	0.47	0.46	0.45

Note: This table reports summary statistics on all workers working at least one full calendar year at strongly connected firms (columns 1 and 3) and those satisfying the baseline sample restrictions (columns 2 and 4). Information on full-time/part-time employment only available for 2004–2011 panel. The industry classification is based on NACE Rev. 2 main sections. I combined section D & E (Utilities), O & P (Public admin./education) and N & S (Services). Not shown: Agriculture, forestry and fishing, Mining, Arts and entertain., 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.

in the entire Austrian labor market, we see that my sample has a lower share of workers in industries at the lower end of the non-wage value distribution, while workers in my sample

Table A.27: BY INDUSTRY – NUMBER OF HIRES AND QUITS 1996–2004, INCLUDING/EXCLUDING PART-TIME WORKERS AND WOMEN

	All		Full-time Men	
	Hires (1)	Quits (2)	Hires (3)	Quits (4)
Manufacturing	11,119	10,159	8,366	7,958
Utilities	594	394	444	300
Construction	1,802	1,852	1,563	1,629
Retail trade, cars	5,334	6,030	3,076	3,276
Transportation	2,703	3,052	1,846	2,094
Hotel and restaurant	252	347	122	154
Information and communication	2,757	2,302	1,824	1,454
Finance and insurance	4,497	4,385	2,470	2,424
Real estate	968	1,101	672	786
Prof./scientific/tech. services	2,328	2,522	1,587	1,766
Services	2,721	3,864	1,531	2,246
Public admin./education	6,479	5,842	2,538	2,027
Health and social	2,681	2,530	741	725

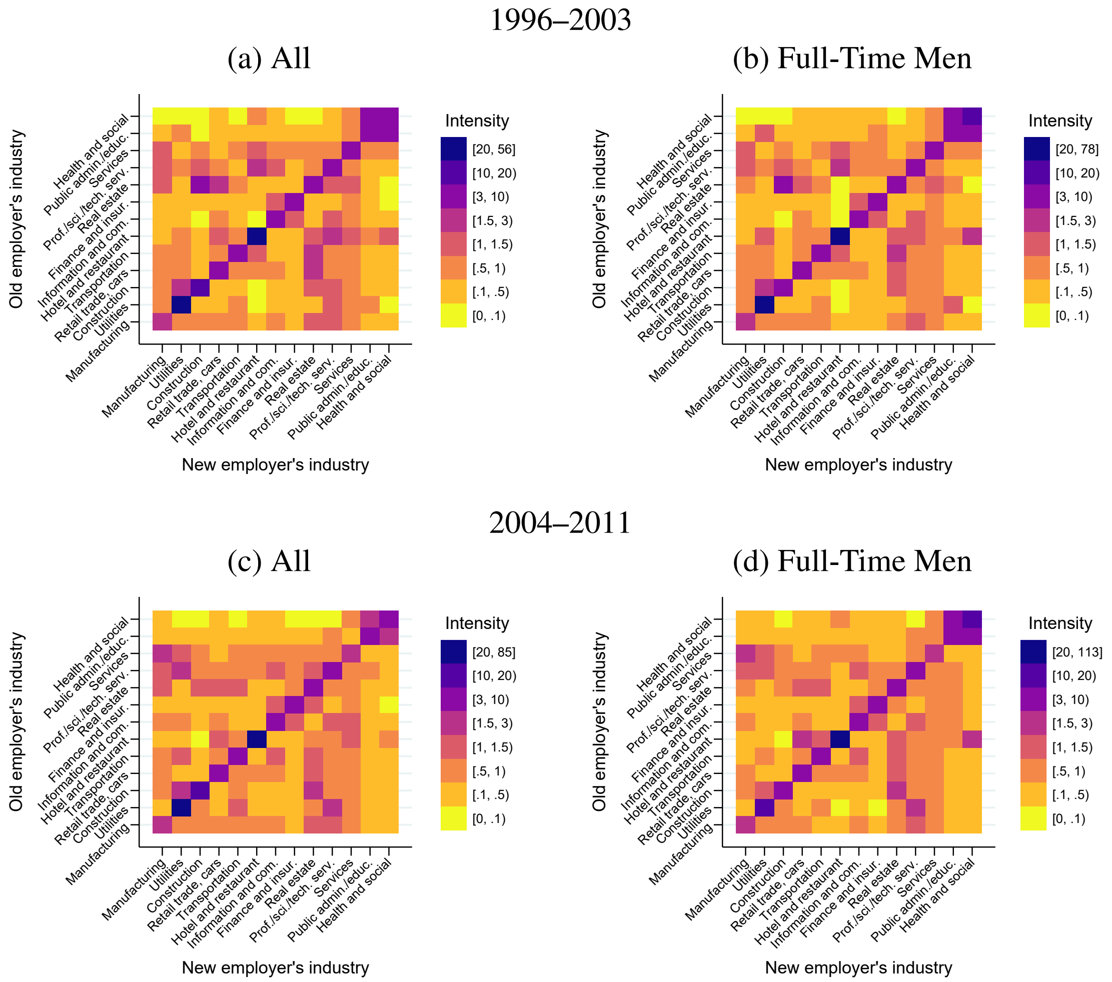
Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for 1996–2003. Columns 1 and 2 show the number of hires and quits in sample with part-time workers and women (column 1 of Table A.26). . Columns 3 and 4 show the number of hires and quits in baseline sample (column 2 of Table A.26). All statistics after downweighting quits from contracting firms according to procedure explained in Section 2.

Table A.28: BY INDUSTRY – NUMBER OF HIRES AND QUITS 2004–2012, INCLUDING/EXCLUDING PART-TIME WORKERS AND WOMEN

	All		Full-time Men	
	Hires	Quits	Hires	Quits
	(1)	(2)	(3)	(4)
Manufacturing	16,826	13,165	12,659	10,256
Utilities	1,355	774	1,034	580
Construction	2,767	2,970	2,365	2,553
Retail trade, cars	7,695	8,575	4,193	4,425
Transportation	4,288	4,507	3,049	3,120
Hotel and restaurant	277	383	137	178
Information and communication	3,049	3,163	2,106	2,083
Finance and insurance	5,829	5,454	3,050	2,739
Real estate	951	1,013	601	686
Prof./scientific/tech. services	4,045	3,586	2,529	2,220
Services	4,881	10,281	2,600	6,434
Public admin./education	11,611	9,727	4,025	3,140
Health and social	3,548	3,575	840	806

Note: This table reports totals of employer-to-employer hires and employer-to-employer quits by industry for 2004–2011. Columns 1 and 2 show the number of hires and quits in sample with part-time workers and women (column 3 of Table A.26). . Columns 3 and 4 show the number of hires and quits in baseline sample (column 4 of Table A.26). All statistics after downweighting quits from contracting firms according to procedure explained in Section 2.

Figure A.14: INTENSITY OF EMPLOYER-TO-EMPLOYER TRANSITIONS, INCLUDING/EXCLUDING PART-TIME WORKERS AND WOMEN



Notes: Figure a and c show the intensity of employer-to-employer transitions between firms in the strongly connected sample using all workers. Figure b and d show the intensity of employer-to-employer transitions between firms in the strongly connected sample using male full-time workers only. If mobility was random, the intensity would be equal to 1 for each cell. Intensities above 1 indicates that there are more transitions from row-industry to column-industry than expected under random mobility. See text in Section 2 for formal definition of intensity. Based on employer-to-employer transitions shown in Table A.26 (Figure a: Column 1 Table A.26; Figure b: Column 2 Table A.26; Figure c: Column 3 Table A.26; Figure d: Column 4 Table A.26).

are more strongly sorted to employers at the upper end of the non-wage value distribution.[76] This industry level statistics suggests that while the average non-wage value in my sample is higher than in the Austrian labor market overall, the dispersion of non-wage values is probably similar. Indeed, on the industry level, the variance of non-wage value is only slightly lower when weighting the industries by the number of person-year observations in my sample (var $= .051$ in 1996–2003 and 2004–2011), than when using full-time workers (var $= .054$ in 1996–2003 and var $= .055$ in 2004–2011), or all workers (var $= .058$ in 1996–2003 and var $= .059$

[76]This difference is probably explained by employer size and employment duration, which is both substantially higher for employers in the manufacturing or public administration/education industries than for employers in the hotel and restaurant, services, or retail trade and cars industries.

Table A.29: Total Firm Value Offers, Including/Excluding Part-Time Workers and Women

	1996 – 2003		2004–2011	
	All	**Full-Time Men**	**All**	**Full-Time Men**
	(1)	(2)	(3)	(4)
Model				
Transitions	92,902	58,349	117,855	74,271
Firm size (people-years)	1,376	799	1,310	727
Hires from non-employment	123	60	126	59
Correlation of firm value offers				
Corr. U all, U full-time men	0.90		0.90	

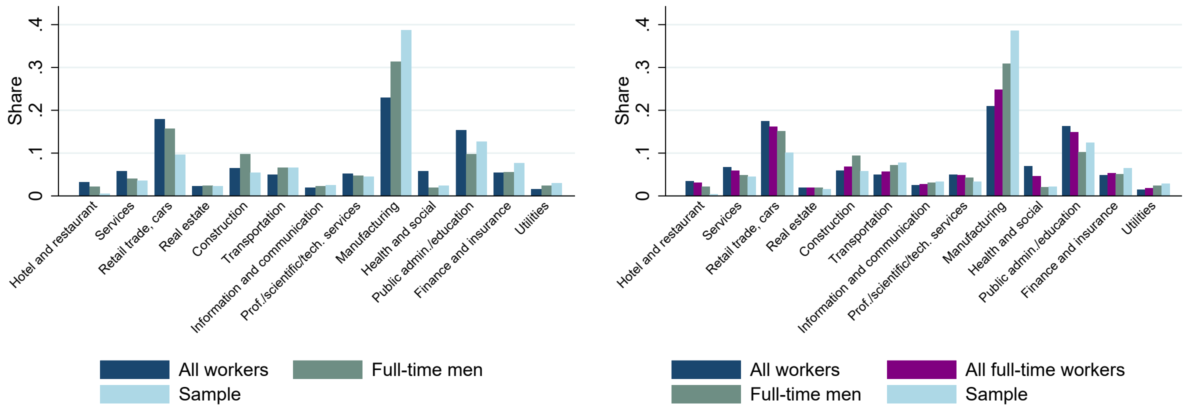
Notes: The panel *Model* of this table reports the number of employer-to-employer transitions and parameter values of the model estimated with all workers (columns 1 and 3 of Table A.26), and with full-time male workers only (columns 2 and 4 of Table A.26). The panel *Correlation of firm value offers* shows the correlation of employer values estimated with the two sets of transitions using the likelihood function in Proposition 1 with $U = ln(w) + ln(a)$, that is, $\mathcal{L} = \prod_{s=1}^{S} \Phi[U_j - U_k]^{\frac{1}{f_j^{NE}} \frac{1}{g_k}}$.

in 2004–2011).

Figure A.15: Sorting of Workers in Sample and in Austrian Labor Market

(a) 1996–2003 Panel

(b) 2004–2011 Panel



Notes: These figures show the share of person-year observations by industry for different subsamples for the 1996–2003 and the 2004–2011 panel. *All workers* denotes the sample containing all workers employed at a single employer for the full calendar year. *All full-time workers* denotes the sample containing all workers employed full-time at a single employer for the full calendar year (only available for 2004–2011). *Full-time men* denotes all male workers employed full-time at a single employer fo the full calendar year (columns 1 and 3 of Table 1). *Sample* denotes my baseline sample, which are all male workers employed full-time at a single employer that is in the strongly connected set for the full calendar year (columns 2 and 4 of Table 1).

Overall, I conclude that employers in my sample likely offer similar non-wage values to part-time workers as to full-time workers. I find no evidence that workers from the entire

Austrian labor market are sorted to employers in a way that would alter my conclusions on non-wage value dispersion and its implications for inequality.

E Derivation of the Model and Estimator in Section 3

E.1 Value Functions

Employed Workers Employed workers' value of being at employer k is characterized by the following Bellman equation:[77]

$$\underbrace{(\alpha_{it} + \tilde{\psi}_k) + ln(a_k)}_{\text{value of being at } k} = \underbrace{v(\alpha_{it}, \tilde{\psi}_k, a_k)}_{\text{flow payoff}}$$

$$+ \underbrace{\beta}_{\text{discounter}} \Bigg[\underbrace{\delta_k(1-\rho_k)V^n}_{\substack{\text{exogenous employer-}\\ \text{to-non-employment}}} + \underbrace{\delta_k \rho_k \sum_j \int_\eta \int_\epsilon ((\alpha_{it} + \tilde{\psi}_j + \eta_j) + ln(a_j) + \epsilon_j) dF(\eta) dF(\epsilon) f_{jk}}_{\text{exogenous employer-to-employer}} + \underbrace{(1-\delta_k)}_{\text{no exogenous transition}} *$$

$$\underbrace{\lambda_1 \sum_j \int_{\eta_1} \int_{\eta_2} \int_{\epsilon_1} \int_{\epsilon_2} \max\{(\alpha_{it} + \tilde{\psi}_k + \eta_k) + ln(a_k) + \epsilon_k, (\alpha_{it} + \tilde{\psi}_j + \eta_j) + ln(a_j) + \epsilon_j\} dF(\eta_1) dF(\eta_2) dF(\epsilon_1) dF(\epsilon_2) f_{jk}}_{\text{receive job offer and make binary choice}}$$

$$+ \underbrace{(1-\lambda_1)}_{\text{no job offer}}((\alpha_{it} + \tilde{\psi}_k) + ln(a_k))\Bigg] \tag{A.1}$$

meaning that a worker at employer k has value composed of its wage $(\alpha_{it} + \tilde{\psi}_k)$ and non-wage value $ln(a_k)$, which equals his flow payoff as well as a continuation value which is discounted by β. Workers are laid off with probability δ_k, in which case they go to non-employment with probability $(1 - \rho_k)$ and immediately find a new job with probability ρ_k.

The part of the continuation value relevant for my estimation is the case when the worker receives a job offer and makes a binary choice, which happens with probability $(1-\delta_k)\lambda_1$. The intensity of offers from employer j is f_{jk}. When the worker receives an offer from an outside employer j, he draws a new offer from employer k, compares the two offers, and selects the one offering him greater value. This process is represented by the two terms in the max-function. There are two stochastic elements associated with the decision to choose the maximum: first, there is randomness in the wages the two employers offer η, and second, there is the workers' idiosyncratic valuation for each employer's offer ϵ.

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

$$\underbrace{V^n}_{\text{value of non-empl.}} = b + \beta \left(\underbrace{\lambda_0 \sum_j \int_\eta \int_\epsilon ((\alpha_{it} + \tilde{\psi}_j + \eta_j) + ln(a_j) + \epsilon_j) dF(\eta) dF(\epsilon) f_{ji}^{ne}}_{\text{receive job offer}} + \underbrace{(1-\lambda_0)V^n}_{\text{no offer}} \right) \tag{A.2}$$

[77] Following Arcidiacono and Ellickson (2011, p. 368) I write the value function as the value of being at employer k just before the first idiosyncratic draws ι and η are revealed, which is why the idiosyncratic draws do not show up in the flow utility.

Where λ_0 represents the probability with which non-employed workers receive job offers and f_{ji}^{NE} represents worker i's probability of receiving an offer from employer j. Hence, non-employed workers probability of receiving a job offer from a particular employer is allowed to depend on his characteristics (e.g., education, skills). In the case when non-employed workers receive an employment offer, they draw η and ϵ and accept the offer.

Following Bonhomme and Jolivet (2009) and Sorkin (2018) I assume that $\lambda_0 * \mathbb{E}_i[f_{ji}^{NE}] = \lambda_1 * \mathbb{E}_{ik}[f_{j,ik}]$, that is, that non-employed workers in expectation receive offers with the same relative intensity from a particular employer as employed workers. This allows me to estimate the intensity with which employers make offers to employed workers from where non-employed workers get hired.

E.2 Proof of Proposition 1

The idea of the proof is to show that in the limit (that is, when the number of periods in which firms make offers gets large) under my model's assumptions and correcting for employers' size and offer intensity, accepted job offers (leading to employer-to-employer transitions) made by employer j to workers at employer k are equivalent to rejected job offers (not leading to employer-to-employer transitions) made by employer k to workers at employer j, why the non-wage value of employer j is pairwise (over-)identified from employer-to-employer transitions with any other employer connected to employer j.

Start by noting that by equation A.1 the probability in a given time period that a worker at employer k who has not made an exogenous transition (either to non-employment or another employer) receives an offer from employer j equals $\lambda_1 f_{jk}$. With g_k workers at employer k who have no exogenous transition, over T time periods there is a sequence $(\lambda_1 f_{jk})_{s \,\in\, T * g_k}$ of offers from employer j of which $\sum_{s \,\in\, Tg_k} \lambda_1 f_{jk} \mathbb{1}(j > k)_s$ are accepted and $\sum_{s \,\in\, Tg_j} \lambda_1 f_{kj} \mathbb{1}(j < k)_s$ are rejected, where

$$\mathbb{1}(j > k)_s = \mathbb{1}((\tilde{\psi}_j + \eta_{js}) + ln(a_j) + \epsilon_{js} > (\tilde{\psi}_k + \eta_{ks}) + ln(a_k) + \epsilon_{ks})$$

by the assumption that $\epsilon \sim i.i.d.N(0, \sigma^2)$, we obtain that the expected number of workers per period at employer k receiving and <u>accepting</u> an offer from employer j is

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T}(\lambda_1 f_{jk} \mathbb{1}(j > k))_{s \,\in\, Tg_j} = \\ \lambda_1 g_k f_{jk} \int_{(\eta_j - \eta_k)} \Phi(((\tilde{\psi}_j - \tilde{\psi}_k) + (\eta_j - \eta_k)) + ln(a_j) - ln(a_k)) dF(\eta_j - \eta_k) \end{aligned} \tag{A.3}$$

where Φ denotes the cumulative distribution function of a normal distribution with mean zero and variance $2\sigma^2$. The expected number of workers per period at employer k receiving and <u>rejecting</u> an offer from employer j is

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T}(\lambda_1 f_{jk} \mathbb{1}(j < k))_{s \,\in\, T g_j} = \\ \lambda_1 g_k f_{jk} &\int_{(\eta_k-\eta_j)} \Phi(((\tilde{\psi}_k - \tilde{\psi}_j) + (\eta_k - \eta_j)) + ln(a_k) - ln(a_j))dF(\eta_k - \eta_j) \end{aligned} \tag{A.4}$$

Following the same logic, the the expected number of workers per period at employer j receiving and <u>accepting</u> an offer from employer k is

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T}(\lambda_1 f_{kj} \mathbb{1}(k > j))_{s \,\in\, T g_j} = \\ \lambda_1 g_j f_{kj} &\int_{(\eta_k-\eta_j)} \Phi(((\tilde{\psi}_k - \tilde{\psi}_j) + (\eta_k - \eta_j)) + ln(a_k) - ln(a_j))dF(\eta_k - \eta_j) \end{aligned} \tag{A.5}$$

while the expected number of workers per period at employer j receiving and <u>rejecting</u> an offer from employer k is

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T}(\lambda_1 f_{kj} \mathbb{1}(k < j))_{s \,\in\, T g_k} = \\ \lambda_1 g_j f_{kj} &\int_{(\eta_j-\eta_k)} \Phi(((\tilde{\psi}_j - \tilde{\psi}_k) + (\eta_j - \eta_k)) + ln(a_j) - ln(a_k))dF(\eta_j - \eta_k) \end{aligned} \tag{A.6}$$

Plugging in all expected offers made in a period by employer k to workers at employer j, which are the offers in equations A.5 and A.6, into the likelihood function of Proposition 1, we obtain the likelihood of all offers received by workers of employer j from employer k as:

$$\begin{aligned} \mathcal{L} =& exp\Big(\int_{(\eta_k-\eta_j)} log(\underbrace{\Phi[(\tilde{\psi}_k - \tilde{\psi}_j) + (\eta_k - \eta_j) + ln(a_k) - ln(a_j)]^{g_j f_{kj} \Phi[(\tilde{\psi}_k - \tilde{\psi}_j) + (\eta_k - \eta_j) + ln(a_k) - ln(a_j)] * \frac{1}{f_k^{NE}} \frac{1}{g_j}}}_{\text{Accepted offers}} * \\ & \underbrace{\Phi[(\tilde{\psi}_j - \tilde{\psi}_k) - (\eta_k - \eta_j) + ln(a_j) - ln(a_k)]^{g_j f_{kj} \Phi[(\tilde{\psi}_j - \tilde{\psi}_k) - (\eta_k - \eta_j) + ln(a_j) - ln(a_k)] * \frac{1}{f_k^{NE}} \frac{1}{g_j}})dF(\eta_k - \eta_j)}_{\text{Rejected offers}}\Big) \end{aligned}$$

Where I use that due to the symmetry of η_k and η_j, $dF(\eta_k - \eta_j) = dF(\eta_j - \eta_k)$. We see immediately that g_j cancels out. Furthermore, we can take logs:

$$ln(\mathcal{L}) = \int_{(\eta_k-\eta_j)} \Big(\underbrace{\frac{f_{kj}}{f_k^{NE}}\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+(\eta_k-\eta_j)+ln(a_k)-ln(a_j)]ln(\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+(\eta_k-\eta_j)+ln(a_k)-ln(a_j)])}_{\text{Accepted offers}} + \underbrace{\frac{f_{kj}}{f_k^{NE}}\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)-(\eta_k-\eta_j)+ln(a_j)-ln(a_k)]ln(\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)-(\eta_k-\eta_j)+ln(a_j)-ln(a_k)])}_{\text{Rejected offers}}\Big)dF(\eta_k-\eta_j) \tag{A.7}$$

Following the same idea, we can write the log-likelihood of all offers received by workers of employer k from employer j as:

$$ln(\mathcal{L}) = \int_{(\eta_k-\eta_j)} \Big(\underbrace{\frac{f_{jk}}{f_j^{NE}}\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)-(\eta_k-\eta_j)+ln(a_j)-ln(a_k)]ln(\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)-(\eta_k-\eta_j)+ln(a_j)-ln(a_k)])}_{\text{Accepted offers}} + \underbrace{\frac{f_{jk}}{f_j^{NE}}\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+(\eta_k-\eta_j)+ln(a_k)-ln(a_j)]ln(\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+(\eta_k-\eta_j)+ln(a_k)-ln(a_j)])}_{\text{Rejected offers}}\Big)dF(\eta_k-\eta_j) \tag{A.8}$$

As by Assumption 1 $\frac{f_{kj}}{f_k^{NE}} = \frac{f_{jk}}{f_j^{NE}}$, we have that the likelihood contributions of accepted offers in equation A.7 equals the likelihood contributions of rejected offers in equation A.8. Thus, the total log-likelihood of all binary choices made over offers between employer k and employer j can be written as a function of accepted offers only:

$$ln(\mathcal{L}) = \frac{f_{jk}}{f_j^{NE}} \int_{(\eta_k-\eta_j)} \Big(\underbrace{\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)-(\eta_k-\eta_j)+ln(a_j)-ln(a_k)]ln(\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)-(\eta_k-\eta_j)+ln(a_j)-ln(a_k)])}_{\text{employer-to-employer transitions from employer } k \text{ to } j} + \underbrace{\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+(\eta_k-\eta_j)+ln(a_k)-ln(a_j)]ln(\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+(\eta_k-\eta_j)+ln(a_k)-ln(a_j)])}_{\text{employer-to-employer transitions from employer } j \text{ to } k}\Big)dF(\eta_k-\eta_j) \tag{A.9}$$

As this holds for any pair of employers $j \in J$ and $k \neq j \in J$, it also holds for the joint likelihood of all observed transitions between all employers in J. Consistent estimates of the parameter σ,which is identified, because the coefficient on wage λ is normalized to 1, and the vector of a's are obtained under standard regularity conditions of MLE.[78] QED.

[78]This becomes more clear when canceling out the constant $2\frac{f_{jk}}{f_j^{NE}}$ and ignoring the random part of the wage offer, in which case equation A.9 equals the standard probit likelihood function (with free variance parameter): $ln(\mathcal{L}) = \underbrace{\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)+ln(a_j)-ln(a_k)]ln(\Phi[(\tilde{\psi}_j-\tilde{\psi}_k)+ln(a_j)-ln(a_k)])}_{\text{employer-to-employer transitions from employer } k \text{ to } j} + \underbrace{\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+ln(a_k)-ln(a_j)]ln(\Phi[(\tilde{\psi}_k-\tilde{\psi}_j)+ln(a_k)-ln(a_j)])}_{\text{employer-to-employer transitions from employer } j \text{ to } k}$. The random part of wage η is required to be non-degenerate for at least one firm in sample to avoid that firms' non-wage value is collinear with firms' wage offer.

F Shrinkage of Non-Wage Values

I rely on the empirical Bayes approach by Morris (1983) for the shrinkage of employers' non-wage values. The following exposition follows Sorkin (2018, Appendix H).

Let j be an employer, and n_J the number of employers. Let $n_{t(j)}$ be the sum of incoming and outgoing transitions of employer j. Let $ln(a_j)$ be employer j's true non-wage value, and $ln(a_j)^{raw}$ be the estimate of employer j's non-wage value. Let Q be the $n_J \times 1$ vector of $ln(a_j)^{raw}$. Let π_j^2 denote the variance of the estimate. Let κ^2 denote the estimate of the true variance of $ln(a_j)$. Let $\boldsymbol{x_j}$ be an $n_x \times 1$ vector of federal state by industry dummies.[79] Let X be the stacked vector of the $\boldsymbol{x_j'}$. Let $\boldsymbol{\lambda_0}$ be a $n_x \times 1$ vector of coefficients. Finally, let w_j the weight of employer j and W be the $n_J \times n_J$ matrix with w_j on the diagonal. These terms relate as follows:

$$w_j = n_{t(j)} \frac{1}{\hat{\pi}_i^2 + \hat{\kappa}^2} \tag{A.10}$$

$$\hat{\kappa}^2 = \max\left\{0, \frac{\sum_j w_j \{\frac{n_J}{n_J - n_x}(ln(a_j)^{raw} - \boldsymbol{x_j'}\hat{\boldsymbol{\lambda}})^2 - \hat{\pi}_j^2\}}{\sum_j w_j}\right\} \tag{A.11}$$

$$\hat{\boldsymbol{\lambda}} = (X^W X)^{-1} X' W Q \tag{A.12}$$

where the two unknowns are $\hat{\kappa}^2$ and $\hat{\boldsymbol{\lambda}}$. These are solved for in the following loop: Initialize $w_j = n_{t(j)}$. Then iterate the following until convergence:

1. Compute $\hat{\boldsymbol{\lambda}}$ using equation A.12
2. Compute $\hat{\kappa}^2$ using equation A.11
3. Check if $\hat{\kappa}^2$ has converged. If not, update the weights, w_j, and return to step 1.

The feasible shrinkage estimator then is:

$$\hat{b}_j = \Big(\frac{n_J - n_x - 2}{n_J - n_x}\Big)\Big(\frac{\hat{\pi}_j^2}{\hat{\pi}_j^2 + \hat{\kappa}^2}\Big) \tag{A.13}$$

$$ln(a_j)^{shrinked} = (1 - \hat{b}_j) ln(a_j)^{raw} + \hat{b}_j \boldsymbol{x_j'}\hat{\boldsymbol{\lambda}} \tag{A.14}$$

Where $ln(a_j)^{shrinked}$ is the estimate of employers' non-wage value on which I rely throughout my analyses.

[79] Industry classification: NACE main section (see Table 1). In case there are fewer than 10 employers in a federal state $\times$ industry cell, I merge it with a geographically adjacent federal state. I do so in a loop until there are 10 employers in every cell, or until all federal states of that industry are merged.

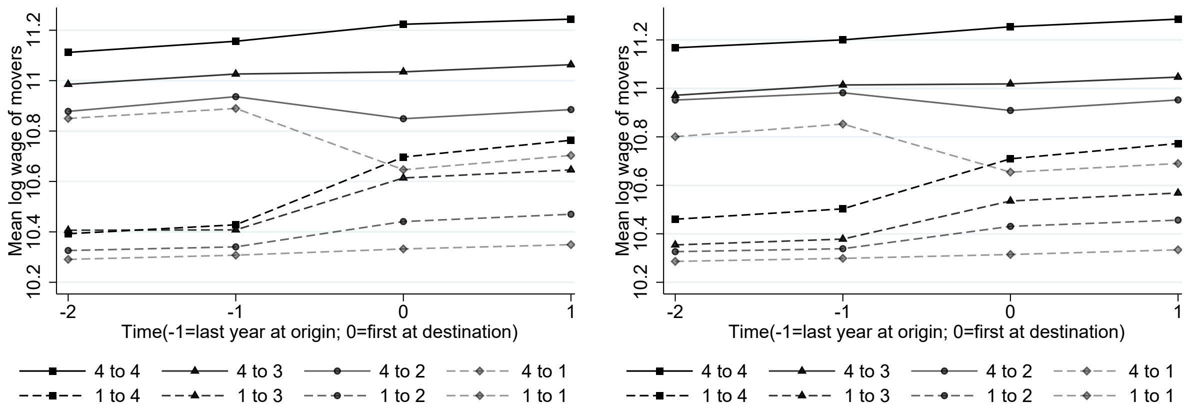
G Search Model and AKM

G.1 Specification Check for AKM Wage Model

Figure A.16: EVENT STUDY AROUND JOB MOVES

(a) 1996–2003 Panel

(b) 2004–2011 Panel



Notes: The figure shows the evolution of wages of workers who moved from employers in the top and bottom wage quartile groups to destination employers in any of the other quartile groups. The sample is based on workers in column 2 (1996–2003) and column 4 (2004–2011) of Table 1 who were reemployed in the destination employer in the next or the following year following separation from the origin employer, and were employed at the origin and destination employer for 2+ consecutive years. Origin/destination employers are based on the quartile of the average wage of co-workers.

G.2 Estimation of Wage Components

I proceed as follows to decompose the wage and obtain its covariance with job value components. First, I estimate a standard two-way fixed effect model of the following form using reghdfe in Stata (Correia, 2017):

$$ln(w_{it}) = \alpha_i + \boldsymbol{X}'_{it}\beta + \psi_{\boldsymbol{J}(i,t)} + r_{it} \tag{A.15}$$

With the coefficients estimated on age squared, age cubic and the year dummies, I correct the wage for the effects of these variables, that is,

$$ln(\tilde{w}_{it}) = ln(w_{it}) - \beta_{ageSq} * \text{age}^2_{it} - \beta_{ageCub} * \text{age}^3_{it} - \text{year}'_{it} * \boldsymbol{\beta_{year}} \tag{A.16}$$

I then use $ln(\tilde{w}_{it})$ and apply the estimator of Kline et al. (2020) to it, using their MATLAB-package (Kline et al., 2019). This returns unbiased estimates of the following moments: $Var(\psi_{\boldsymbol{J}(i,t)})$, $Var(\alpha_i)$ and $Cov(\psi_{\boldsymbol{J}(i,t)}, \alpha_i)$. I use these estimates in my decomposition. To calculate the other moments of the variance decomposition including the covariances with employers' non-wage values, I rely on the estimates from reghdfe.[80]

G.3 Nesting AKM Identifying Assumptions in Search Model

As Card et al. (2013, pp. 988–992) show, the critical necessary condition for OLS to identify the parameters of interest in equation 4 is that $\mathbb{E}[f^{j\prime} r] = 0$, where f^j is a vector of dummies defining the assignment of workers to firms and r is a vector of workers' wage error term. $\mathbb{E}[f^{j\prime} r] = 0$ holds if the assignment of workers to establishments $\boldsymbol{J}(i,t)$ is strictly exogenous with respect to r:

$$P(\boldsymbol{J}(i,t) = j|r) = P(\boldsymbol{J}(i,t) = j) \;\; \forall i, t \tag{A.17}$$

To see how this condition translates into my search model, consider two firms k and j, and consider workers at k. Equation A.17 holds for every worker at k if

$$\begin{aligned}
&(\delta_k(1-\rho_k)f_j^{NE} + \delta_k\rho_k f_{jk}) * (\tilde{\psi}_j - \psi_j) - \\
&((1-\delta_k)\lambda_1 f_{jk}) * \Bigg(\int_{\eta_j}\int_{\eta_k}\int_{\epsilon_j}\int_{\epsilon_k} \mathbb{I}((\alpha_{it} + \tilde{\psi}_j + \eta_j) + ln(a_j) + \epsilon_j > \\
&(\alpha_{it} + \tilde{\psi}_k + \eta_k) + ln(a_k) + \epsilon_k) dF(\eta_j)dF(\eta_k)dF(\epsilon_j)dF(\epsilon_k) * (\tilde{\psi}_j - \psi_j + \eta_j)\Bigg) = 0
\end{aligned} \tag{A.18}$$

where the terms embraced in the first parenthesis represent the probability that a worker at k makes an exogenous transition to firm j, potentially via non-employment (see workers' value function in Appendix E), and $\tilde{\psi}_j - \psi_j$ describes the wage residual in case of these transitions to employer j (see Section 3.2). The terms embraced in the first paranthesis on the second line of equation A.18 represent the probability that a worker at k does not make an exogenous transition and receives an offer from firm j. The remaining terms represent the wage residual the worker will have if the value of the offer from firm j exceeds the value of the offer from firm k.

Intuitively, equation A.18 says that if a worker at k is reassigned to j, the expected wage residual the worker will have at j will be $= 0$. Thus, assuming that A.18 holds for every firm-pair k and j in the sample (including when k=j), the search model indeed nests the condition in equation A.17.[81]

[80] The limited mobility bias only affects the three moments which I calculate using the estimator by Kline et al. (2020). For all other variance and covariances of the decomposition calculating them based on the estimates from reghdfe yields consistent estimates.

[81] Assuming that the initial state of the search model is that every worker is assigned to a firm and has a wage

H Lower Bound for Var(ϵ)

Due to the binary choice made by workers, the distribution of realized ϵ_{it} is truncated below for all workers that have either obtained their job through an employer-to-employer transition or that have rejected at least one job offer from an outside employer since they started working at their current employer. To calculate a lower lower bound on the variance of ϵ_{it}, I assume all workers have received at least one outside offer (or have been hired through an employer-to-employer transition) and estimate the distribution of ϵ_{it} to be truncated from below at -0.421 in the 1996–2003 sample and at -0.410 in the 2004–2011 sample, which equals the average lower bound on ϵ_{it} from all employer-to-employer transitions. I then calculate $Var(\epsilon_{it}|\epsilon_{it} \geq$ lower bound$)$ according to Greene (2000, p. 876). I obtain that $Var(\epsilon_{it}|\epsilon_{it} \geq -0.421) = .032$ for 1996–2003 and $Var(\epsilon_{it}|\epsilon_{it} \geq -0.410) = .034$ for 2004–2011.

residual equal to zero.

I Excluding Wage Growth from Non-Wage Value

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. I will now show how my results change if expectations about future wage growth are removed from my estimates of firm non-wage value.

Specifically, I modify Equation 3 by multiplying the offered wage used in the main specification with the expectation about wage growth over the next s years, i.e.,

$$P(V_{ik} > V_{ij}) = P(ln(w_{ik} * \mathbb{E}[r_{jpq}|q = u]) + ln(a_k) + \epsilon_{ik} > ln(w_{ij} * \mathbb{E}[r_{jpq}|q = 0]) + ln(a_j) + \epsilon_{ij}) \tag{A.19}$$

where $\mathbb{E}[r_{kpq}|q = u])$ denotes the wage growth workers at firm k expect over the next p years given that they have been employed at firm k for $q = u$ years ($q = 0$ at firm j as the worker would start a new employment spell at firm j).

I assume that $\mathbb{E}[r_{kpq}|q = u])$ equals the average wage growth workers at firm k with $q = u$ have experienced in the subsequent p years, condition on staying at firm k. I estimate $\mathbb{E}[r_{kpq}|q = u])$ as follows: First, I estimate the one-year wage growth for every pair k, q in my sample. I then build the product of all p one-year wage growth estimates for each year of tenure between k, q and $k, q + p$, which results in $\mathbb{E}[r_{kpq}|q = u])$.[82]

I re-estimate my model for the 1996–2003 period and the 2004–2011 period using $q = 4, 6, 8$ as time horizon. Table A.30 shows the decomposition results for $q = 8$ analogous to the baseline decomposition in Table 4.[83] As we might expect, removing expectations about future wage growth from $ln(a)$ reduces the dispersion of non-wage value $Var(ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{ij})$. However, it does not change the result that workers with higher wages have higher non-wage value ($2Cov(w_{ij}, ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{ij})$), and that the covariance between firms' wage premium and firm non-wage value changed from negative in the 1996–2003 panel to positive in the 2004–2011 panel ($2Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a_{\boldsymbol{J}(i,t)}))$). Thus, the conclusions regarding inequality and its evolution over time do not change when wage growth up to 8 years into the future is removed from the non-wage value estimates.

Note that there will still be some expectations of wage growth left to be captured by $ln(a_k)$ in this specification, for example wage growth at firm k beyond 8 years of tenure, or wage

[82]For some high values of q I cannot estimate the wage growth and thus assume that wage growth is equal to 0. This is innocuous as these cases are rare, and wage growth is typically close to 0 after long tenure.

[83]I only report the results for $q = 8$, noting that the results for $q = 4$ and $q = 6$ are similar.

Table A.30: ACCOUNTING FOR WAGE GROWTH: JOB VALUE VARIANCE 1996–2003 AND 2004–2011

	1996–2003 (1)	2004–2011 (2)	(2)-(1)
$Var(V_{ij})$	0.425	0.458	0.033
$Var(ln(w_{ij}))$	0.195	0.197	0.002
$Var(ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{ij})$	0.185	0.192	0.007
$2Cov(w_{ij}, ln(a_{\boldsymbol{J}(i,t)}) + \epsilon_{ij})$	0.046	0.069	0.024
$2Cov(\alpha_i, ln(a_{\boldsymbol{J}(i,t)}))$	0.059	0.066	0.007
$2Cov(\psi_{\boldsymbol{J}(i,t)}, ln(a_{\boldsymbol{J}(i,t)}))$	-0.017	0.001	0.018

Notes: Variance of job value and covariances of components of job value 1996–2003 and 2004–2011, estimating the model with wages that include the expected wage growth over the coming 8 years.

growth coming from the opportunity for outside job offers with higher wages at firm k (compared to the job offers that arrive at firm j). However, the fact that the immediate wage growth after a job switch has little impact on results make it unlikely that expectations about wage growth in $ln(a)$ are important for the conclusions drawn in this article.

J Derivations for Section 4.2

J.1 Rent Dispersion Under Monopsony

In following I show how in a general equilibrium monopsony model rent dispersion can arise due to differences in labor productivity across firms. The following line of reasoning follows the exposition in Manning (2021).

Suppose that, in logs, the revenue of firm j can be written as:

$$y_j = z_j + (1-\eta)ln(g_j) - ln(1-\eta) \tag{A.20}$$

where y_j is log revenue and z_j is a shifter of the revenue function. η captures a parameter that is influenced by returns to scale in the production function and the elasticity of the product demand curve.

Firm j's labor supply in period t is given by:

$$\begin{aligned} g_{j,t} = & \underbrace{(1-\delta_j)g_{j,t-1}}_{\text{no job destruction}} + \underbrace{\bar{f}_j}_{\text{hires from non-employment}} \\ & + \sum_{k=1}^{J} \underbrace{\lambda_1(f_{jk}(1-\delta_k)g_{k,t-1}\Phi((\psi_j + ln(a_j)) - (\psi_k + ln(a_k))}_{\text{hires from employer } k} \\ & - \underbrace{\lambda_1 f_{kj}(1-\delta_j)g_{j,t-1}\Phi((\psi_k + ln(a_k)) - (\psi_j + ln(a_j)))}_{\text{quits to employer } k} \end{aligned} \tag{A.21}$$

that is, the labor supply to firm j in period t equals the share of its labor supply from the previous period that was not exogenously destroyed $(1-\delta_j)g_{j,t-1}$, plus its hires from non-employment $\bar{f}_j$, plus the net hires resulting from workers making voluntary employer-to-employer transitions. $\Phi()$ denotes the cumulative distribution function of $\mathrm{N}(0, 2\sigma^2)$. We see immediately that g_j is increasing in firm j's value offer $\psi_j + ln(a_j)$. Let's assume that the relationship between $ln(g_j)$ and $\psi_j + ln(a_j)$ is linear and thus write

$$ln(g_j) = \epsilon * (\psi_j + ln(a_j)) \tag{A.22}$$

where ϵ denotes the elasticity of firm j's labor supply with respect to its firm value offer.

The profit-maximizing value offer $\psi_j + ln(a_j)$ will equate marginal cost of one additional unit of labor g_j with the marginal revenue of one additional unit of labor. Thus, taking the derivative with respect to g_j in equation A.20 and combining it with the derivative with respect to $exp(\psi_j)$ in equation A.22, this can be written as:[84]

$$z_j - \eta ln(g_j) = ln(\frac{1+\epsilon}{\epsilon}) - \frac{1}{\epsilon} ln(g_j)$$

and thus, will imply the following level of employment:

$$ln(g_j) = \frac{\eta}{1+\epsilon\eta}(z_j - ln(\frac{1+\epsilon}{\epsilon})) \tag{A.23}$$

where we see immediately that for any η and $\epsilon > 0$, $\frac{\partial ln(g_j)}{\partial z_j} > 0$, and $\frac{\partial(\psi_j + ln(a_j))}{\partial z_j} > 0$ thus more productive firms will have more employees and offer greater value.

J.2 Effect of a Decrease in Labor Supply Elasticity on Rent Dispersion

Consider firms' labor supply given by equation A.22. More productive firms offer greater value, that is, $\frac{\partial(\psi_j + ln(a_j))}{\partial z_j} > 0$. Thus, a sufficient condition for an decrease in labor supply elasticity to increase rent dispersion across firms is that $\frac{\partial^2(\psi_j + ln(a_j))}{\partial\epsilon\partial z_j} < 0$. Note that:

$$\frac{\partial^2(\psi_j + ln(a_j))}{\partial\epsilon\partial z_j} = -\frac{\eta(1+2\epsilon\eta)}{(\epsilon(1+\epsilon\eta))^2} \tag{A.24}$$

which is strictly negative for any η and $\epsilon > 0$.

J.3 Effect of Increase in Search Frictions on Labor Supply Elasticity

Consider firms' labor supply given by equation A.21, and suppose in $t-1$ all firms are equally of size and make each other equally many offers ($g_{j,t-1} = 1$ and $f_{jk} = 1$ for all j and k), and that there is no job destruction and no hires from non-employment.

Then equation A.21 simplifies to

$$g = J\lambda_1 f(2\Phi((\psi_j + ln(a_j)) - (\psi_k + ln(a_k))) - 1)$$

Where J denotes the number of firms in the economy. I will omit the constants J and f in what follows. The labor supply elasticity then is

[84] Here, I assume that $exp(\psi_j)$ captures the full wage of a worker at firm j. Then, by the envelope theorem, the marginal cost of increasing ψ_j will equal the marginal cost of increasing total firm value $\psi_j + ln(a_j)$.

$$\frac{\partial ln(g)}{\partial g}\frac{\partial g}{\partial(\psi_j + ln(a_j))} = \frac{2}{g}\lambda_1\phi((\psi_j + ln(a_j)) - (\psi_k + ln(a_k))) \tag{A.25}$$

where $\phi()$ denotes the probability density function of $\mathrm{N}(0, 2\sigma^2)$.

An increase in search frictions is reflected by a decrease in the frequency with which workers are receiving offers λ_1. The impact of λ_1 on the labor supply elasticity is:

$$\frac{\partial \frac{\partial ln(g)}{\partial g}\frac{\partial g}{\partial(\psi_j + ln(a_j))}}{\partial \lambda_1} = \frac{2}{g}\phi((\psi_j + ln(a_j)) - (\psi_k + ln(a_k))) \tag{A.26}$$

Indeed, equation A.26 is strictly positive. Thus, when λ_1 decreases (more search frictions), the elasticity of labor supply decreases.

J.4 Effect of Decrease in Segregation on Labor Supply Elasticity

Consider firms' labor supply elasticity given by equation A.25. A decrease in segregation implies that workers receive offers from firms that are more different than their current one in terms of the value they offer, which means that $|(\psi_j + ln(a_j)) - (\psi_k + ln(a_k))|$ in equation A.25 increases. Substituting $(\psi_j + ln(a_j)) - (\psi_k + ln(a_k))$ with ΔV, we thus have:

$$\frac{\partial \frac{\partial ln(g)}{\partial g}\frac{\partial g}{\partial(\psi_j + ln(a_j))}}{\partial |\Delta V|}\Bigg| \Delta V \geq 0 = -\frac{2}{g}\lambda_1\phi(\Delta V)\frac{\Delta V}{\sqrt(2)\sigma} \tag{A.27}$$

and

$$\frac{\partial \frac{\partial ln(g)}{\partial g}\frac{\partial g}{\partial(\psi_j + ln(a_j))}}{\partial |\Delta V|}\Bigg| \Delta V < 0 = \frac{2}{g}\lambda_1\phi(\Delta V)\frac{\Delta V}{\sqrt(2)\sigma} \tag{A.28}$$

Where equation A.27 and A.28 are strictly negative. Thus, a decrease in segregation leads to a decrease in labor supply elasticity.

J.5 Effect of Increase in Idiosyncrasy of Preferences on Labor Supply Elasticity

Consider firms' labor supply elasticity given by equation A.25. An increase in the idiosyncrasy of workers' preferences is reflected in my model by an increase in σ. Thus, we are interested in:

$$\frac{\partial \frac{\partial ln(g)}{\partial g}\frac{\partial g}{\partial(\psi_j+ln(a_j))}}{\partial\sigma} = \frac{\sqrt{2}}{g}\lambda_1\phi((\psi_j+ln(a_j))-(\psi_k+ln(a_k)))\frac{(\psi_j+ln(a_j))-(\psi_k+ln(a_k))^2-2\sigma^2}{2\sigma^3} \quad \text{(A.29)}$$

which is negative if

$$2\sigma^2 > (\psi_j+ln(a_j))-(\psi_k+ln(a_k))^2 \quad \text{(A.30)}$$

I use the offer distribution for the 1996–2003 panel I estimate in Appendix K to evaluate whether it is plausible to expect that inequality A.30 holds, and an increase in the idiosyncrasy of preferences will decrease labor supply elasticity. While I find that inequality A.30 holds only for close to half of all offers, I still find that the derivative in equation A.29 is negative when summing over all offers, because $\phi(\psi_j+ln(a_j))-(\psi_k+ln(a_k))$ is much greater for offers with low $|(\psi_j+ln(a_j))-(\psi_k+ln(a_k))|$. I thus conclude that the increase in idiosyncrasy in preferences likely lead to a reduction of labor supply elasticity.[85]

J.6 Firms' Cost Minimization in Value Provision

Employers solve:

$$\min_{\psi_j,a_j} c_j(a_j)+exp(\bar{\alpha}_j+\psi_j) \qquad s.t.\ \ \psi_j+ln(a_j)=V_j^E$$

where $\bar{\alpha}_j$ denotes the average wage component net of the firm-specific component of every worker at firm j. The cost-minizing quantities of $ln(a_j)$ and ψ_j thus solve:

$$ln(a_j^*)=\frac{V_j^E-ln(c'(a_j^*))+\bar{\alpha}_j}{2} \qquad \psi_j^*=\frac{V_j^E+ln(c'(a_j^*))-\bar{\alpha}_j}{2}$$

meaning I can estimate the log of employers' marginal cost of non-wage value provision using

$$ln(c'(a_j^*))=\psi_j-ln(a_j)+\bar{\alpha}_j$$

and replacing the right-hand-side terms with the estimates from my search model and the AKM regression. Figure A.5 summarizes the marginal cost of non-wage value provision on the industry-level.

[85] An additional argument in favor of a decrease of labor supply elasticity is that measurement error in my empirical estimates inflates $(\psi_j+ln(a_j))-(\psi_k+ln(a_k))^2$.

K Estimating the Distribution of Offers

In my search model I assume that (see Assumption 1)

$$\sum_{k \in J} f_{jk} g_k = f_j^{NE} \quad \forall j \in J \tag{A.31}$$

that is, that I can estimate the total number of offers firm j makes equals the number of workers firm j hires from non-employment f_j^{NE}.[86] Moreover, I assume that

$$\frac{f_{jk}}{f_j^{NE}} = \frac{f_{kj}}{f_k^{NE}} \;\; \forall j, k \in J \tag{A.32}$$

that is, the ratio of intensities with which firm j and k make offers to each others' workers must be equal to the ratio of the intensities with which they hire from non-employment.

There does not necessarily exist a unique solution (unique vector $f_{(j*k\times 1)}$) to the system of equations implied by equations A.31 and A.32.

A feasible way of estimating f is

$$\min_{f_{(j*k\times 1)}} \sum_{j\in J} (\sum_{k\in J} f_{jk} g_k - f_j^{NE})^2 \;\text{ s.t. }\; \frac{f_{jk}}{f_j^{NE}} = \frac{f_{kj}}{f_k^{NE}} \;\; \forall j, k \in J \tag{A.33}$$

that is, to minimize the quadratic difference of the total number of offers made by firm j to employees of any other firm k, minus the total number of offers firm j has made, as estimated from the hires from non-employment, subject to equation A.32.

I solve the minimization problem equation A.33 separately for the 1996–2003 and the 2004–2011 sample.[87]. The resulting solution for the total number of offers implied by $\sum_{k\in J} f_{jk} g_k$ is remarkably close to f_j^{NE} (correlation of 1.00 in both sample periods).

[86] More precisely, I assume that $\sum_{k \in J} f_{jk} g_k = a * f_j^{NE} \quad \forall j \in J$, where I set, w.l.o.g., $a = 1$ in equation A.31 to ease exposition.

[87] I use the quadprog command in MATLAB. I only estimate f_{jk} if firm j and k are connected to each other by at least one employer-to-employer transition. Otherwise, I assume firm j and k are unconnected and thus set $f_{jk} = 0$

CHAPTER 2

IMPROVING MATCHING EFFICIENCY IN TWO-SIDED MARKETS: A MUTUAL POPULARITY RANKING APPROACH

Improving Matching Efficiency in Two-Sided Markets: A Mutual Popularity Ranking Approach *

Tobias Lehmann† Camille Terrier‡ Rafael Lalive §

October 2022

Abstract

Congestion is a widespread phenomenon in two-sided markets, but evidence on its costs and benefits is scarce. Using data from a dating platform, we document significant congestion, as measured by a large excess of demand for some women. By exploiting exogenous variation in the number of men and women using the dating app, we show that congestion slows down the matching time for men, but it benefits women by increasing their choice set. This asymmetric effect raises a natural question: Is there a way to reduce congestion without impending women matching outcomes? The second part of the paper builds on an intuitive answer. Men who *like* a woman who does not *like them* back generate congestion without improving women choice set. To identify these congestion-increasing contacts, we infer men and women "popularity" as their eigenvector centrality in the network of likes. We show that *likes* from men who are significantly less popular than a woman have a close-to-zero match probability, but they generate a congestion cost for other men in the market.

Keywords: Congestion, two-sided markets, online platforms, eigenvector centrality, revealed preferences.

JEL classification: D4, D47, D62, D83.

*We are grateful to Jérôme Adda, David Autor, Michele Belot, Bruno Crépon, Dominik Hangartner, John Horton, Thomas Le Barbanchon, Thierry Magnac, and Isaac Sorkin for useful comments and discussions. We also thank seminar participants at Bocconi University, ETH Zurich, LMU Munich, and conference participants at the CESifo Area Conference on Economics of Digitization, SSES conference, EEA conference, AMLD conference, and EALE conference for their helpful comments.

†University of Lausanne. Email: tobias.lehmann.1@unil.ch.

‡University of Lausanne. Email: camille.terrier@unil.ch.

§University of Lausanne. Email: rafael.lalive@unil.ch.

1 Introduction

From Tinder and match.com to Upwork, LinkedIn or university admission platforms, there has been a fast development of two-sided online markets in recent years. The transition from offline to online markets, by favoring market thickness and superstar effects (Rosen, 1981), has made congestion a widespread problem. Many markets suffer form an excess demand for some agents or objects that cannot be regulated by price variations (Roth, 2018).[1] On Upwork, a large freelancing platform, fewer than 10% of job applications get a response while over half of job openings remain unfilled (Horton, 2017). Congestion is also a concern in public sector markets. New York City has to match over 90,000 students a year to over 500 high school programs. The city tackled congestion by adopting a centralized assignment system in 2005 to get rid of the 30,000 students who remained unmatched every year (Abdulkadiroğlu et al., 2005). However, adopting a fully centralized solution is often not possible.

Then congestion becomes not only widespread, but also costly, especially when prices cannot play their regulatory role.[2] Congestion costs can take multiple forms. In labor markets, firms adopt strategic behaviors by not interviewing the top candidates considered as too unlikely to accept an offer (Roth and Xing, 1997). Many entry-level labor markets also suffer from unravelling, a process by which companies make offers to candidates as early as possible to avoid competition (Roth, 2008; Roth and Xing, 1994; Avery et al., 2001).[3] Despite this rich evidence on unravelling, there is surprisingly little empirical evidence on the causal effect of congestion on other outcomes, such as matching chances and matching time.

Yet, how much congestion affects matching chances and matching time is unclear. In two-sided platforms, congestion could have an asymmetric effect on the opposite sides of the market. For job seekers, fierce competition for a job might reduce their interview chances, but for recruiters more applications could mean increased matching chances and matching quality (Lazear et al., 2018; Peters, 2010). Congestion costs and benefits also depend on search and screening costs (Kanoria and Saban, 2021; Arteaga et al., 2022). Increasing the time each employer spends on an application could reduce the benefit of receiving numerous applications if employers no longer have time to go through all the applications. Despite these intuitions, we

[1] Roth (2018) defines congestion as "the accumulation of more time-consuming activities than can easily be accommodated in the time available." ([p.1613).

[2] Many markets do not rely on prices (or imperfectly do so) to regulate over-demand. School choice, social housing, day care allocation, or the dating market are only a few examples. Even in labor markets, Banfi and Villena-Roldan (2019) document that job ads with hidden wages “account for 86.6% of all job ads in www.trabajando.com, 75.2% in www.monster.com (Brenčič, 2012), 80% in www.careerbuilder.com (Marinescu and Wolthoff, 2020), and 83% in www.zhaopin.com, a Chinese online job board (Kuhn and Shen, 2012)”.

[3] Typical examples include the American market for new physicians (Roth, 1991), the market for American specialty residencies such as neurosurgery, or ophthalmology (Roth and Xing, 1994) or the entry-level labor market for Federal court clerkships in the U.S (Niederle and Roth, 2003).

lack empirical evidence on congestion costs and benefits and on how screening and application costs influence those.

The first objective of this paper is to fill this gap by bringing causal evidence on the costs and benefits that congestion generates in two-sided markets, and to quantify the role played by screening costs. We use novel data from a dating platform on which 8,700 heterosexual men and 4,200 heterosexual women check each other's profiles, like each other, and chat when they mutually like each other. We show that there is an excess demand for women, as measured by the number of likes women receive from men. Excess demand arises because there are twice as many men on the dating app, but also because men like women much more often than women like men.[4] By the time of their first connection to the dating app, women have received 61 likes from men on average, which is more than the 43 male profiles they check per day on average.

To quantify congestion costs and benefits, we use a unique feature of the dating app. Each new profile has to be approved by a moderator before being shown to other users. Because users creating profiles typically do not connect to the platform immediately after their profile is approved, this generates a period of time during which they can accumulate likes. Women typically accumulate large numbers of likes in that period, sometimes up to 300. We exploit this stockpiling of likes to investigate the costs and benefits it generates. Specifically, for each man m who likes a woman on the dating app, we estimate how much increasing the number of likes from other men affects (i) the probability that the woman sees the man m on her feed, (ii) the probability that she matches with him, and (iii) the time it takes for the match to happen. Our measure of congestion—how many men like a woman—is correlated with a woman's unobservable traits (like her physical attractiveness, education level, overall charisma), which might in turn determine her liking behavior. To address this endogeneity, we instrument the number of likes a woman receives with the quasi-random variation in the number of men who use the dating app in the 24 hours that follow the creation of the woman's profile.

We find large congestion costs for men, but also congestion benefits for women. Starting with congestion costs, for each man m who likes a woman w, increasing by 100 the number of likes from other men results in (i) a 23 point reduction in the probability that woman w sees man m on her feed (46.4% drop), (ii) a 2 point reduction in their matching chances (48.9% drop), and a twofold increase (1.92 days) in the time it takes to reach a match. On the other hand, our results suggest that some women benefit from receiving a large number of likes. Women with large screening costs do not benefit from congestion, but women with small screening costs, by screening quickly through each man's profile, largely benefit from receiving more likes from men. When receiving 100 additional likes, the share of a woman's feed that is composed of

[4] Men like 53% of the profiles they see, whereas women like 11% of the profiles.

men who like her increases by 37.9 percentage points.

All in all, our results confirm that, in two-sided markets, congestion can have an opposite effect on both sides of the market. This asymmetric effect raises a challenging issue: Policies that aim at reducing congestion could harm the side of the market that benefits from congestion. This typically happens if some of the men who refrain from liking a woman would have been liked by the woman. In contrast, reducing congestion would not hamper women's matching chances if the men who refrain from liking a woman are those she does not like. Said differently, to reduce congestion in a way that improves matching outcomes on both sides of the market, one would like to take into account the preferences of agents on both sides of the market. Yet, estimating agent mutual preferences is often difficult due to a lack of micro data on agent characteristics.[5]

The second contribution of this paper is therefore to develop a new ranking method, based on revealed-preference approaches, to estimate the preferences of agents on both sides of a market (men and women). Our ranking method adapts the eigenvector centrality to a two-sided environment. This centrality measure has been widely used to compute rankings in markets that are one-sided, for instance to compute journal, website, and firm rankings (Page et al., 1999; Palacios-Huerta and Volij, 2004; Sorkin, 2018).[6] However, the one-sided eigenvector centrality measure is not well suited to two-sided markets in which interactions happen *between two sides* of a market rather than *within one side*. We provide the first adaptation of the eigenvector centrality to a two-sided environment. In this setting, a woman who receives many likes from highly-ranked men receives a high popularity rank herself. Similarly, a man who receives many likes from highly-ranked women receives a high popularity rank himself.

We use our two-sided eigenvector centrality method to compute popularity rankings for men and women in the dating market. This additional information on male and female popularity helps to shed new light on the sources and costs of congestion. Our results show three interesting facts. First, men's preferences are highly correlated with those of other men. The probability that they like a woman increases with women's popularity. Importantly, this is true irrespective of the man's popularity, suggesting that men do not direct search based on their matching chances. Undirected likes result in many unpopular men liking women who are sig-

[5]Many two-sided platforms do not collect extensive information on their participants. The dating app we use collects information on users' age, gender, the number of Facebook friends, and preferences regarding age and gender, which only provides limited inputs for preference estimations. Similarly, Tinder only asks users to report gender and age, and preferences regarding age, gender, and geographic location. Scarcity of micro data also applies to two-sided online job boards. Glassdoor and CareerBuilder, for instance, allow job seekers to search and apply for jobs without providing any information apart from a job title and geographic preferences.

[6]In these applications, a web-page gets a high rank if it is linked to by many other highly-ranked web-pages, a journal gets a high rank if it is cited by a large number of other highly-ranked journals, and a firm gets a high rank if it receives many applications from other highly-ranked firms.

nificantly more popular than they are. Yet, the probability that these men are liked back by the popular woman is close to zero. To sum up, we identified men's likes that have no personal return.

We further show that shooting for the stars (unpopular men liking popular women) is not only ineffective, it also generates congestion for other more likely matches. To reach this conclusion, we estimate the congestion cost imposed specifically by over-shooting men, that is, by men who are less popular than the woman. As previously, we instrument the number of likes the woman receives from these men with the number of less-popular men using the dating app after a woman creates her profile. Our results show that increasing the number of likes from over-shooting men reduces the matching chances between the liked woman and other men who are more likely matches for her, i.e., men who are more popular than the woman. This last finding closes the circle. Limiting likes from over-shooting men would hardly reduce their matching outcomes, but it would increase the matching outcomes of other men with better matching chances by reducing congestion. All in all, the additional information we generate on the popularity of men and women allows us to identify ways to reduce congestion that would improve matching outcomes on both sides of the market. We conclude that there is room for two-sided Pareto-improving policy intervention.

This paper contributes to three main strands of literature. First, we bring novel causal evidence on congestion costs and benefits. Congestion in matching markets has been studied in laboratory experiments (Kagel and Roth, 2000), in the field (Roth and Xing, 1994, 1997; Roth, 2008; Avery et al., 2001), and more recently in online markets (Horton, 2019; Fradkin, 2017).[7] Yet, there has been surprisingly little empirical evidence until now on the causal effect of congestion on matching chances and matching time, especially in markets that are neither fully centralized nor fully reliant on prices, typically job markets and dating markets. In such markets, our empirical results, by shedding new light on the costs of bottlenecks and the role played by screening costs, are also of immediate interest to both public sector agencies (like unemployment agencies or university admission units) and private sector marketplaces (such as Tinder, match.com, Upwork, or LinkedIn).

Second, by developing a popularity ranking for two-sided markets that is based on eigenvector centrality, this paper contributes to recent developments in preference estimation techniques based on revealed preferences (Sorkin, 2018; Page et al., 1999; Palacios-Huerta and Volij, 2004). Hedonic models are often used for preference estimations (Rosen, 1986; Hwang

[7] Roth and Xing (1994, 1997), Roth (2008) and Avery et al. (2001) empirically studied unravelling in labor markets due to congestion. Kagel and Roth (2000) carried out related laboratory experiments. Fradkin (2017) has documented large reductions in the number of bookings on AirBnB when the initial contact made by searchers went to hosts who reject the offer.

et al., 1998; Hitsch et al., 2010; Chan and Wang, 2018).[8] However, the accuracy of these estimations strongly depends on the richness of the micro data available. Although the transition of many marketplaces from offline to online has increased the richness of available micro data, accurately estimating preferences through hedonic models is still often difficult (Marinescu and Wolthoff, 2020; Horton, 2017).[9] This makes it all the more important to design alternative preference estimation methods based on revealed preferences. In doing so, we directly build on the work of Sorkin (2018) by adapting eigenvector centrality to a two-sided environment.

Finally, our conclusions enrich an active research field that studies congestion-related policies. The popularity ranking we design can help improve two types of policies. First, policies that aim at reducing congestion by limiting the number of applications. Typical examples of such policies include the adoption of application costs (He and Magnac, 2020; Arnosti et al., 2021), signaling (Coles et al., 2010, 2013; Lee and Niederle, 2015), restrictions to choices and actions (Halaburda et al., 2018; Kanoria and Saban, 2021), or information provision (Belot et al., 2019; Gee, 2019; Bhole et al., 2021; Arteaga et al., 2022). Second, the popularity ranking we design can guide policies that help the congested side (e.g., recruiters, women, universities) screen through the numerous applications received (Horton, 2017).[10] These two types of policies can effectively reduce applications from men, job seekers, or students, but without considering the matching chances between the discouraged person and the targeted woman, job, or university. In contrast, the mutual popularity ranking we design, by accounting for preferences on both sides of the market, is tailored to improve matching efficiency on both sides.

The rest of the paper is organized as follows. The next section presents the dating platform. Section 3 reports descriptive statistics on congestion. Section 4 introduces the research design we use to estimate congestion costs and benefits (4.1) and presents the results (4.2). We introduce the two-sided eigenvector centrality method in section 5, and we estimate men and women popularity in Section 5.2. Finally, we check how well the popularity rankings predict matching probability and congestion costs in Section 6, before concluding in section 7.

[8]Several papers estimate preferences based on individuals' characteristics in the dating market (e.g., Wong, 2003; Choo and Siow, 2006; Hitsch et al., 2010) and in the labor market (e.g., Chan and Wang, 2018; Combe et al., forthcoming; Hangartner et al., 2021).

[9]Marinescu and Wolthoff (2020) show that estimates of workers' preference for wage are wrong signed, unless job-title fixed effects are included in estimations. This stresses a common limitation of hedonic preference estimations: Much of the relevant information is only available in unstructured text format, from which obtaining meaningful structured data is not straightforward. Horton (2017) show that recommending job seekers to employers based on hedonic preference estimates has no effect on employer behavior and vacancy outcomes in non-technical jobs.

[10]Many platforms and recruiters use algorithms to rank applicants based on match probability. Horton (2017) finds that employers that received algorithmic recruiting recommendations had a 20% higher fill rate compared to the control.

2 A Two-Sided Dating Platform

We use data from a Swiss dating app that is very similar to Tinder. Users create a profile using their Facebook login. The app sources the user name, age, and sex from Facebook.[11] Users also add pictures, introduce themselves in a few lines[12], and they can specify their preferences for a partner's sex, age, and geographical location. Men and women browse profiles that appear on their smartphone (see Figure 1a and 1b). When a user likes a profile, he presses the "HI" button. When he does not like the profile, he presses the "BYE" button. These two options are identical to the right and left swipe on Tinder. A user has to like or dislike a profile before she or he can see the next profile. When a man and a woman mutually like each other, they form a match (illustrated in Figure 1c) and they can start chatting (Figure 1d). There is no limit on the number of profiles a user can browse or like.

The dating app has a unique feature. Users cannot start browsing other users' profiles right after they create their profile. To filter out fake profiles, a moderator verifies each profile and validates it before a user can start using the app.[13] After a profile is approved, it is posted online, and all users can start liking it. Users do not know how long the approval process will take, so their first connection often happens several hours after their profile has been approved.[14] This waiting period is very useful to study congestion as it allows the number of likes to pile up before the first connection of a user, a feature we will exploit for our research design.

The accumulation of likes before a user's first connection is amplified by the boost of visibility the platform gives to newly created profiles. The dating platform uses a recommendation algorithm that determines which profiles users see first (Schaffner, 2016). It shows first profiles that match the criteria selected by each user, profiles that have liked the user, profiles that have been liked by a large share of other users, those that have recently used the platform, and those that are geographically close to the user. The ordering process is done each time a user opens the app. Importantly for us, the fact that the platform boosts the visibility of newly created profiles often leads to a large accumulation of likes before a user logs in for the first time.

[11] The app also imports the list of Facebook friends, the schools a user has attended, the Facebook objects a user is interested in, and the places a user has marked as visited on Facebook. However, this information was neither made public on user profiles nor used in the matching algorithm. The app was active between July 2013 to February 2017.

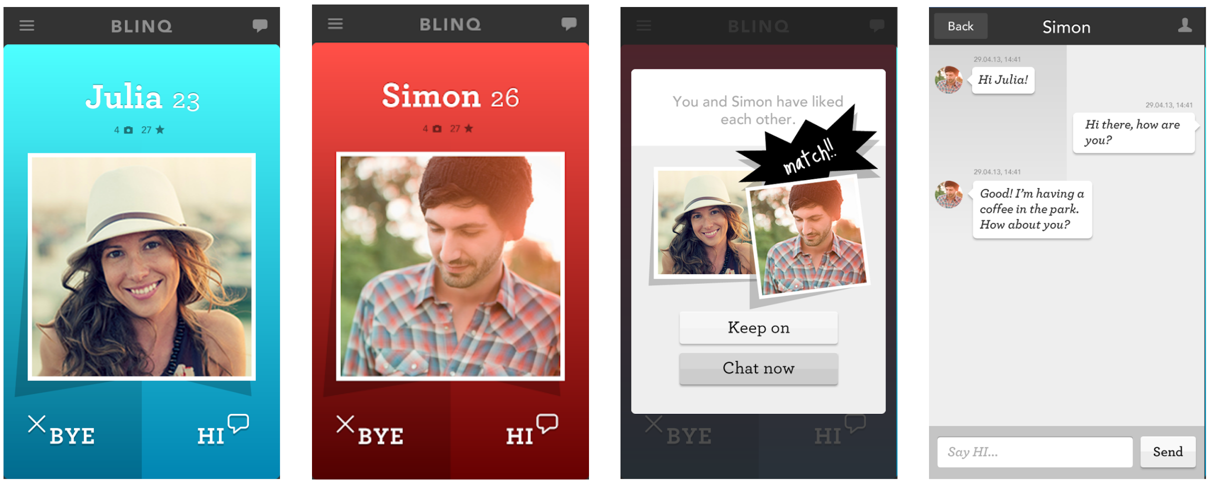
[12] E.g., "I play volleyball, hang out with friends, I love cats", "I am a sports addict and adventurer" or "Love exploring, Passionate about trucks and beer. 1.69m".

[13] Unfortunately, we do not have information on how long it takes for each profile to be approved.

[14] The approval process and the resulting waiting period do not exist on other platforms, including Tinder, where users can start swiping right after they create their profile.

Figure 1: Illustration of the dating platform

(a) (b) (c) (d)



Notes: This Figure illustrates the matching process on the dating app: (a) and (b) show men and women profiles, and how they can like a profile by clicking "HI" or not like it by clicking "BYE". When a man and woman like each other they form a match (c) which gives them the opportunity to chat (d).

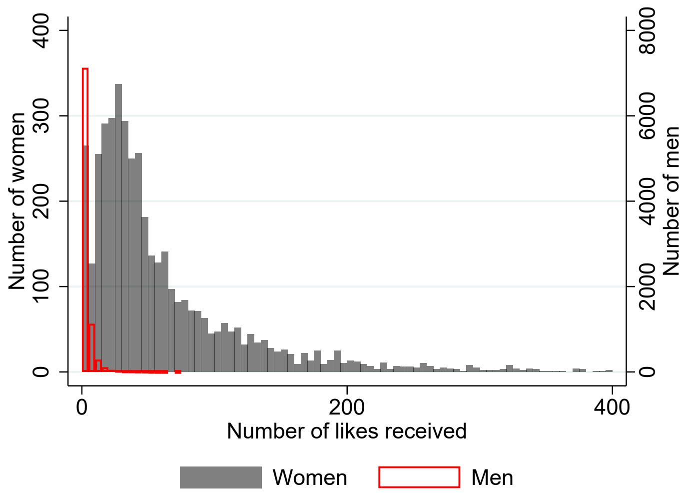
3 Descriptive Statistics

First signs of congestion on the dating platform. Table 1 compares the characteristics, preferences, and activity of men and women who used the dating app between January 2014 and December 2015, the period for which we have data. A few striking differences emerge. First, there are twice as many men as women using the app (8,788 men versus 4,238 women). This creates a large imbalance between the two sides of the market. Men are also significantly more active on the app. On average, they log in 1.8 times a day, versus 1.6 times for women, and when they connect, men check on average 54.3 women profiles per day when women only check 43.9 profiles. In addition to being more active, men are also five times more likely than women to like the profiles they see (53.0% for men versus 11.1% for women). This large difference in liking behavior implies that, by the time of their first login to the platform, women have received 61.5 likes from men on average, while men have only received 3.0 likes.

A second interesting fact emerges from the data. Because the dating app boosts the visibility of newly created profiles, men and women receive most of their likes in the days that follow their profile creation. They receive about half of their likes within the first week after they create a profile (see Table 1, Panel D and Figure A.1).

All in all, these statistics show that congestion, as measured by an excess demand for some women, affects women much more than men, and that it is particularly prevalent in the first days that follow a profile creation. This motivates our definition of congestion: We use the number of likes a woman receives between her profile creation and her first connection.[15] It takes 10.7 hours for the median woman to connect for the first time after she creates a profile (see Table 1, Panel B). During part of that period, a woman's profile is online and it starts accumulating likes from men. Figure 2 shows that there is a large variation across women in the number of likes they receive between their profile creation and their first connection. We will use this variation to analyse the effect of increasing the number of likes on matching outcomes for men and women. Naturally, there are several reasons why the number of likes a woman receives might be correlated with her unobservable traits and attractiveness. Our research design addresses this endogeneity in Section 4.

Figure 2: Number of likes received between profile creation and first login



Notes: This Figure shows the distribution of the number of likes women and men have accumulated by the time of their first connection to the dating platform. Bars represent frequencies by 5-like wide bins. The distributions are plotted for the 4,238 women in our sample (who receive on average 61.5 likes), and the 8,788 men in our sample (who receive on average 3.0 likes).

[15] To make sure that congestion has time to build up, in our analysis we discard women who log in within the first hour after they created a profile and women whose first connection happens more than one week after the profile creation (or who never log in to the platform). These women receive a very large number of likes, a substantial number of which are from men that have already found another match on the platform by the time the woman connects for the first time. These two restrictions respectively drop 773 and 332 women from the sample (14% and 6% of the sample). Table A.1 shows that they have almost no effect on the characteristics of the women we consider.

Table 1: Descriptive statistics on women and men

	All (1)	**Women** (2)	**Men** (3)	Difference (4)=(3)-(2)
A. User Characteristics				
Age	28.1	27.2	28.6	1.4***
Profile contains text (%)	0.118	0.091	0.131	0.040***
User has set age filter	0.171	0.162	0.176	0.014*
Number of Facebook friends	562.0	499.7	592.0	92.3***
B. User Activity				
Number of hours btw profile creation and first login	14.3	10.7	16.7	5.9***
Number of logins per day	1.7	1.6	1.8	0.2***
Number of profiles seen per day	50.9	43.9	54.3	10.4***
Number of minutes active per login	4.3	5.1	4.0	-1.1***
Seconds spent per profile	5.3	6.6	4.7	-1.8***
C. Preferences				
Share of profiles liked	0.393	0.111	0.530	0.419***
Share profiles matched	0.025	0.046	0.015	-0.031***
Share of profiles with chat	0.003	0.005	0.002	-0.004***
D. User Popularity				
Share of likes received in first week	0.473	0.515	0.453	-0.061***
Number of likes received (at first login)	22.0	61.5	3.0	-58.5***
Share of users with 0 likes (at first login)	0.175	0.027	0.246	0.219***
E. Metrics on Congestion Costs and Benefits				
Prob. of seeing partner who liked own profile	0.764	0.558	0.893	0.335***
Prob. of liking a partner who liked own profile	0.264	0.062	0.391	0.329***
Mean days to match	8.3	11.2	6.2	-5.1***
Share of matches happening within first month	0.958	0.933	0.968	0.035
Mean days to match | match within first month	1.165	1.434	0.973	-0.460***
Number of individuals	13,026	4,238	8,788	
Number of likes received (at first login)	286,563	260,434	26,129	

Notes: This table shows descriptive statistics on men and women who created a profile between January 1st 2014 and December 31st 2015, and who logged in for the first time more than an hour after profile creation but less than a week after profile creation. The variable *Days to match* represents the time elapsed between when a user receives a like and when he likes back. For *Number of hours btw profile creation and first login* we report the median. All variables in Panel D. are calculated based on activity in the first 2 weeks after the first login. *Share of likes received in first week* shows, out of all likes received within 10 weeks after profile creation, the share received within the first week after profile creation. Column (4) shows the mean difference between column (3) and column (2). Stars indicate significance as follows: *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

First signs of congestion costs. Table 1 also brings suggestive evidence of congestion costs. The statistics reported in Panel B show that women check on average 43.9 profiles per day. This is significantly less than the average number of likes they receive by the time of their first login (61.5). Women may not have time to check all the profiles of the men who have liked them, especially as women tend to spend more time on a profile (6.6 seconds) than men (4.7 seconds). These first signs of congestion costs are consistent with the statistics we report in Panel E. When they login to the app, women are 33.5 percentage points less likely than men to see the profile of a man who liked her (55.8% for women versus 89.3% for men). Figure 3a confirms that this might be due to congestion. The larger the number of likes a woman receives (x-axis), the lower the chances that she will see the profile of a man who liked her (y-axis).

First signs of congestion benefits. If congestion seems costly for men, there are reasons to believe that, on the other hand, it might benefit women for whom receiving more likes can result in higher match probability. This is especially true when screening costs are limited, that is, when women are able to check the profiles of all the men who have liked them. Figure 3b confirms this intuition by plotting the share of a woman's feed (at first login) that is composed of men who previously liked the woman. This share constantly increases with the number of likes a woman receives.[16,17]

4 Empirical Evidence on Congestion Costs and Benefits

4.1 Research design

Endogeneity. A key difficulty in analyzing the effect of congestion is the non-random number of likes that women receive. Some female traits which we do not observe (such as attractiveness, education level, charisma, etc.) are likely to drive both the number of likes a woman receives and the outcomes we are interested in, such as the chances that a woman likes a man and matches with him. Attractive women might be more picky when it comes to liking men on the app. Another source of endogeneity is the time it takes women to log in for the first time after they create a profile. This is the period over which we count the number of likes a woman receives; our measure of congestion. The longer a woman waits, the larger the number of likes she mechanically accumulates. But waiting time can also be a signal of how keen a woman is

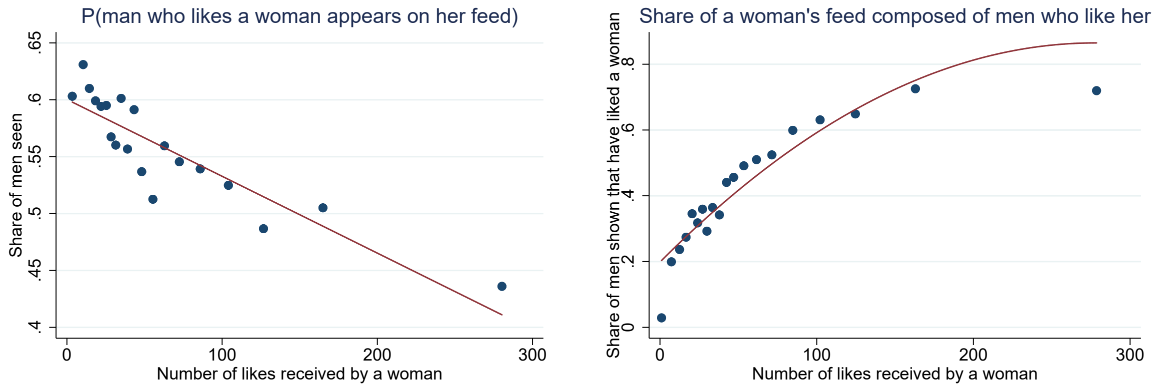
[16] We define a first connection as the first session without any interruption that lasts more than 15 minutes. If an interruption lasts more than 15 minutes, we define it as a second connection. We consider all profiles a woman sees during that first connection.

[17] Figure A.2 shows that the same conclusion applies when we consider men instead of women, i.e., when we plot the share of a man's feed (at first login) that is composed of women who previously liked the man.

to find a match. Women who log in quickly after creating a profile might spend more time on the platform and like more men.

Figure 3: Descriptive evidence on costs and benefits of congestion

(a) Congestion costs (for men) (b) Congestion benefits (for women)



Notes: Panel (a) shows the probability that a man who liked a woman before her first login ever appears on the feed of the woman (y-axis), as a function of the number of likes the woman receives before her first login (x-axis). Panel (b) reports the share of a woman's feed at her first login that is composed of men who have liked her (y-axis), as a function of the number of likes the woman receives before her first login. The figures are binned scatter plots, with women assigned to 20 equally sized bins by the number of likes they receive before their first login. The red line represents the regression line from a linear (panel a) / quadratic (panel b) regression of the y-axis variable on the number of likes the woman receives before her first login. Panel (b) is based on all women in the sample (column (2) of Table 1), and panel (a) is based on the 4,122 women in column (2) of Table 1 who receive at least one like before their first login.

Instrumental variable research design. To deal with the endogeneity of congestion, we exploit quasi-random variation in the number of men who use the dating app during the 24 hours that follow the creation of the woman's profile. Figures A.3 and A.4 show that this number of men varies quite substantially across women, and that the variation comes primarily from day-to-day variation in the number of users. We exploit this variation by instrumenting the number of likes a woman j receives (noted L_j) with the number of men using the dating app in the 24h after woman j creates a profile (noted M_j).

The second-stage equation of our IV research design is:

$$Y_{ij} = \alpha + \beta L_j + \gamma X_{ij} + \epsilon_{ij} \quad (1)$$

where Y_{ij} is the outcome of interest (for instance the probability that woman j sees man i on her feed), L_j is the number of likes that woman j received before her first login, and X_{ij} is a vector of control variables.[18] ϵ_{ij} is an error term that reflects the influence of the unobserved characteristics of woman i and man j on the outcome. β identifies the causal effect of congestion.

The first stage for this two-stage least squares (2SLS) procedure is:

$$L_j = \theta + \eta M_j + \gamma X_{ij} + \epsilon_j \tag{2}$$

where M_j is the number of men using the app in the 24h after woman j creates a profile, and L_j is the number of likes that woman j receives before her first login. X_{ij} contains the same control variables as in Equation 1, and ϵ_j is an error term that captures idiosyncratic shocks affecting the number of men using the dating app. η is our first stage coefficient of interest. It indicates how much a change in the number of men using the app in the 24h after woman j creates a profile affects the number of likes she receives. Columns 1 and 2 of Table 2 show that this correlation is large and significant. Increasing the number of men who use the app by 100 raises the number of likes a woman gets by 10.4.[19]

Outcomes of interest. We build three outcomes to examine the costs of congestion for men. For each man i who likes a woman j, we estimate how much the number of likes that woman j receives *from other men* affects (i) the probability that woman j sees the man i on her feed, (ii) the probability that woman j matches with the man i, and (iii) the time it takes for woman j and man i to match.

We proceed similarly to measure the potential benefits of congestion for women. For each woman i in our sample, our outcome of interest is the share of woman i's feed (at first login) that is composed of men who previously liked woman i. We then estimate how much the number of likes that woman i receives from men affects the above outcome.

Identifying assumption. Our instrumental variable methodology relies on the assumption that the number of men using the dating app the day after a woman creates her account is independent of that woman's unobserved characteristics, especially the characteristics that might affect the outcomes like her chances of liking a man. Although this assumption is not empirically testable, we check if women's observable characteristics are correlated with the number

[18]Control variables are woman j's age, whether she defined an age filter, whether her profile contains text, and her number of Facebook friends, and man i's age, whether he defined an age filter, whether his profile contains text, and his number of Facebook friends. We show that all our results are very similar when controlling for these variables, a subset of them, or none of them.

[19]The first-stage F-statistics for these estimates are larger than the rule-of-thumb threshold of 10 commonly used to diagnose weak instruments.

of men using the app the day after a woman creates her account. When running this balance test, we are particularly interested in the characteristics of a woman that could reflect her attractiveness (for instance her age and number of Facebook friends) or her dating preferences and eagerness to find a match (for instance whether she specified an age filter, and whether she presents herself in her profile).

Table 2: First stage regressions and test of instrument independence

	First stage		Test of instrument independence			
	# likes received at 1st login		# of friends on Facebook	Age	Has age filter set	Has profile text
	(1)	(2)	(3)	(4)	(5)	(6)
Nb men using app 24h after woman profile creation	0.1037***	0.0538***	-0.0167	0.0020***	-0.0000	0.0001***
	(0.0141)	(0.0112)	(0.0276)	(0.0005)	(0.0000)	(0.0000)
Observations	260,434	10,990	4,122	4,122	4,122	4,122
F-statistic	53.87	22.90	0.34	14.54	0.11	12.72
R^2	0.04	0.02	0.00	0.00	0.00	0.00
Second stage's depvar	P(being seen), P(being liked)	Days to match				

Notes: This table shows, in columns (1) and (2), the first stage estimated using Equation 2. We regress the number of likes that woman j receives before her first login (L_j) on the number of men using the app in the 24h after woman j creates a profile (M_j). We report coefficients from the specification without any control variables, but obtain almost identical coefficients when we control for woman j's and man i's age, whether she/he defined an age filter, whether her/his profile contains text, and her/his number of Facebook friends, or a subset of these covariates. We cluster standard errors at the woman level. The corresponding second stage results are reported in Table 3, Panel A. Columns (3) to (6) test the instrument independence assumption. We show coefficients from bivariate regressions of woman characteristics on the instrumental variable, i.e., the number of men using the app in the 24h after a woman creates a profile (M_j). *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Columns 3 to 6 of Table 2 shows that most of these variables are unrelated to the number of men using the platform.[20] Women who have more Facebook friends and women who set up an age filter do not create accounts in periods when more men use the dating app. We find a small correlation with women's age, suggesting that slightly older women tend to create accounts when more men use the app. However, the magnitude of the coefficient is very small—an additional 100 men using the app would increase women's age from 27.2 to 27.4—and most importantly, women's age does not predict whether a man appears on a woman's feed (see Table A.2).[21,22] These two arguments also hold for the small correlation we find between the

[20]In Table A.3 we show that the number of men using the platform is also urnrelated to the number of profiles women check at their first login and the time women take from profile creation to their first login.

[21]This last fact explains why our results on congestion costs and benefits remain the same when we control for women's age (see Table 3).

[22]We performed a second test to show that the magnitudes are small. We regressed the number of likes received at first login on the four women's observable characteristics (number of Facebook friends, age, whether the woman has set an age filter, and whether she has a profile text). We then regress the fitted value from this regression on the instrument. The magnitudes of the correlation between the instrument and women observable characteristics (more specifically how these characteristics predict the number of likes) are therefore in interpretable units. We

number of men using the app and whether a woman presents herself in her profile.

4.2 Empirical results on congestion costs

Men's probability of being seen by a woman. Table 3 reports our estimates of congestion costs. We focus the discussion on the 2SLS coefficients reported in columns 2, 4, and 6. First, when a man i likes a woman j, he has 51% chances of appearing on her feed, hence to be seen by the woman. The estimates in column 2 show that increasing the number of likes that woman j receives from other men by 100 reduces the probability that she sees the man i on her feed by 23 percentage points. This corresponds to a 46.4% drop.

Another way to read this result is to calculate the number of men it takes for one man who likes a woman to no longer be seen by that woman. It takes four additional men liking a woman for a like not to be seen.[23] The fact that only four additional men are enough to eliminate any chances of a man matching with a woman indicates that congestion costs for men are substantial on the dating app.

Men's probability of being liked by a woman. We move to our next outcome, the probability that a man who likes a woman is liked back by that woman. On average, 4.6% of the men who like a woman are liked in return by the woman. Said differently, 4.6% of the men who like a woman match with her. Again, congestion significantly reduces these matching chances. The coefficient reported in column 4 shows that increasing the number of men who like a woman by 100 leads to a 2.2 percentage point reduction in matching chances, which corresponds to a 48.9% drop.

Part of this effect mechanically stems from the effect of congestion on the probability of being seen by a woman discussed in the previous paragraph. The probability of being liked back is the combined effect of the chances of being seen by a woman times the share of men the woman likes. The latter effect captures the effect of a greater choice set on women's selectivity. Comparing the coefficient in column 4 with the coefficient in column 2 suggests that the effect of congestion on the probability of being seen passes through almost 1:1 to the probability of being liked back, meaning that women hardly change their selectivity when they get more likes.[24]

find that each additional man using the app in the first 24h after a woman creates her profile only changes the predicted number of likes a woman receives by 0.0011.

[23] On average, when a man likes a woman, there are 113 other men who also like that woman. When four additional men like this woman, this woman will see on average 4*(-0.2344/100)*113=1 fewer men.

[24] To see this, take the ratio of the coefficient in column 4 over the coefficient in column 2, which is $\frac{0.0224}{0.2344} = 0.10$. This is almost identical to the ratio of the variables indicating whether a man liked back and whether a man is seen, $\frac{0.062}{0.558} = 0.11$

Matching time. The results reported in column 6 show that increasing the number of likes a woman receives by 100 raises the time it takes to reach a match by 1.92 days. Given that it takes 1.78 days on average for men and women to match on the app, the congestion effect more than doubles the matching time.[25]

For all outcomes, the coefficients estimated using OLS are smaller than those estimated using 2SLS, which reflects selection bias. We underestimate the cost of congestion when we naively correlate the number of likes a woman receives and the probability that a man is seen by the woman, probably because more popular women enjoy using the platform more, spend more time on the platform, and therefore have higher chances of seeing men who liked them.

Robustness checks. We run several robustness checks to verify how stable our estimates are across specifications. First, the results reported in Panel B of Table 3 control for women age and their number of Facebook friends. The coefficients are almost identical. In Panel C, we report estimates that further control for whether a woman defined an age filter, and whether her profile contains text. Again, the estimates are mostly unaffected.

Finally, in Panel D, we further control for the characteristics of the men who use the dating app in the 24h that follow a woman profile creation. The characteristics we control for include men age, whether they defined an age filter, whether their profile contains text, and their number of Facebook friends. These controls are important as Table A.4 suggests that some men characteristics change slightly when the number of men using the app increases. However, the magnitudes of these correlations are small, so that changes in men characteristics are unlikely to drive our results.[26] The results we report in Panel D confirm that controlling for men characteristics leaves our estimates of congestion costs unchanged.

[25]For this outcome, we only consider matches that happen within 30 days of a woman creating her profile. This represents 90% of the matches (see Panel D of Table 1). A few outliers drive the regression results when we do not restrict the matching time window. We tried different time frames between 14 days after profile creation and 60 days after profile creation. We obtained similar results for all our analyses that incorporate the "days to match" variable.

[26]A hundredfold increase in the number of men using the dating app leads to a drop in the number of Facebook friends of 7 (which represents a 0.9% reduction based on the 592 friends men have), an increase in men average age of 0.3 years, and a 0.03 and 0.01 point increase in the probability that men have set up an age filter and that their profile contains a descriptive text.

Table 3: Congestion costs for men

	P(being seen)		P(being liked)		Days to match	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No controls						
Number of likes	-0.0627***	-0.2344***	-0.0122***	-0.0224***	0.0117***	0.0192***
received	(0.0061)	(0.0324)	(0.0009)	(0.0051)	(0.0012)	(0.0063)
Panel B: Controlling for woman's age and nb of Facebook friends						
Number of likes	-0.0620***	-0.2306***	-0.0125***	-0.0224***	0.0116***	0.0193***
received	(0.0062)	(0.0314)	(0.0010)	(0.0050)	(0.0012)	(0.0062)
Panel C: Controls from Panel B, and whether woman defined an age filter and a profile text						
Number of likes	-0.0629***	-0.2295***	-0.0125***	-0.0225***	0.0115***	0.0194***
received	(0.0059)	(0.0312)	(0.0010)	(0.0050)	(0.0012)	(0.0064)
Panel D: Controls from Panel C, and controlling for same characteristics of the man liking the woman						
Number of likes	-0.0624***	-0.2143***	-0.0124***	-0.0204***	0.0115***	0.0190***
received	(0.0059)	(0.0299)	(0.0010)	(0.0049)	(0.0012)	(0.0065)
Observations	260,434	260,434	260,434	260,434	10,990	10,990
Mean. Dep.var	51.00	51.00	4.58	4.58	1.78	1.78
# of women	4,122	4,122	4,122	4,122	2,848	2,848
First stage F		53.87		53.87		21.79

Notes: This table reports the β coefficients from the following regression: $Y_{ij} = \alpha + \beta L_j + \gamma X_j + \epsilon_{ij}$ (Equation 1), where L_j denotes the number of likes a woman receives between when she creates her profile and when she logs in to the platform for the first time. $\mathbf{X_j}$ denotes a vector of control variables. Column (1), (3) and (5) report coefficients from an OLS regression. In column (2), (4) and (6), the number of likes is instrumented by the number of men using the platform in the first 24h after a woman creates her profile. The unit of observation is the like of a man for a woman. The dependent variable Y_{ij} is, in columns 1-2, an indicator for whether a man who liked a woman ever appears on the woman's profile feed, in columns 3-4 an indicator for whether a man who liked a woman is liked back by the woman, and in columns 5-6 the days elapsed between when a woman receives a like and when she likes the man back. We exclude matches that happen later than 30 days after a woman created her profile in columns 5-6. Control variables in Panel D include all control variables from Panel C, plus a control for the age of the man liking the woman, his number of Facebook friends, whether the man has defined an age filter, and whether the man's profile contains text. First stage results are reported in Table 2. First stage F corresponds to the lowest Kleibergen-Paap F-statistic of the four specifications in Panel A-D. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

4.3 Empirical results on congestion benefits

We show in the previous section that congestion is costly for men. However, congestion might, on the other hand, benefit women for whom receiving more likes from men can result in higher match probability. This is what we test next. Table 4 reports our estimate of the effect of increasing congestion—that is, the number of likes a woman receives—on the *share* of profiles shown to that woman that are from men who liked the woman. As before, we instrument the total number of likes a woman receives by the number of men using the platform in the 24h after the woman creates her profile. The 2SLS coefficients reported in column 2 are close to zero, and the precision of the estimate rules out substantial congestion benefits for women. This stands in contrast with the OLS estimates in column 1, which confirms once more the selection bias that OLS estimates suffer from.

The overall absence of congestion benefits masks heterogeneous effects. We have shown before that women receive a large number of likes before their first login (on average 61.4), so an increase in the number of men who like a woman might not have any effect on many women who are not able to see all the men who like them. However, women who screen through profiles particularly fast might benefit from having additional men liking them. In other words, the congestion benefits would be larger when the screening costs faced by women are limited. To test this, we split the sample of women in two groups based on the time they spend on men profiles. Women who spend less than the median time of 7.6 seconds are considered to have small screening costs, while women who spend more than the median time have large screening costs.

We present the results for these two groups separately in columns 4 and 6 of Table 4. A clear difference in congestion benefits emerges. As expected, women with large screening costs do not benefit from receiving more likes from men. Women with small screening costs, on the other hand, by screening quickly through men profiles, largely benefit from receiving more likes from men. When receiving 100 additional likes, the share of profiles a woman sees that are from men who like her increases by 37.9 percentage points.

Finally, we conducted the same analyses for men. As men face much lower demand, we expect them to experience substantial benefits from an increase in the number of women liking them. Table A.5 shows that this is indeed the case. Men substantially benefit from having more women liking them.

4.4 Designing two-sided welfare-improving congestion policies

All in all, our results show that, in two-sided markets, congestion can have an opposite effect on both sides of the market. By generating tougher competition, congestion is costly for the congestion-generating side (men in our case), but it benefits the congested side (women in our case) by increasing their choice set, especially when screening costs are limited. This asymmetric effect implies that policies that aim at reducing congestion to limit its costs are likely to harm the side of the market that benefits from congestion. It is simple to see why using the dating market as an example. Some of the men who are discouraged from liking a woman would potentially have been liked by the woman, whose matching chances go down when her preferred partner no longer likes her.

This type of welfare-reducing outcome for the side that benefits from congestion (women) can typically happen when congestion-reducing policies are designed to improve the welfare of the congestion-generating side only (men). These are policies that aim at reducing congestion

Table 4: Congestion benefits for women

	All women		Women with high screening costs		Women with low screening costs	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No controls						
Number of likes	0.2185***	-0.0257	0.2214***	-0.0934	0.2014***	0.3787***
received	(0.0075)	(0.0589)	(0.0100)	(0.0741)	(0.0100)	(0.0795)
Panel B: Controlling for age and # of Facebook friends						
Number of likes	0.2184***	-0.0109	0.2200***	-0.0530	0.1998***	0.3870***
received	(0.0076)	(0.0562)	(0.0101)	(0.0650)	(0.0100)	(0.0788)
Panel C: Controls from Panel B, and whether woman defined an age filter and a profile text						
Number of likes	0.2187***	-0.0109	0.2203***	-0.0516	0.2001***	0.3809***
received	(0.0076)	(0.0563)	(0.0101)	(0.0654)	(0.0100)	(0.0782)
Panel D: Controls from Panel C, and controlling for same characteristics of the man liking the woman						
Number of likes	0.1859***	-0.0061	0.1940***	0.0187	0.1667***	0.3390***
received	(0.0077)	(0.0617)	(0.0101)	(0.0620)	(0.0105)	(0.0988)
Observations	4,238	4,238	1,987	1,987	1,987	1,987
Mean. Dep.var	42.24	42.24	49.53	49.53	30.46	30.46
First stage F		88.15		55.88		35.63

Notes: This table reports the β coefficients from the following regression: $Y_j = \alpha + \beta L_j + \gamma \mathbf{X_j} + \epsilon_j$, where Y_j is the share of the profiles appearing on a woman's feed that are from men who have liked her, at the time when the woman logs in to the platform for the first time. L_j denotes the number of likes a woman receives between her profile creation and her first connection to the platform. $\mathbf{X_j}$ is a vector of control variables. Columns (1), (3) and (5) report coefficients from an OLS regression. In columns (2), (4) and (6), the number of likes is instrumented by the number of men using the platform in the first 24h after a woman creates her profile. The unit of observation is a woman. The regressions in columns (1) and (2) are estimated on the full sample of women. Columns (3) and (4) show regression results for the subsample of women who take more than the median time to evaluate a profile ($>$ 7.2 seconds). Columns (5) and (6) show regression results for the subsample of women who take less than the median time to evaluate a profile ($\leq$ 7.2 seconds). For 264 women we do not know the time they take to evaluate a profile, because those women never evaluate more than one profile per session. Control variables in Panel D include all control variables from Panel C, plus a control for the average age of men liking the woman, those men's average number of Facebook friends, the share of men who have defined an age filter, and the share of those men whose profile contains text. First stage F corresponds to the lowest F-statistic of the four specifications in Panel A-D. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

costs without considering the potential benefits of congestion for the congested side. For instance, introducing application costs (He and Magnac, 2020; Arnosti et al., 2021), encouraging agents to signal their top preferences (Coles et al., 2010; Lee and Niederle, 2015), giving agents information on congestion (Gee, 2019; Bhole et al., 2021), or encouraging them to diversify their preferences (Belot et al., 2019) are policies that aim at reducing congestion by acting on the preferences of the agents on the congestion-generating side of the market.

Reducing congestion without harming the benefiting side is possible if one adopts a solution that accounts for preferences of both sides of the market. Again, the intuition is simple using the dating market as an example: The men we discourage from liking a woman should be the men a woman does not like. In other words, policies would discourage congestion-generating agents by acting not only on their preferences (e.g., how much a man likes a woman), but also on the preferences of agents on the other side (e.g., the probability that the man is liked back by a woman). Under such policies, the matching chances would not go down for the congested side (women), but they should improve for the congestion-generating side (men) as congestion goes down. Pareto-wise speaking, the match would be more efficient for both sides of the market.

Reaching a two-sided welfare-improving matching requires identifying sources of congestion that (1) generate a congestion cost, but (2) have a small or null individual return. Continuing with our dating example, these are the likes from men that generate a congestion cost, but that are not liked by the targeted woman, and hence have no chance of resulting in a match. Identifying these sources of congestion is challenging as it requires information on agents' preferences on both sides of the market; an information that is often unavailable.[27] In the next section, we therefore design a new ranking method, based on revealed-preference approaches, to estimate the preferences of agents on both sides of a market. We then compute this ranking for men and women in the dating market. This allows us to identify woman-man pairs that have low matching chances. We then show that, in these pairs, holding men back from liking the women would reduce congestion without harming women matching chances.

5 Ranking Agents in Two-Sided Markets

We design a ranking method that adapts the eigenvector centrality to a two-sided environment. This centrality measure has been widely used to compute rankings in markets that are one-sided, for instance to compute journals rankings (Palacios-Huerta and Volij, 2004) or to rank websites using Google's PageRank algorithm (Page et al., 1999).[28] More recently, Sorkin (2018) used workers' transitions between firms to build a ranking of firms. All these applications refer to one-sided markets, and require connections among websites, papers, or firms to compute the popularity ranking. A good web-page is linked to by many other good web-pages, a good paper is cited by a large number of other good papers, and a highly-ranked firm receives many applications from other high-rank firms (assuming that job search is done on-the-job and that transitions are between firms).

[27] Preferences in markets that are not fully centralized are typically not directly observable, and can be difficult to estimate in the absence of detailed micro data.

[28] Starting with journal citations, Narin et al. (1976) and Palacios-Huerta and Volij (2004) used citations between journals to derive a ranking of journal influence. The PageRank algorithm used by Google exploits links between webpages to compute a ranking of all pages (Page et al., 1999).

However, the one-sided eigenvector centrality measure is not well suited to two-sided markets in which interactions happen *between* two sides of a market rather than *within* one side. For instance, in a heterosexual dating market, a man cannot be ranked based on his interactions with other men. He could, however, get a high rank if many high-ranked women liked him.[29] This simple example provides the intuition of the two-sided eigenvector centrality method we develop. An agent on one side of the market is ranked high if he is linked to by numerous high-rank agents on the other side of the market.

5.1 Eigenvector centrality in two-sided markets

We present our two-sided eigenvector centrality method more formally in this section. We follow the standard approach by writing the popularity of a woman as a function of the men who like the woman, or more precisely, as a function of the share of likes the woman receives from every man:

$$w_i = \frac{y_{i,1} * m_1}{\sum_{\iota \in I} y_{\iota,1}} + \frac{y_{i,2} * m_2}{\sum_{\iota \in I} y_{\iota,2}} + ... \frac{y_{i,K} * m_K}{\sum_{\iota \in I} y_{\iota,K}} \quad \forall i \in I \tag{3}$$

The popularity w_i of a woman $i \in I$ is a function of the sum of the (unknown) popularities m_k of the men $k \in K$ who like this woman $y_{i,k} = 1$, each divided by the total number of women man k likes $\sum_{\iota \in I} y_{\iota,k}$. As this holds for every woman, we can write this as a linear system of equations:

$$\begin{pmatrix} w_i \\ \vdots \\ w_I \end{pmatrix} = \underbrace{\begin{pmatrix} y_{1,1} & \cdots & y_{1,K} \\ \vdots & \ddots & \vdots \\ y_{I,1} & \cdots & y_{I,K} \end{pmatrix}}_{Y} \underbrace{\begin{pmatrix} \frac{1}{\sum_{\iota \in I} y_{\iota,1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sum_{\iota \in I} y_{\iota,K}} \end{pmatrix}}_{Y_w} \begin{pmatrix} m_1 \\ \vdots \\ m_K \end{pmatrix} \tag{4}$$

Following the same idea, we write the popularity of a man as a function of the share of likes the man receives from every woman:

$$m_k = \frac{x_{k,1} * w_1}{\sum_{\kappa \in K} x_{\kappa,1}} + \frac{x_{k,2} * w_2}{\sum_{\kappa \in K} x_{\kappa,2}} + ... \frac{x_{k,I} * w_I}{\sum_{\kappa \in K} x_{\kappa,I}} \quad \forall k \in K \tag{5}$$

[29]The example applies to other markets. An Uber cab driver cannot be ranked based on his interactions with other cab drivers. He could, however, be ranked as a good cab driver if he was booked by many good customers. Similarly, a person renting a flat on Airbnb cannot be ranked based on her interactions with other flat renters. However, that person could get a high ranking if she receives many booking requests from top hosts.

The popularity m_k of a man $k \in K$ is a function of the sum of the (unknown) popularities w_i of the women $i \in I$ who like this man $x_{k,i} = 1$, each divided by the total number of men woman i likes $\sum_{\kappa \in I} x_{\kappa,i}$. As this holds for every man, we can write this as a linear system of equations:

$$\begin{pmatrix} m_k \\ \vdots \\ m_K \end{pmatrix} = \underbrace{\begin{pmatrix} x_{1,1} & \dots & x_{1,I} \\ \vdots & \ddots & \vdots \\ x_{K,1} & \dots & x_{K,I} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \frac{1}{\sum_{\kappa \in K} x_{\kappa,1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sum_{\kappa \in K} x_{\kappa,I}} \end{pmatrix}}_{X_w} \begin{pmatrix} w_1 \\ \vdots \\ w_I \end{pmatrix} \tag{6}$$

Observing that the popularities of men are a (linear) function of the popularities of women and vice versa means that we can write the popularities of women as a recursive function of their own popularity by plugging Equation 6 into Equation 4:

$$w = YY_w XX_w w$$

In other words, the popularity ranking of a woman (resp. man) is defined recursively and depends on both the number of men who like the woman, and their popularity rank. A woman who receives many likes from highly-ranked men receives a high popularity rank herself. Importantly, the value (or popularity rank) transferred from a given man to all the women he likes is divided equally among all women so that a top-ranked man who likes only one woman will contribute more to the popularity rank of that woman than if he had liked numerous women.[30]

YY_w and XX_w are both row-stochastic matrices, meaning that their rows sum up to one. Hence, the product of the two $YY_w XX_w$ is itself row stochastic. In addition, if we restrict the sample of men and women to the subset of strongly connected men and women, the matrix $YY_w XX_w$ is an irreducible non-negative matrix. A woman (a man) is in the strongly connected set if she (he) has been liked by at least one man (woman) in the strongly connected set and if she (he) likes at least one man (woman) in the strongly connected set. By the Perron-Frobenius Theorem for irreducible non-negative stochastic matrices, the eigenvector associated with $YY_w XX_w$ is unique up to scaling and corresponds to an eigenvalue of one (Minc, 1988, theorem 3.2 and theorem 4.1). The popularity of women w therefore corresponds

[30]By taking into account not only the number of likes between men and women, but also the popularity of men and women who like each other, our popularity ranking bears a striking resemblance with the desirability rating designed by the dating app Tinder to rate their users. Although the code of the ranking has never been released, Tinder's CEO, Jonathan Badeen, compared it to the ELO scoring used for chess players: "Whenever you play somebody with a really high score, you end up gaining more points than if you played someone with a lower score." (The Atlantic, 2016). This suggests that Tinder might have used a similar eigenvector-centrality based method to design their ranking.

to this eigenvector, and because it is unique up to scaling, the ranking of the women in w is unique. Following the same reasoning, the popularity of men m corresponds to the eigenvector associated with XX_wYY_w.

The eigenvector w is the fixed point of the function $YY_wXX_w : \mathbb{R}^{|I|} \to \mathbb{R}^{|I|}$. To reach this fixed point, we start with an initial guess and we repeatedly apply the function YY_wXX_w until the vector w has converged.

5.2 Estimating popularity of men and women

We use our two-sided eigenvector centrality method to compute men and women popularity ranking in the dating market. The market we consider changes every day due to entries and exits of agents. To account for this dynamic feature, we estimate the rankings on a rolling basis. For every day t in our period, we calculate the ranking for that day using interactions between men and women that happened in the 60 days preceding day t.[31] We therefore end up estimating 730 rankings, one for every day between January 1st 2014 and December 31st 2015. We standardize each ranking on a scale from 0 (lowest) to 100 (highest).

Leave-out rank estimations. A natural measure of the overall popularity of men and women would be their average popularity rank over all days in which they are ranked. Calculating popularity in this way, however, raises two concerns: First, we measure congestion for a woman as the number of *likes* she receives from the moment she creates her profile until she logs in to the platform. We then use the same *likes* to calculate her and other's popularity rankings for the following 60 days, so there is a mechanical relationship between the congestion outcomes we study and the popularity of men and women. To prevent this mechanical relationship, we employ a leave-out approach. When calculating popularity, we leave-out rankings of men or women which are based on likes given in the first week after either the woman or the man created her profile.[32]

The second concern is that there may be a mechanical relationship between men's popularity and women's popularity because the popularity ranking is estimated from *likes* between men and women. When a man receives an additional like, the women he liked will be ranked higher.

[31] Our results and conclusions are not sensitive to alternative choices of this time frame between 30 and 120 days.

[32] Recall that our sample contains women who log in within one week after creating their profile. Also, as the ranking on day t relies on interactions that happened in the 60 days preceding t, this means that we exclude any ranking on days that lie between the day when the woman created her profile and 67 days after. For example, if we consider a man who created his profile on May 1st 2014 and a woman who created her profile on October 1st 2014, then we exclude the rankings between May 2nd and July 7th 2014 and between October 1st and December 7th 2014 when calculating the average ranking of the man and the woman.

We have two reasons to believe that this concern is negligible. First, our leave-out approach omits likes happening after a woman and a man create their profile, which is the period where most likes are received (see Figure A.1). Second, because there are so many likes between men and women, a single like makes up for less than 1% of all likes used to calculate a woman's or a man's leave-out rank.[33]

Descriptive statistics on rankings. Table 5 Panel A shows descriptive statistics on the 730 rankings between January 1st 2014 and December 31st 2015. To compute all rankings, we use over half a million likes from women and over five million likes from men, which corresponds to 96.9% of all likes made by women and 86.5% of likes made by men. The share of likes is smaller for men due to their lower chances of being in a strongly connected set of agents. On a given day, the rankings for men and women are based on 47,779 likes from women and 380,838 likes from men, which determine the popularity rank of 1,401 women and 2,997 men (see Table 5 Panel B). The average leave-out ranking is based on 221 ranking days for women and 249 ranking days for men. We can estimate a leave-out ranking for 4,824 women and 9,381 men (see Table 5 Panel C).[34]

To validate our popularity ranking, we also compute a simpler popularity ranking for men and women that treats all likes received with the same weight. The most popular woman is the woman with the highest share of likes. Specifically, for each woman, we consider all the men who have seen the woman's profile, and we calculate the percentage of these men who liked the profile. We proceed identically to compute the men's popularity ranking. As we will show in the following, the popularity ranking that gives the same weight leads to very similar results as our eigenvector centrality based popularity ranking.

[33]The average number of likes received by men and used to calculate the popularity ranking of a particular day is 34. The average number of ranking days used to calculate the leave-out ranking is 249 (Table 5). A single like can be used to calculate the ranking of at most 60 ranking days. Thus a single like can make up for no more than $1 * 60/(249 * 34) = .7\%$ of likes men receive. Following the same line of reasoning, a single like can make up for no more than $1 * 60/(221 * 127) = .2\%$ of likes women receive.

[34]The number of men and women differs from the numbers reported in Table 1 because Table 5 uses all platform *users* between November 1st 2013 (for the ranking valid on January 1st 2014) and December 30th 2015 (for the ranking valid on December 31st 2015), whereas Table 1 describes the sample of men and women creating a profile between January 1st 2014 and December 31st 2015. The women and men considered in Table 5 are very similar to those described in Table 1 with regard to characteristics and statistics on behavior on the platform.

Table 5: Descriptive statistics on men and women's popularity ranking

	Women	**Men**
Panel A: Statistics on all rankings		
Nb of likes used to compute the rankings	592,205	5,163,455
Share of all likes used to compute the ranking	96.9%	86.5%
Panel B: Statistics on ranking per day		
Nb of likes used to compute the ranking	47,779	380,838
Nb of users ranked	1,401	2,997
Panel C: Statistics on leave-out rankings		
Nb of ranking days	221	249
Nb of users ranked	4,824	9,381

Notes: Notes: This table shows descriptive statistics on the popularity rankings for every day from January 1st 2014 to December 31st 2015. Panel A shows descriptive statistics on all rankings valid between January 1st 2014 and December 31st 2015. The *Share of all likes used to compute the ranking* refers to the share of all likes made on the dating app between November 1st 2013 and December 30th 2015 that are used at least once to estimate the popularity ranking. Panel B shows the following statistics on the popularity rankings: On how many likes a ranking is based on average; How many men and women are ranked on average. *Nb of ranking days* shows the average number of ranking days between January 1st 2014 and December 31st 2015 over which the leave-out average popularity rank is calculated. *Nb of users ranked* shows the number of users for which we can calculate the leave-out popularity ranking for at least one profile they check.

6 Towards a Two-Sided Efficient Matching: Role of a Mutual Popularity Ranking

Our motivation for estimating men and women popularity rank is to identify likes from men that might generate a social congestion cost, but these likes have no individual return as they are not liked back by the targeted woman, and hence have no chance of resulting in a match. The relative popularity ranking helps us to identify man-woman pairs that might have low matching chances because the woman is significantly more popular than the man (or vice versa). In this section, we successively test if the relative popularity ranking helps to identify men and women's likes that have (1) a small or null individual return, but (2) a positive congestion cost.

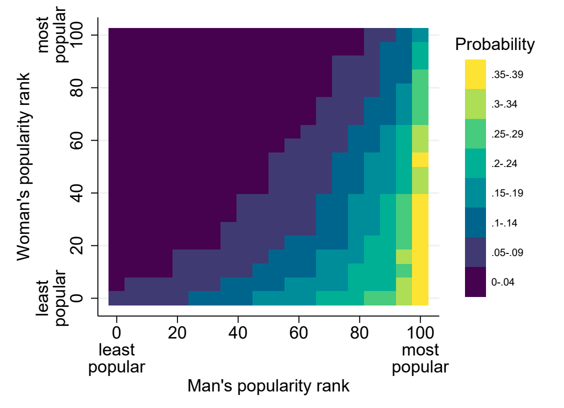
6.1 Mutual popularity ranking and matching probability

Result 1: Correlated preference for popular partners. We start by investigating the relationship between popularity ranking and matching probability. Figure 4b shows the probability that a woman likes a man as a function of both her own rank (y-axis) and the man's rank (x-axis). Figure 4a reports the same probabilities, but for men. These heat maps are useful to visually check if men and women's likes are directed or not. Under directed search, we expect men (resp. women) to take into account the matching chances when they like a woman.[35] The probability that a man likes a woman would therefore increase when a woman's popularity is close to the man's popularity (on the 45 degree line of each figure), and it would move down when women become more popular than the man (when moving to the right of the 45 degree line).

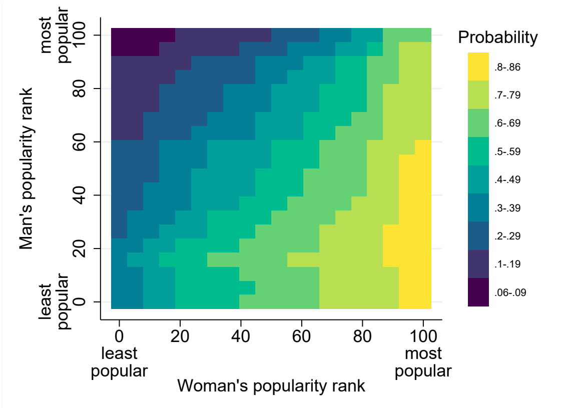
The heat maps do not support directed search. Panel b shows that all men, i.e., irrespective of their own popularity rank, tend to prefer popular women, as shown by the increasing like probability when moving horizontally from the left to the right in Figure 4b. This finding is more true for men than for women, meaning that the probability that a man likes a woman increases significantly more with her popularity than when considering women's likes as a function of men's popularity. Women have a close-to-zero chances of liking men who are less popular than they are, as shown by the dark blue area on the left of the 45 degree line in Figure 4a. There are also large differences between men and women in their overall liking probabilities, which range from 0 to 39% for women versus 6 to 86% for men.[36,37]

Figure 4: Like probability as a function of men and women popularity

(a) Women liking men



(b) Men liking women



Notes: Figure (a) shows the probability a woman likes a man that appears on her feed depending on the woman's and the man's popularity rank. Figure (b) shows the probability a man likes a woman that appears on his feed as a function of the man's and the woman's popularity rank. We take the popularity rank of the woman (resp. the man) on the day when she sees the man's (resp. woman's) profile. Both figures show shares within cells of a 20x20 grid. Figure (a) is based on 5,398,981 men profiles checked by 4,824 women between January 1st 2014 and December 31st 2015. Figure (b) is based on 8,561,333 women profiles checked by 9,381 men between January 1st 2014 and December 31st 2015.

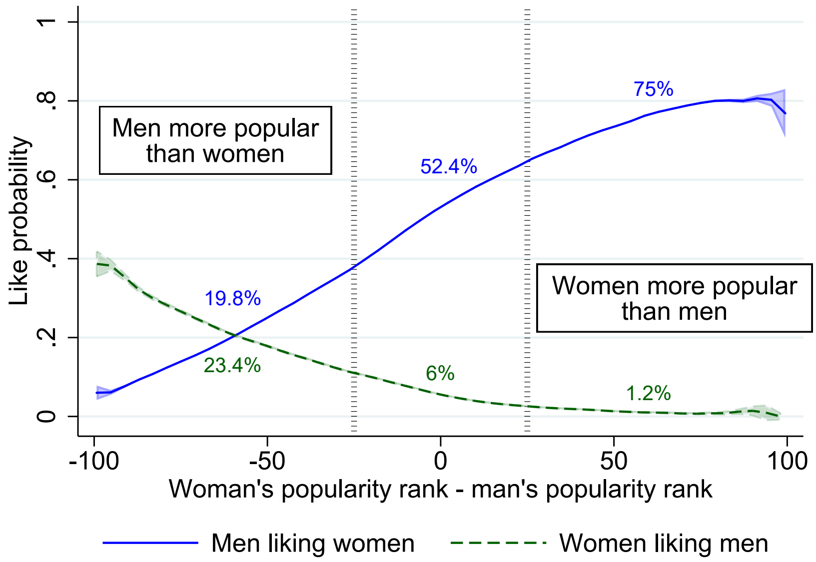
[35]This is sometimes also referred to as competitive search (see Wright et al. (2021) for a detailed review of directed and competitive search theory).

[36]Figure A.5 shows that we obtain similar results when we use a popularity ranking that gives the same weight to all likes.

[37]Following this result we derive an estimate of the extent to which preferences of women and men are vertical, meaning that they are explained by the popularity ranking (Appendix A). We estimate that 31 percent of women preferences and 32 percent of men preferences are vertical.

The fact that men and women likes are determined by their partners' absolute popularity more than by their relative popularity suggests that users of the dating app either have limited information on the relative popularity of their partners, or that they are not discouraged by it. Said differently, men and women show limited signs of directed search in our market. They do not factor in their matching chances when they like others. This explains the formation of congestion for the most popular women on the dating app.

Figure 5: Like probability as a function of men and women relative popularity



Notes: This figure shows the probability a woman (man) likes a man (woman) that appears on her (his) feed, depending on the difference between the popularity rank of the woman and the man. We take the popularity rank of the woman (resp. the man) on the day when she sees the man's (resp. woman's) profile. The lines represent local polynomial estimates, and the surrounding shaded areas 95% confidence intervals. The figure is based on 5,398,981 men profiles checked by 4,824 women, and 8,561,333 women profiles checked by 9,381 men between January 1st 2014 and December 31st 2015.

Result 2: No personal benefit of overshooting. The fact that most men prefer popular women leads many unpopular men to like popular women. Figure 5 shows the probability that a man likes a woman (y-axis) as a function of his relative popularity, that is, as a function of the difference between the popularity rank of the woman and the popularity rank of the man (x-axis). The left side of the graph corresponds to pairs in which the man is more popular than the woman, while the right side corresponds to the opposite.

We see that, when a woman is much more popular than a man (at least 25 ranks more popular), men like the woman 75% of the time (blue line). Naturally, the fact that men like women who are significantly more popular than themselves is not necessarily irrational, especially if they have a fair chance of being liked back despite the popularity difference. Yet, this is not what the green line suggests. When a man is much less popular than a woman (at least 25 ranks less popular), women only like the man 1.2% of the time. Importantly, this low probability of being liked back only applies when the woman is significantly more popular than the man. When the man and woman are equally popular, women like 6% of the men, and when a man is more popular than a woman (at least 25 ranks more popular), women like the man 23.4% of the time.[38]

The low matching chances for men who "over-shoot" suggests that some men's likes have a close-to-zero individual return. Before concluding that the market would be better off had these unpopular men not liked the popular women, we need to understand if these likes generate a negative externality for other men in the market. Hence, the question we address in the next section is: Do men who like a woman, despite being significantly less popular than she is, generate congestion costs for other men in the market?

6.2 Mutual popularity ranking and congestion cost

We saw in Section 4 that the demand for women suffers from congestion, and that this congestion is costly for men. With the new information we have on agent popularity, we can now test whether the congestion cost is caused by men who have low or no chances of being matched with the woman they like. Specifically, we want to check if the likes from men who are less popular than a woman induce a congestion cost for men with a higher chance of getting a match, i.e., men who are more popular than the woman.

Method. To shed light on this question, we use a similar instrumental variable methodology as in Section 4. We consider each man i who likes a woman j who is less popular than he is—this is the man with a reasonable chance of being liked back by the woman—and we estimate how much increasing the number of likes from other men who are significantly less popular than woman j affects: (i) the probability that the woman j sees the man i on her feed, (ii) the probability that the woman j matches with the man i, and (iii) the time it takes for them to reach a match. Our main equation is:

$$Y_{ij}|r_i \geq r_j = \alpha + \beta L_{j,r_{-i} \geq r_j} + \gamma L_{j,r_{-i} < r_j +} + \lambda X_j + \epsilon_{ij} \tag{7}$$

where Y_{ij} is the outcome of interest, $r_i \geq r_j$ is a condition indicating that a man i has

[38]Figure A.6 shows that we obtain very similar results when we use a popularity ranking that gives the same weight to all likes.

popularity rank r_i that is higher than woman j's popularity rank r_j. On the right hand side, $L_{j,r_{-i}<r_j}$ captures the number of likes that woman j receives from men (excluding i) who are less popular than she is, and $L_{j,r_{-i}\geq r_j}$ the number of likes that woman j receives from men (excluding i) who are more popular than she is.[39] X_j is a vector of control variables for woman j that includes her popularity rank r_j, her age, whether she defined an age filter, whether her profile contains text, and her number of Facebook friends. Controlling for a woman's popularity is important to account for unobserved determinants of a woman's preference, behavior, and activity on the dating app. λ thus identifies the effect of increasing the number of likes that woman j receives from less-popular men on the probability that man i will be seen by the woman on her feed, holding constant the number of likes from men ranked higher than the woman as well as the woman's popularity rank r_j.

As previously, the number of men who are less popular than woman j and like her ($L_{j,r_{-i}<r_j}$) is likely to be endogenous, as is the number of men who are more popular than woman j and like her ($L_{j,r_{-i}\geq r_j}$). To address this endogeneity, we exploit quasi-random variation in the number of men who are less popular than woman j (resp. more popular than her) and who use the dating app during the 24h after she creates her profile.

When designing the instrument, we need to account for the mechanical relationship that exists between a woman's popularity and the number of less- and more-popular men who use the app. The higher a woman's popularity, the lower the number of men on the app who are more popular than her, a correlation that violates the independence assumption of the instrument.[40] To get rid of this mechanical correlation, we control for a woman's popularity in the first stage regression and we construct our instrumental variable as the interaction between the woman's popularity rank and the number of more- and less-popular men using the dating app during the 24h after a woman creates her profile. Ultimately, the variation we exploit therefore comes from day-to-day variation in the number of more-popular and less-popular men using the platform *across women who have the same popularity*.[41]

[39]The popularity ranking of woman j enters both the left- and right-hand-side of Equation 7. To avoid the resulting mechanical relationship between dependent and independent variables, as explained in section 5.2, for each man-woman pair in our sample, we calculate the man and woman's popularity rank as the average rank over all days when they are ranked, excluding the week following the profile creation of the woman.

[40]The independence assumption requires that the instrument is unrelated to a woman's unobservable characteristics that determine the outcomes (e.g., her chances of seeing or liking a more-popular man on her feed).

[41]More formally, the first stage equation for this two-stage least squares (2SLS) procedure is: $L_j = \theta + \lambda r_j + \eta M_{j,r_{-i}\geq r_j} * r_j + \nu M_{j,r_{-i}<r_j+} * r_j + \rho X_j + \epsilon_j$ where r_j is woman j's popularity rank, $M_{j,r_{-i}\geq r_j} * r_j$ (resp. $M_{j,r_{-i}<r_j+} * r_j$) corresponds to the interaction between the woman's popularity rank and the number of men who are more popular (resp. less popular) than woman j and who use the app in the 24h after her profile creation. λ captures how much a woman's rank determines the number of likes she receives. η and ν capture the heterogeneous effect of the rank depending on the quasi-random variation in the number of more- and less-popular men on the app. Table A.6 shows that these first stage coefficients are statistically significant. The first-stage F-statistics are also larger than 10.

Table 6: Congestion costs generated by men who are less- and more-popular than a woman

	P(being seen)		P(being liked)		Days to match	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
Panel A: Controlling for woman's rank						
Nb likes from	-0.0757***	-0.2669***	-0.0459***	-0.1003***	0.0158***	0.0179**
more-popular men	(0.0210)	(0.0485)	(0.0059)	(0.0164)	(0.0031)	(0.0076)
Nb likes from	-0.0154	-0.0850**	0.0034	0.0244*	0.0117***	0.0181***
less-popular men	(0.0153)	(0.0380)	(0.0054)	(0.0146)	(0.0020)	(0.0067)
Panel B: Controlling for woman's rank, age and number of Facebook friends						
Nb likes from	-0.0834***	-0.2603***	-0.0466***	-0.0961***	0.0158***	0.0185**
more-popular men	(0.0208)	(0.0464)	(0.0058)	(0.0161)	(0.0031)	(0.0077)
Nb likes from	-0.0078	-0.0653*	0.0044	0.0274*	0.0119***	0.0191***
less-popular men	(0.0151)	(0.0373)	(0.0054)	(0.0148)	(0.0020)	(0.0070)
Panel C: Controls from Panel B, and whether woman defined an age filter, and a profile text						
Nb likes from	-0.0842***	-0.2696***	-0.0468***	-0.0985***	0.0158***	0.0189**
more-popular men	(0.0208)	(0.0471)	(0.0058)	(0.0162)	(0.0031)	(0.0077)
Nb likes from	-0.0083	-0.0655*	0.0044	0.0269*	0.0119***	0.0193***
less-popular men	(0.0145)	(0.0371)	(0.0054)	(0.0147)	(0.0020)	(0.0071)
Panel D: Controls from Panel C, and controlling for same characteristics of the man liking the woman						
Nb likes from	-0.0802***	-0.2609***	-0.0460***	-0.0942***	0.0158***	0.0196**
more-popular men	(0.0208)	(0.0464)	(0.0058)	(0.0160)	(0.0031)	(0.0078)
Nb likes from	-0.0099	-0.0583	0.0041	0.0286*	0.0120***	0.0191***
less-popular men	(0.0145)	(0.0364)	(0.0054)	(0.0147)	(0.0020)	(0.0069)
Observations	87,739	87,739	87,739	87,739	7,140	7,140
Mean. Dep.var	68.09	68.09	8.83	8.83	1.77	1.77
# of women	2,808	2,808	2,808	2,808	2,029	2,029
First stage F		33.54		33.54		17.85

Notes: This table reports the β and γ from estimating Equation 7. Columns (1), (3) and (5) report coefficients from an OLS regression. In columns (2), (4) and (6), we instrument the number of likes from more-popular men (resp. less-popular men) with the number of men using the platform during the 24 hours after a woman profile creation that are more-popular (resp. less-popular) than the woman. The unit of observation is the like of a man for a woman. The regressions are estimated on the subsample of men who are, on average over the sample period (excluding the period of interaction between the men and women, as described in the text), more popular than the woman they like. The dependent variable Y_{ij} is, in columns (1) and (2), an indicator for whether a man who liked the woman ever appears on the woman's profile feed, in columns (3) and (4) an indicator for whether a man who liked a woman is liked back by the woman, and in columns (5) and (6) the number of days elapsed between when a woman receives a like and when she likes the man back. We exclude matches that happen later than 30 days after the woman created her profile. Control variables in Panel D include all control variables from Panel C, plus a control for the age of the man who likes the woman, his number of Facebook friends, whether the man has defined an age filter, and whether the man's profile contains text. First stage results are reported in Table 2. First stage F stats correspond to the lowest Kleibergen-Paap F-statistic of the four specifications in Panel A-D. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Results. Table 6 reports our estimates of the cost of congestion inflicted by men who are less popular than a woman—referred to as *less-popular men*—on men who are more popular than that woman—referred to as *more-popular men*. The coefficient on *Nb likes from less-*

popular men confirms that less-popular men impose a congestion cost for their more-popular peers. Raising the number of likes a woman receives from less-popular men by 100 reduces the chances that she will see a more-popular man on her feed by 8.6 percentage points. It also increases the number of days it takes to reach a match by 1.8 (+102.3%).

It is also interesting to notice that the congestion costs generated by more-popular men are about three times as large as the congestion costs generated by less-popular men. This difference is driven by the fact that the platform's recommendation algorithm prioritizes men that have been liked by a large share of other women. As these men also tend to be more popular according to our ranking, it is not surprising that likes from *more-popular men* impose greater congestion costs. The current prioritization of popular men on the dating app also implies that our estimates of the congestion costs generated by *less-popular men* most likely under estimate the congestion costs we would find on any platform that does not prioritize popular men.

Summing up, our results confirm that men who are less popular than a woman have very small matching chances, but they generate negative externalities on other men when liking the woman. Reducing likes from these men would most likely increase the matching chances of other men without affecting their own matching chances.

Robustness checks. We run several robustness checks to verify how stable our estimates are across specifications. First, we progressively add controls for women age and their number of Facebook friends (Panel B), as well as whether a woman defined an age filter, and whether her profile contains text (Panel C). The estimates are mostly unaffected, which is reassuring as Table A.7 suggests that some women characteristics are correlated with our instrumental variables (the interaction between the woman popularity rank and the number of more-popular and less-popular men using the app).[42] Although our results show that this correlation does not affect our results, it raises the risk that there are other unobserved characteristics of the women that might be correlated with our instrument, which would violate the independence assumption of the instruments. Our results on the congestion cost generated by less-popular men should therefore be interpreted with more caution. Finally, we also control for men characteristics (Panel D), and the results remain unchanged.

[42]Conditional on a woman's rank, women who connect for the first time on a day with a larger number of less-popular men tend to have fewer Facebook friends, to be slightly older, and to be more likely to have set up an age filter or to present themselves in their profile.

7 Conclusion

We analyze congestion in two-sided markets using data from an online dating platform. After documenting that some women receive very large number of likes from men—a phenomenon we term congestion—we estimate the causal costs and benefits of congestion using an instrumental variable research design. We exploit variation in the number of likes women receive that stem from exogenous variation in the number of men using the platform. Our results reveal substantial congestion costs for men, as measured by men's lower probability of appearing on a woman's feed, and increased matching time. On the other hand, women who are fast at screening profiles, that is, women with low screening costs, benefit from congestion. The share of their feed that is composed of men who like them goes up. Finding that congestion is costly for men, but beneficial for women, raises a challenge when designing policies that aim at reducing congestion. Women might see their matching chances drop when they receive fewer likes, unless these likes are from men they do not like. Said differently, a congestion-reducing policy that improves the matching outcomes of both men and women should account for the preferences of both men and women.

We develop a new ranking method, based on a revealed-preference approach, to estimate the preferences of agents on both sides of a market. Our method adapts eigenvector centrality measures to two-sided markets. We then use our method to compute men and women popularity ranking in the dating market. Equipped with this additional information, we show that congestion arises because men tend to have correlated preferences and to like popular women, irrespective of their own popularity. Yet, our results reveal that many of the likes from low-popularity men to high-popularity women (i) have very little chances of resulting in a match, but (ii) they create a costly congestion for other higher-popularity men. We conclude that information on agent relative popularity is useful to design targeted policies that reduce congestion, but without negatively affecting matching outcomes on any side of the market.

Before discussing the broader policy relevance of our results, we should stress the environments in which our approach is most relevant. The ranking method we design produces a unique popularity ranking of men and women. Using this ranking to proxy preferences is therefore particularly appropriate in environments in which men and women's preferences are partly aligned. This raises the question of the relevant group of agents on whom to compute the ranking as the degree of correlation in agent preference largely depends on it. Considering all men and women in a country might produce large heterogeneity in preferences due to geographical preferences over a partner. In contrast, considering preferences of men and women who are in the same city would produce preferences that are much more aligned. Our popularity ranking is most suitable for markets with vertical preferences, like university admission markets, or in smaller size markets in which preferences are more aligned, like local labor markets or local dating markets.

Keeping this in mind, the mutual popularity ranking approach we propose has a number of applications in two-sided markets, notably in the public sector. Unemployment agencies from many countries use matching technologies to match unemployed workers to vacancies or training programs (see, for example, Belot et al. (2019) for the UK, Hensvik et al. (2020) for Sweden, and Dhia et al. (2022) for France). University admission systems also use matching tools to help match students to universities. In an attempt to better direct searches, online job boards and university admission platforms have even started to provide participants with information on the popularity of jobs and universities (Gee, 2019; Bhole et al., 2021; Parcoursup, 2021).[43] However, job seekers and university applicants are likely to react differently to information on the number of applicants than to information on the number of applicants combined with information on how much a recruiter or university values and prioritizes their profile. The mutual popularity ranking we design opens the door to interesting applications and interventions in several decentralized two-sided markets.

[43] Gee (2019) shows that providing job seekers information on the number of applicants increased the likelihood of application by 1.9% to 3.6%. Similarly, the French university match system shows students information on the share of admitted students in previous years (Parcoursup, 2021).

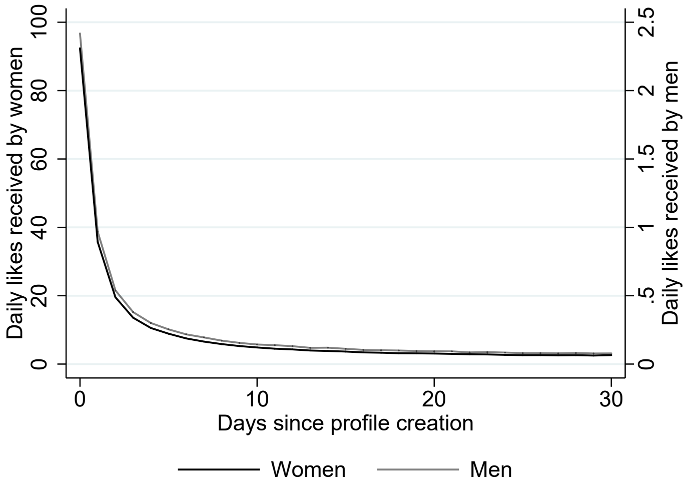
ONLINE APPENDIX

Table A.1: Balance test of sample restrictions

	All		Sample	
	Women (1)	Men (2)	Women (3)	Men (4)
A. User Characteristics				
Age	27.3	28.3	27.2	28.6
Profile contains text (%)	0.090	0.128	0.091	0.131
User has set age filter	0.164	0.176	0.162	0.176
Number of Facebook friends	504.4	598.6	499.7	592.0
B. User Activity				
Number of hours btw profile creation and first login	9.6	16.3	10.7	16.7
Number of logins per day	1.6	1.7	1.6	1.8
Number of profiles seen per day	44.2	53.8	43.9	54.3
Number of minutes active per login	5.2	4.0	5.1	4.0
Seconds spent per profile	6.6	4.7	6.6	4.7
C. Preferences				
Share of profiles liked	0.111	0.526	0.111	0.530
Share of profiles matched	0.046	0.015	0.046	0.015
Share of profiles with chat	0.005	0.002	0.005	0.002
D. User Popularity				
Share of likes received in first week	0.506	0.443	0.515	0.453
Number of likes received (at first login)	61.8	2.8	61.5	3.0
Share of users with 0 likes (at first login)	0.096	0.336	0.027	0.246
E. Metrics on Congestion Costs and Benefits				
Prob. of seeing partner who liked own profile	0.556	0.887	0.558	0.893
Prob. of liking a partner who liked own profile	0.062	0.390	0.062	0.391
Mean days to match	14.1	11.0	11.2	6.2
Share of matches happening within first month	0.910	0.935	0.933	0.968
Mean days to match | match within first month	1.683	1.331	1.434	0.973
Number of individuals	5,343	11,010	4,238	8,788
Number of likes received at first login	330,408	30,533	260,434	26,129

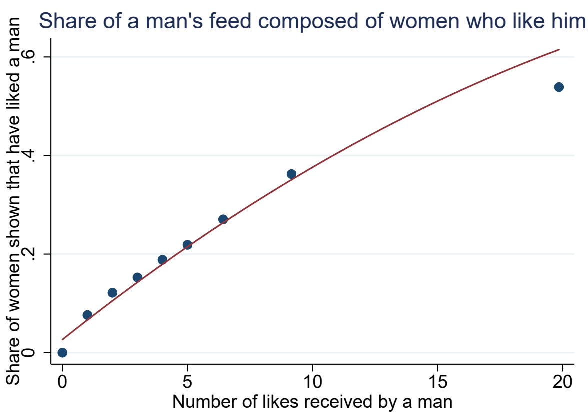
Notes: Columns (1) and (2): All profiles created between January 1st 2014 and December 31st 2015. Columns (3) and (4): Profiles created between January 1st 2014 and December 31st 2015 and first login $\leq$ 1 week after profile creation and $\geq$ 1 hour after profile creation, i.e., the sample shown in Table 1. The variable *Days to match* represents the days elapsed between when a user receives a like and when he likes back. *Share of matches happening within first month* shows, among all likes the user receives before his first login and that ultimately result in a match, the share for which this match happens within 30 days after the user created his profile. For *Number of hours btw profile creation and first login* the median is reported. All other variables in Panel B are calculated based on activity in the first 2 weeks after the first login. *Share of likes received in first week* shows, out of all likes received within 10 weeks after profile creation, the share that is received within the first week after profile creation.

Figure A.1: Timing of likes received



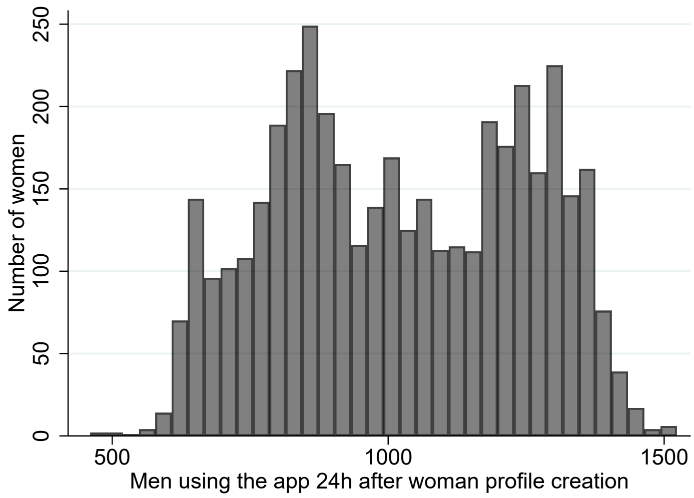
Notes: This figure shows the average number of likes women and men receive on each day in the first month after creating their profile. The samples include all men and women creating their profile between January 1st 2014 and December 31st 2015 (Table A.1, columns (1) and (2)).

Figure A.2: Congestion benefits for men



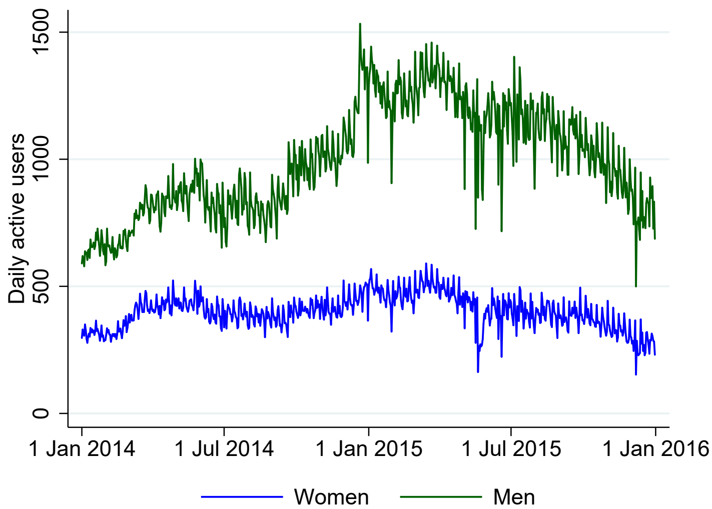
Notes: This figure reports the share of a man's feed at his first login that is composed of women who have liked him (y-axis), as a function of the number of likes the man receives before his first login. The figure is a binned scatter plot, with men assigned to 20 equal sized bins by the number of likes they receive before their first login. Bins with the same value of the x-axis variable merged. The red line represents the regression line from a linear quadratic regression of the share of women shown that have liked the man on the number of likes the man receives before his first login. The figure is based on all men in the sample (column (3) of Table 1).

Figure A.3: Distribution of number of men using the app 24h after woman profile creation



Notes: This figure shows the number of men using the platform in the 24h after women profile creation for the 4,122 women in column (2) of Table 1 who receive at least one like before their first login. The mean of the distribution shown in this figure is $1,024$, and the standard deviation 226.

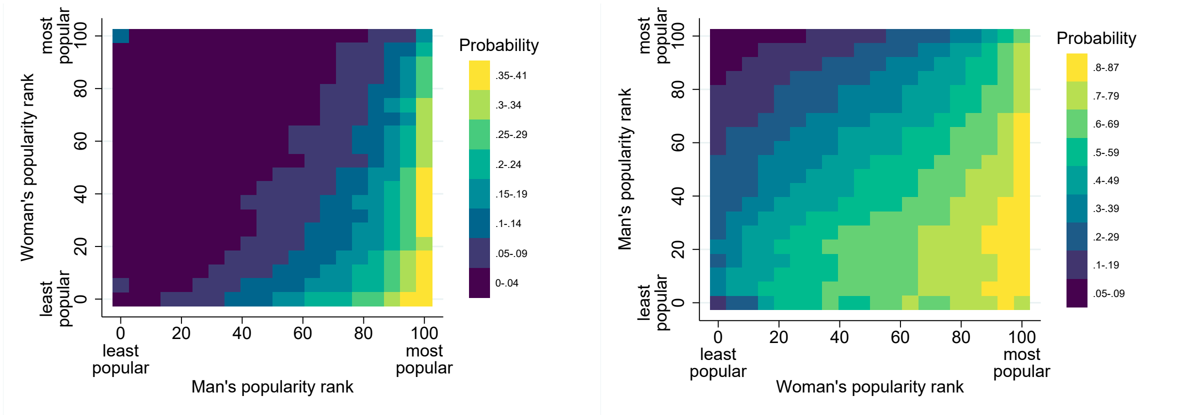
Figure A.4: Daily number of men and women using the dating app



Notes: This Figure shows the daily number of distinct men and women using the platform between January 1st 2014 and December 31st 2016.

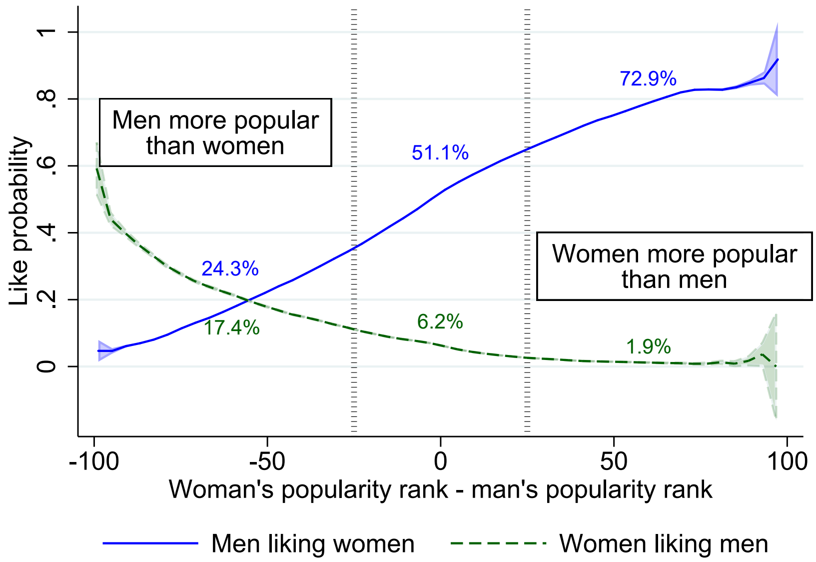
Figure A.5: Like probability as a function of men and women popularity - Popularity measured as the share of men (resp. women) who like a woman (resp. man)

(a) Women liking men (b) Men liking women



Notes: Figure (a) shows the probability a woman likes a man that appears on her feed and has not yet liked the woman, depending on the woman's and the man's popularity. Figure (b) shows the probability a man likes a woman that appears on his feed and has not yet liked the man as a function of the man's and the woman's popularity. The popularity is measured using the share of ratings received within the past 60 days that are likes, where 100 is the woman (man) who has received the greatest share of likes and 0 is the woman (man) who has received the smallest share of likes. We take the popularity rank of the woman (resp. the man) on the day when she sees the man's (resp. woman's) profile. Both figures show shares within cells of a 20x20 grid. Figure (a) is based on 5,398,981 men profiles checked by 4,824 women between January 1st 2014 and December 31st 2015. Figure (b) is based on 8,561,333 women profiles checked by 9,381 men between January 1st 2014 and December 31st 2015.

Figure A.6: Like probability as a function of men and women relative popularity - Popularity measured as the share of men (resp. women) who like a woman (resp. man)



Notes: This figure shows the probability a woman (man) likes a man (woman) that appears on her (his) feed and has not yet liked the woman, depending on the difference between the popularity rank of the woman and the man. The popularity is measured using the share of ratings received within the past 60 days that are likes, where 100 is the woman (man) who has received the greatest share of likes and 0 is the woman (man) who has received the smallest share of likes. We take the popularity rank of the woman (resp. the man) on the day when she sees the man's (resp. woman's) profile. The lines represent local polynomial estimates, and the surrounding shaded areas 95% confidence intervals. The figure is based on 5,398,981 men profiles checked by 4,824 women, and 8,561,333 women profiles checked by 9,381 men between January 1st 2014 and December 31st 2015.

Table A.2: Correlation between women characteristics and congestion outcomes

	P(being seen)	P(being liked)	Days to match
	(1)	(2)	(3)
Number of Facebook friends	-0.0038***	0.0001	0.0003
	(0.0015)	(0.0004)	(0.0003)
Age	0.0251	0.0356**	-0.0179
	(0.0850)	(0.0180)	(0.0142)
Has age filter	12.8554***	0.4888	0.4252
	(1.4609)	(0.3007)	(0.2836)
Has profile text	5.9703***	0.7948*	0.2612
	(1.7546)	(0.4450)	(0.3308)
Observations	260,434	260,434	10,990
Mean. Dep.var	51.00	4.58	1.78

Notes: This table shows the coefficients from bivariate regressions of the respective outcome (column) of a like given by a man to a woman on the characteristics of the woman receiving the like (row). Each coefficient corresponds to a separate regression. The corresponding balance tests are shown in Table 2, columns (3) to (7). Days to match represents the days elapsed between when a woman receives a like and when she likes the man back, where we exclude matches that happen later than 30 days after the woman created her profile. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.3: Test of instrument independence – women's behavior

	Number of hours between profile creation and first login	Number of profiles checked at first login
	(1)	(2)
Nb men using app 24h after woman profile creation	0.0026	-0.0115
	(0.0018)	(0.0097)
Observations	4,122	4,122
F-statistic	2.05	1.41
R^2	0.00	0.00

Notes: This table shows coefficients from bivariate regressions of measures of women's behavior on the platform on the instrumental variable, i.e., the number of men using the app in the 24h after a woman creates a profile (M_j). *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.4: Correlation between nb of men using the app and men characteristics

	Friends on Facebook	Age	Has age filter set	Profile text
	(1)	(2)	(3)	(4)
Number of men using the app 24h after woman profile creation	-0.0770** (0.0350)	0.0030*** (0.0003)	0.0003*** (0.0000)	0.0001*** (0.0000)
Observations	514,154	514,154	514,154	514,154
Distinct men	7,030	7,030	7,030	7,030
F-statistic	4.83	85.06	68.70	42.01
R^2	0.00	0.01	0.01	0.01

Notes: This table shows the coefficients from bivariate regressions of the characteristics of men liking a woman (between when the woman creates her profile and when she logs in for the first time) on the number of men using the app 24h after the woman profile creation. Each coefficient corresponds to a separate regression. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.5: Congestion benefits for men

	All men		Men with high screening costs		Men with low screening costs	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Nb of likes received at 1st login from women	2.5120*** (0.0498)	5.6255*** (0.8672)	2.4842*** (0.0594)	3.9224** (1.8910)	1.7063*** (0.0604)	2.9875 (2.0173)
Observations	8,788	8,788	4,184	4,184	4,184	4,184
Mean. Dep.var	12.96	12.96	16.70	16.70	5.53	5.53
First stage F		42.00		4.71		4.16

Notes: This table reports the β coefficients from the following regression: $Y_j = \alpha + \beta L_j + \epsilon_j$, where Y_j is the share of a man's feed that is composed of women who have liked him (at the time when he logs in to the platform the first time). L_j denotes the number of likes the man receives between his profile creation and his first connection to the platform. Columns (1), (3) and (5) report coefficients from an OLS regression. In columns (2), (4) and (6), we instrument the number of likes L_j by the number of women using the platform in the first 24h after a man creates his profile. The regressions in columns (1) and (2) are estimated on the full sample of men (column (3) of Table 1). Columns (3) and (4) show regression results for the subsample of men who take more than the median time to evaluate a profile (> 5 seconds). Columns (5) and (6) show regression results for the subsample of men who take less than the median time to evaluate a profile (≤ 5 seconds). The unit of observation is a man. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.6: First stage – Congestion costs from less-popular and more-popular men

	Nb likes from more-popular men	Nb likes from less-popular men	Nb likes from more-popular men	Nb likes from less-popular men
	(1)	(2)	(3)	(4)
Woman's popularity rank	-0.9552***	0.2350	-0.5374***	0.4595**
	(0.1130)	(0.1607)	(0.0903)	(0.1815)
Nb more-popular men using app	0.0018***	0.0005**	0.0013***	0.0001
× Woman's rank	(0.0002)	(0.0002)	(0.0001)	(0.0002)
Nb less-popular men using app	0.0004***	0.0017***	0.0001	0.0013***
× Woman's rank	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Observations	87,739	87,739	7,140	7,140
F-statistic	73.17	326.48	67.53	199.73
# of women	2,808	2,808	2,029	2,029
Second stage's depvar	P(being seen), P(being liked)		Days to match	

Notes: This table shows the first stage results corresponding to the second stage results in Table 6. The dependent variable is the number of likes a woman j receives before her first login from more-popular men (columns (1) and (3)) or from less-popular men (columns (2) and (4)). The right-hand-side variables are the excluded instruments, that is, the number of more-popular and less-popular men using the app in the 24h after woman j creates a profile, interacted with woman j's popularity rank. Regressions control for woman j popularity rank. We obtain almost identical coefficients when we control for woman j's and man i's age, whether she/he defined an age filter, whether her/his profile contains text, and her/his number of Facebook friends, or a subset of these covariates. We cluster standard errors at the woman level. Standard errors are clustered by woman. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Table A.7: Test of independence assumption – Congestion costs from less-popular and more-popular men

	# of friends on Facebook	Age	Has age filter set	Has profile text
	(1)	(2)	(3)	(4)
Nb more-popular men using app	-0.1966**	0.0001	0.0002**	0.0004***
× Woman's rank /100	(0.0811)	(0.0016)	(0.0001)	(0.0001)
Nb less-popular men using app	-0.1303**	0.0059***	0.0001**	0.0002***
× Woman's rank /100	(0.0618)	(0.0012)	(0.0001)	(0.0001)
Observations	2,850	2,850	2,850	2,850
F-statistic	64.47	101.33	6.01	15.75
R^2	0.06	0.10	0.01	0.02

Notes: This table shows coefficients from regressions of woman j's characteristics on the number of higher and lower ranked men using the app in the 24h after a woman j creates a profile, interacted with the woman j's popularity rank. The sample of women is the same as the one we use in Table 6. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

A Vertical vs. Horizontal Preferences in the Dating Market

The ranking method we design produces a unique popularity ranking of men and women. Using this ranking to proxy preferences is therefore appropriate in environments in which men and women's preferences are partly aligned. Said differently, some of men and women's preferences should be vertical for the popularity ranking to bring meaningful information on preferences. In this section, we quantify the extent to which men and women's preferences in our dating market are vertical versus horizontal. We use a simple model, inspired by Menzel (2015), in which a woman i derives the following utility from matching with a man j:[44]

$$V_{ij} = \beta P_j + \epsilon_{ij} \tag{1}$$

We assume that the common value of man j is a linear function of his popularity rank P_j. A man's popularity rank captures the vertical component of women's preferences. ϵ_{ij} represents the idiosyncratic preference of woman i for man j. This term captures the horizontal component of women's preferences. If woman i remains single, her utility would be:

$$V_{ii} = \gamma P_i + \epsilon_{ii} \tag{2}$$

We assume that the common outside option value of woman i is a linear function of her popularity rank P_i. This is justified by the empirical evidence in Section 6 showing that the probability a woman likes a man is a decreasing function of her own popularity rank.[45] ϵ_{ii} represents the idiosyncratic preference of woman i for remaining single.

When man j appears on woman i's feed, she faces the binary choice of either liking the man or staying single. Because a like is costless, a woman will like a man when matching with him provides a greater utility than remaining single. Thus, the probability that woman i likes the man j depends on her idiosyncratic preference for man j (ϵ_{ij}), as well as her idiosyncratic preference for remaining single (ϵ_{ii}):

$$P(i \text{ likes } j) = P(V_{ij} > V_{ii}) = P(\beta P_j - \gamma P_i > \epsilon_{ii} - \epsilon_{ij}) \tag{3}$$

Assuming that idiosyncratic preferences ϵ are i.i.d. normally distributed with mean 0 and variance $\frac{1}{2}$, we can identify the coefficients β and γ by estimating a probit model on all *likes* and *dislikes* made by women on men (Table A.8). We then use these estimates to compute the distribution of the vertical component of women's utility ($\hat{\beta} P_j - \hat{\gamma} P_i$) across all woman-man pairs and the distribution of the horizontal component of women's utility ϵ_{ij} across all pairs.

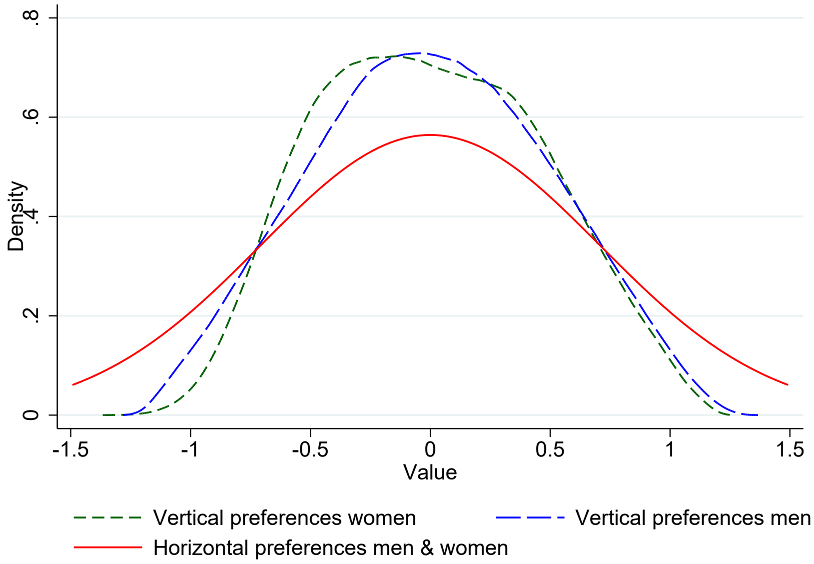
Figure A.7 plots these two distributions. The variance of women's vertical preferences (0.22) is smaller than the variance of the horizontal preferences (0.5). However, this exercise suggests that at least 31 percent of the variance in women and men's preferences is due to verti-

[44]We present the model that determines women's utility, but the same model applies for men.

[45]The assumption that the outside option value of an agent is a linear function of the own value also follows theoretical results in Menzel (2015) and Peters (2010). Menzel (2015) shows, in a very similar environment, that the match an agent on one side of the market attains in equilibrium is an increasing function of his own value. Peters (2010) shows in a directed search model of the labor market with two-sided heterogeneity that the reservation wage of workers is an increasing function of their productivity.

cal preferences.[46] This result means that the popularity ranking brings meaningful information on men and women's preferences in our dating market.

Figure A.7: Dispersion of vertical and horizontal preferences



Notes: This figure plots the dispersion in vertical and horizontal preferences. We estimate vertical preferences of women using a probit model with $y_{ij} = \beta P_j - \gamma P_i + (\epsilon_{ii} - \epsilon_{ij})$, where y_{ij} represents an indicator for whether woman i liked man j, P_j represents man j's popularity, P_i represents woman i's popularity, and $(\epsilon_{ii} - \epsilon_{ij})$ represents a standard normally distributed error term. We estimate women and men's preferences based on all ratings between January 1st 2014 and December 31st 2015 that are made between pairs with a leave out rank, which corresponds to 5,398,981 ratings of 4,824 women on men and 8,561,333 ratings of 9,381 men on women (Table A.8). We then use the estimates from the probit model to calculate $\hat{\beta}P_j - \hat{\gamma}P_i$ for every possible woman-man pair. The distribution of $\hat{\beta}P_j - \hat{\gamma}P_i$ represents the vertical preferences of women in this Figure. We estimate vertical preferences of men analogously. The estimated variance of horizontal preferences is equal to .5.

[46]Figure A.7 also shows the distribution of men's vertical preferences (0.24), which we obtain analogously. We calculate the variance in women and men's preferences that is due to vertical preferences as $.22/(.22 + .5) = .31$ for women and $.24/(.24 + .5) = .32$ for men.

Table A.8: Estimation of vertical preferences

	Woman rating man	Man rating woman
	(1)	(2)
Partner popularity ($\hat{\beta}$)	0.0173***	0.0158***
	(0.0000)	(0.0000)
Own popularity ($-\hat{\gamma}$)	-0.0091***	-0.0109***
	(0.0000)	(0.0000)
Observations	5,398,981	8,561,333

Notes: This table shows the coefficients from estimating the following regression using probit: $y_{ij} = \beta P_j + (-\gamma)P_i + \eta_{ij}$, where y_{ij} represents an indicator whether woman (man) i likes man (woman) j ($y_{ij} = 1$ and $y_{ij} = 0$ otherwise), P_j denotes the man's (woman's) popularity and P_i denotes the woman's (man's) popularity. $\eta_{ij} = \epsilon_{ii} - \epsilon_{ij}$ denotes a standard normally distributed error term. We estimate the regression on all ratings made by women (men) on men (women) with a leave out rank between January 1st 2014 and December 31st 2015. *** denotes significance at the 1 percent level. ** significance at the 5 percent level. * significance at the 10 percent level.

Chapter 3

How Directed is Search in the Labor Market? Evidence from an Online Job Board

How Directed is Search in the Labor Market? Evidence from an Online Job Board

Tobias Lehmann

October 2022

Abstract

This paper provides empirical evidence on how job seekers and recruiters search for matching partners. I use data from a job board where job seekers can apply for vacancies, and recruiters can apply to job seekers. I derive predictions for recruiter and job seeker behavior if meetings are fully random (random search), if recruiters and job seekers direct search to high quality partners (directed search), and if they direct search to high quality partners, but not to those for which they have low matching chances (competitive search). I show that I can estimate a quality ranking for vacancies and job seekers from the network of applications between them. I use the ranking to empirically study application behavior as a function of own quality and partner quality. I find that job seekers and vacancies are less likely to apply to low quality partners. I thus reject random search. Job seekers and vacancies are also less likely to apply to very high quality partners, in particular if they are themselves of low quality. This suggests that considerations over matching chances implied by competitive search are important.

Keywords: Directed search, competitive search, online job search, two-sided markets, eigenvector centrality, revealed preferences.

JEL classification: D83, J64

1 Introduction

Search models are among the most important tools economists use to study labor markets. The backbone of every search model is an assumption on the process of how job seekers and firms meet each other. The most widely-used models assume *random search* (McCall, 1970; Burdett and Mortensen, 1998), meaning that every job seeker meets with every vacancy with the same probability.[1] The shortcoming of random search models is that they are unable to accommodate that preferences of job seekers and firms might affect whom they meet.[2] An alternative meeting process, typically referred to as *directed* or *competitive search*, is one in which firms and job seekers care about the quality of their matching partner, and take into account that they are competing with others for a match when deciding where to apply (Rogerson et al., 2005; Wright et al., 2021).[3]

From a theoretical point of view, competitive search is preferable over random search because it is a micro-founded description of the search process while random search is a crude approximation. However, it is a priori not clear which one has more empirical content. At least two reasons could lead job seekers (or firms) to behave more like under random search than under competitive search: job seekers might have noisy or biased beliefs about vacancy quality and about competition by other job seekers, and job seekers may be unable or unwilling to make the often quite complex strategic considerations implied by competitive search (Nagel, 1995; Camerer et al., 2004; Camerer, 2011).

In this paper I provide evidence on how job seekers apply for vacancies as a function of vacancy quality and their own quality, and how vacancies apply to job seekers as a function of job seeker quality and their own quality. Two features of my study are novel with respect to the existing literature: first, I estimate job seeker and vacancy quality from revealed preferences, rather than relying on observable characteristics.[4] Second, this paper is, to the best of my

[1] Variants of this model include models where firms post contracts and job seekers apply, and where job seekers and firms negotiate contracts after meeting (Postel-Vinay and Robin, 2002; Mortensen, 2003).

[2] Random search is convenient because it allows for tractable models of the labor market that are able to reproduce important empirical patterns related to unemployment, vacancies, and business cycles (see, e.g., Rogerson et al., 2005; Schmieder et al., 2012; Lalive et al., 2006; Jäger et al., 2020).

[3] More precisely, *directed search* describes the environment, while the resulting equilibrium behavior of agents is called a *competitive search equilibrium* (Wright et al., 2021).

[4] Vacancy quality is typically measured through wage, while several studies show that non-wage amenities represent a large share of the utility that job seekers derive from their job (Hall and Mueller, 2018; Sorkin, 2018; Taber and Vejlin, 2020).

knowledge, the first to provide evidence on directed search on the firm side. This is important because nowadays many job boards are two-sided (e.g., Linkedin, Upwork), meaning that firms can also contact job seekers and invite them to apply for jobs.[5]

I rely on data from a Swiss online job board. On this platform, job seekers create profiles by adding their education and work history, and specifying the type of jobs they are interested in.[6] Firms post vacancies on the platform, which include job characteristics and work requirements, but no wage.[7] Every time they log in to the platform job seekers are shown vacancies that match their profile. They can then apply to these jobs on the platform. Similarly, firms are shown job seekers that match a particular vacancy, and can likewise apply to these job seekers on the platform. I use data on interactions between 7,300 job seekers and 3,900 vacancies in the year 2018.

I measure job seeker and vacancy quality using the revealed preference ranking for two-sided markets developed in Lehmann et al. (2022). I estimate the ranking using the network of applications between vacancies and job seekers. The intuition is that a job seeker is ranked high if many vacancies apply to him, and if these vacancies are ranked high. Similarly, a vacancy is ranked high if many job seekers apply to it, and if those job seekers are ranked high. I show that the ranking has high power in predicting whether an application will lead to a job interview.

I use the information on vacancy and job seeker quality to test whether and how they direct search for matching partners. Specifically, I test three competing hypotheses on the search process: first, random search, which implies that the probability that vacancies and job seekers apply to each other is independent of their quality. Second, directed search with respect to partner quality, which implies that the probability of applying increases in partner quality. Third, competitive search, meaning that job seekers and vacancies take into account both partner quality and the fact that they are competing with other agents for a match. The latter mechanism means that job seekers and vacancies are aware that their matching chances are lower for higher quality partners. Competitive search implies that the probability of applying increases in partner quality, except when partner quality is much higher than own quality.

I find that the probability a job seeker applies for a vacancy increases in vacancy quality

[5] I refer to this using the term "vacancy apply to job seeker" henceforth.

[6] The job board is used by both unemployed as well as employed job seekers.

[7] Like in many other European countries, it is not common in Switzerland to indicate the wage, or a wage range, in a vacancy.

for vacancies in the lower half of the quality distribution. Consistent with previous studies, I thus reject random search on the job seeker side (Banfi and Villena-Roldan, 2019; Marinescu and Wolthoff, 2020; Belot et al., 2022). I show that the probability of applying decreases for vacancies that are of much higher quality than the job seeker. Competitive search thus also plays a role for job seeker behavior. This is consistent with Faberman and Menzio (2018) and Belot et al. (2022).[8] Looking at the behavior of firms, I find that for low-quality job seekers the probability a vacancy applies to a job seeker increases in job seeker quality. I thus reject random search on the firm side. I show that the probability a vacancy applies decreases for job seekers that are of much higher quality than the vacancy. This implies that considerations from competitive search also play a role on the firm side.

This paper tests predictions from the rich theoretical literature on directed search (Wright et al., 2021). This literature typically assumes that firms post utility levels, and job seekers decide where to apply.[9] It derives equilibrium patterns in environments that differ, for example, with respect to the extent of heterogeneity on the job seeker and firm side, or with respect to how job seekers are incentivized to select where to apply.[10] A few studies have provided empirical evidence on the predictions of this theoretical literature (Holzer et al., 1991; Dal Bó et al., 2013; Faberman and Menzio, 2018; Banfi and Villena-Roldan, 2019; Marinescu and Wolthoff, 2020; Belot et al., 2022). A common finding of these studies is that firms offering higher wages receive more applications, thus suggesting that job seekers direct their search to vacancies of higher quality.[11] Evidence provided in Faberman and Menzio (2018) and Belot et al. (2022) suggests that some job seekers do not apply for vacancies offering high wages because they expect more competition from other job seekers.[12]

This paper also relates to the literature evaluating how information affects search in two-

[8]Competitive search behavior has also been found in laboratory experiments (Cason and Noussair, 2007; Anbarci and Feltovich, 2013; Kloosterman, 2016; Helland et al., 2017).

[9]This literature typically summarizes vacancy utility levels in their wages.

[10]For example, Shi (2002) and Shimer (2005) allow for different types of job seekers, while Peters (2000) allows for different vacancy types. Eeckhout and Kircher (2010) and Peters (2010) allow for heterogeneity in both job seeker types and the vacancy types. Job seekers are incentivized to direct applications because either the number of applications is assumed to be limited (e.g., Shimer, 2005; Galenianos and Kircher, 2009; Peters, 2010), or because there is a non-zero cost of applying (e.g., Albrecht et al., 2006).

[11]This statement is true unless higher wages reflect compensating differentials for inferior non-wage characteristics.

[12]Faberman and Menzio (2018) infer this from survey data on vacancy wage and the number of applications vacancies receive. Belot et al. (2022) conduct a field experiment in which they show job seekers the same vacancy with two different wages, finding that some job seekers only apply to the lower wage vacancy.

sided markets. Kuhn et al. (2020), Leibbrandt and List (2018) and Ibañez and Riener (2018) show how employer preference signaling affects how job seekers direct search. Gee et al. (2017) and Bhole et al. (2021) show how job seekers redirect search in response to information on competition. Skandalis (2018) shows that newspaper coverage about firm expansion leads job seekers to redirect search to expanding firms.

The rest of the paper is organized as follows. The next section presents the framework and the hypotheses. Section 3 explains the job search platform and shows descriptive statistics. In Section 4 I show how I measure the quality of vacancies and job seekers. Finally, I test the hypotheses on search behavior in Section 5, before concluding in Section 6.

2 Random, Directed, and Competitive Search

In the following, I will set up a stylized environment in which vacancies and job seekers search for matching partners. The environment matches all key features of the platform in the empirical section. I will then use the environment to illustrate my hypotheses on job seeker and vacancy behavior in three cases: first, when search is random, second, when search is directed with respect to partner quality, and third, when search is directed with respect to partner quality and the probability of matching. I refer to the latter as competitive search in the vein of Wright et al. (2021).

An environment with two-sided search by heterogeneous agents There is a set of job seekers $i \in I$ and a set of vacancies $j \in J$. Each job seeker is of type $s_i \in [0, 1]$, and each vacancy is of type $v_j \in [0, 1]$. Job seekers hold beliefs about vacancy types $\hat{v}_j \in [0, 1]$ and vacancies hold beliefs about job seeker types $\hat{s}_i \in [0, 1]$. Job seekers and vacancies can match with exactly one partner from the other side. When job seeker i matches with vacancy j, the job seeker gets payoff $v_j + \epsilon_{ij}$, where ϵ_{ij} represents the idiosyncratic taste of job seeker i for vacancy j. Similarly, when vacancy j matches with job seeker i, the vacancy gets payoff $s_i + \epsilon_{ji}$, where ϵ_{ji} represents the idiosyncratic taste of vacancy j for job seeker i.

Job seekers act as follows:

1. Search for vacancies and apply to them if $\hat{P}(i \text{ will match with } j) * (\hat{v}_j + \epsilon_{ji}) > c$, that is, if the expected payoff for job seeker i from applying for vacancy j exceeds the cost

c of applying. $\hat{P}(i \text{ will match with } j)$ denotes job seeker i's belief he will match with vacancy j in case he applies.

2. They evaluate all applications they receive from vacancies, and accept them at cost c if $\hat{P}(i \text{ will match with } j) * (\hat{s}_i + \epsilon_{ij}) > c$.

Vacancies act as follows:

1. Search for job seekers and apply to them if $\hat{P}(j \text{ will match with } i) * (\hat{s}_i + \epsilon_{ji}) > c$, that is, if the expected payoff for vacancy j from applying to job seeker i exceeds the cost c of applying.
2. They accept applications from job seekers at cost c if $\hat{P}(j \text{ will match with } i) * (\hat{s}_i + \epsilon_{ji}) > c$.

Actions of stage 1 are taken simultaneously on the job seeker and the vacancy side, followed by actions of stage 2, also simultaneously on both sides. Final one-to-one matches are formed between job seekers and vacancies where either the job seeker has applied for the vacancy and the vacancy has accepted the application, or the vacancy has applied to the job seeker and the job seeker has accepted the application.[13]

Optimal behavior of vacancies and job seekers in stage 1 can be derived using a backwards induction argument. Intuitively, there are two reasons for job seekers (and vacancies) to avoid c in stage 1: first, the believed payoff in case a match will happen $(\hat{v}_j + \epsilon_{ji})$ is too low, and second, the believed probability a match will happen $\hat{P}(i \text{ will match with } j)$ is too low.

Search behavior of vacancies and job seekers I will now illustrate my three hypotheses on the optimal behavior of vacancies and job seekers in stage 1, depending on the belief they hold about s_i, v_j, $P(j \text{ will match with } i)$ and $P(i \text{ will match with } j)$. I will only formulate the hypotheses from the job seekers' point of view, noting that due to the symmetry of the platform the corresponding hypotheses for vacancies follow the same line of argument.

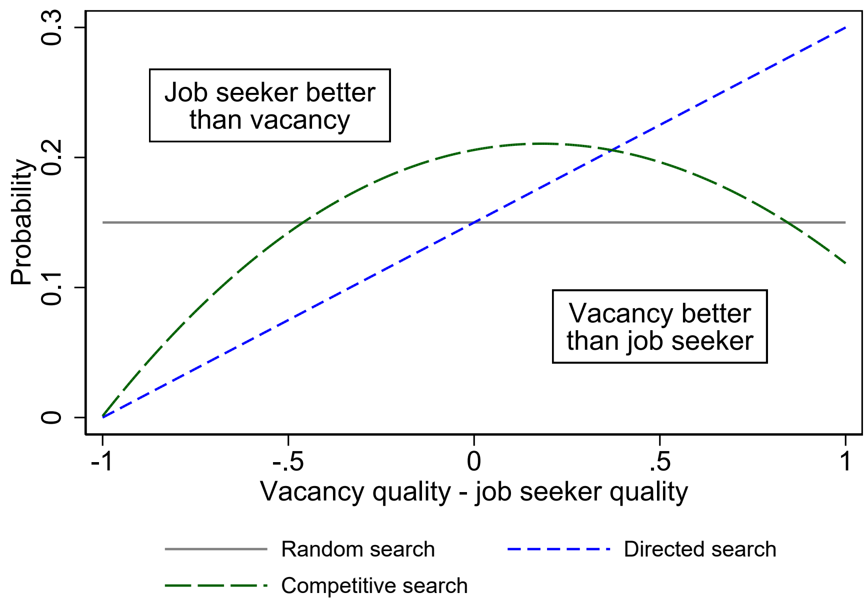
Hypothesis 1 *(random search)*: If job seekers have no information on the quality of vacancies and on the probability they will match with a vacancy, thus $\hat{P}(i \text{ will match with } j) =$

[13]The mechanism with which final matches are formed does not matter, as long as ex ante $P(i \text{ will match with } j)$ increases in s_i and $P(j \text{ will match with } i)$ increases in v_j. A sufficient condition for the mechanism to satisfy this criteria is that it represents a stable matching in the spirit of Gale and Shapley (1962).

$\bar{P}$ and $\hat{v}_j = \bar{v}$ are both constant, then applications of job seekers for vacancies occur if $\bar{P} * (\bar{v} + \epsilon_{ji}) > c$ and are thus "random" with respect to vacancy quality.

This reflects the trivial case where job seekers have no information whatsoever about the quality of a vacancy they are considering, nor about anything that helps them obtain an informed belief about $P(i \text{ will match with } j)$.[14] Thus, they will just apply to any vacancy for which they have a sufficiently high idiosyncratic taste.

Figure 1: Probability of applying under different search hypotheses



Notes: This figure illustrates the probability a job seeker will apply for a vacancy under different hypotheses on search behavior, depending on the difference in quality between the vacancy and the job seeker. Quality is such that 1 denotes the highest quality and 0 denotes the lowest quality.

The gray line in Figure 1 shows the resulting probability a job seeker will apply for a vacancy.[15] The x-axis represents the difference between the quality of a vacancy the job seeker is considering, and his own quality. The value 1 on the horizontal axis represents the case where the lowest-quality job seeker is considering the highest-quality vacancy, and the value -1 the case where the highest-quality job seeker is considering the lowest-quality vacancy. The flat gray line in Figure 1 reflects that the probability a job seeker will apply for a vacancy is unrelated to both the job seeker and the vacancy quality.

[14] $\hat{v}_j$ and $\hat{P}(i \text{ will match with } j)$ do not actually need to be constant across vacancies, but can also be random, for random search to be the resulting strategy of job seekers.

[15] The scale of the vertical axis in Figure 1 is arbitrary.

Hypothesis 2 ***(directed search)***: If job seekers have perfect information on the quality of vacancies, thus $\hat{v}_j = v_j$, but no information on the probability they will match with a vacancy, thus $\hat{P}(i \text{ will match with } j) == \bar{P}$ is constant, then applications of job seekers to vacancies occur if $= \bar{P} * (v_j + \epsilon_{ji}) > c$ and thus monotonically increase in vacancy quality.

This reflects the case where search is directed with respect to vacancy quality. It is represented by the blue line in Figure 1. Intuitively, for job seekers of any quality s_i expected utility from applying for a vacancy monotonically increase in vacancy quality v_j, because the job seeker believes the probability he eventually will get offered a job is independent of the quality of the vacancy as well as his own quality.

Hypothesis 3 ***(competitive search)***: If job seekers have accurate information on the quality of vacancies and the probability they will match with a vacancy, thus $\hat{v}_j = v_j$ and $\hat{P}(i \text{ will match with } j) = P(i \text{ will match with } j)$, then applications of job seekers to vacancies occur if $P(i \text{ will match with } j) * (v_j + \epsilon_{ij}) > c$. Applications will be a non-monotonic function of $v_j - s_i$. For low values of $v_j - s_i$, application probability increases in $v_j - s_i$, and for high values of $v_j - s_i$, application probability decreases in $v_j - s_i$.

This reflects the case in which job seekers and vacancies play a game with full information on players, types, and actions, where the sole private information is each job seeker's (vacancy's) idiosyncratic utility ϵ_{ij} (ϵ_{ji}).

The green line in Figure 1 illustrates the application probability in this case. The application probability increases for low values of $v_j - s_i$, but decreases for high values of $v_j - s_i$. This is due to the trade off job seekers face between the probability of matching $P(i \text{ will match with } j)$ and the value of a match v_i. To see the intuition behind this trade off, let us consider the decision faced by a job seeker of low quality. Contacting low quality vacancies is unattractive for this job seeker because the value in case of a match is relatively low. On the other hand, it is likely that no better job seeker will apply for the vacancy, thus the probability of matching when applying is high.[16] The exact opposite holds for the job seeker when he considers applying for a high

[16] An additional element that enters $P(i \text{ will match with } j)$ is the probability vacancy j ends up being the best vacancy that wants to match with job seeker i, out of the set of all vacancies that want to match with job seeker i.

quality vacancy. While the probability that no better job seeker will apply for the vacancy is low, the value in case of a match is high. The inverse U-shape in Figure 1 reflects this trade off.

A full derivation of equilibrium behavior for job seekers and vacancies in this competitive search environment is beyond the scope of this article.[17] I will instead now relate hypothesis 3 to the existing theoretical literature on competitive search.

Relation to theoretical literature on competitive search The key characteristics of the environment above are non-transferable utility, two-sided heterogeneity, multiple applications, application costs, and two-sided search. While noone has studied competitive search equilibria in an environment that combines these characteristics, all characteristics have appeared in theoretical studies on two-sided matching and directed search. Galenianos and Kircher (2009) and Albrecht et al. (2006) study environments with multiple applications. Peters (2010) and Eeckhout and Kircher (2011) study environments with two-sided heterogeneity, but only one single application.[18] All these papers find that job seekers trade off high wages and low employment probabilities on the equilibrium path, which is in line with my competitive search hypothesis. From the literature on stable matchings in two-sided one-to-one matching markets with non-transferable utility, Menzel (2015) shows that equilibrium matching probabilities are an increasing function of the own quality and a decreasing function of the partner's quality, which is also in line with my competitive search hypothesis.

3 A Two-Sided Job Board

I use data from a Swiss online job board. The platform is fully two-sided in that it allows job seekers to apply for vacancies, and recruiters to apply to job seekers.

Job seekers create a profile using their Linkedin login. The platform sources user name, work history, education history, skills, and the user's language from Linkedin. Every time a job seeker connects to the platform he sees a list of vacancies (Figure 2a). The list contains vacancies that are deemed suitable for the job seeker by yooture's matching algorithm, where

[17] The high complexity of the problem mainly has two sources: first, the fact that search happens on both sides, and second, that each $P(i$ will match with $j)$ and $P(j$ will match with $i)$ are functions of the type and strategies of all job seekers and vacancies in the market.

[18] All these studies assume that firms post wages and that job seekers search and apply.

the most suitable vacancies are listed first.[19] Job seekers can click on a vacancy, in which case they get to the vacancy detail view shown in Figure 2b. Here, job seekers can apply for a vacancy by clicking the *APPLY NOW* button. When doing so they get to a form where they can add a short cover letter, and submit the application.[20]

Recruiters create their profile by adding details on their company's name, industry, and location. After profile creation recruiters can create vacancies.[21] For each of their vacancies, recruiters see a list of suitable job seekers (Figure 3).[22] Recruiters can apply to job seekers by clicking the *SEND MESSAGE* button. When doing so, they get to a form where they can add a short cover letter, and submit the application.

When vacancies and job seekers receive an application, they are notified and can either accept or reject the application. When accepting an application, job seekers and vacancies continue their conversation, for example to provide further details or to arrange a meeting. The platform does not oblige vacancies and job seekers to respond to applications. I treat applications that are rejected and those that remain unanswered both as "not accepted" in my analysis.

I use data from January 1st 2018 to December 31st 2018. Table 1 describes the sample. It consists of 7,314 job seekers and 3,910 vacancies. Job seekers are active on the platform for an average of 28 weeks, and apply to 9.4 jobs in that time. The average vacancy duration is 6.2 weeks, in which time they apply to 11.3 job seekers.[23] The platform hosts vacancies and job seekers across all industries and regions, while there is some concentration of vacancies in information technology and the finance, legal & insurance industry, as well as in the metropolitan area of Zurich (Panel B and C of Table 1).

Panel D of Table 1 reports summary statistics on job seeker education and work experience. While information on work experience is comprehensive, many job seekers do not report

[19]The matching algorithm calculates a matching score for every vacancy with the job seeker. The following elements lead to a higher score: An overlap in job title with the titles of the jobs in the job seeker's work history, geographical proximity, an overlap in skills parsed from the vacancy and the job seeker profile, and a general text overlap between the vacancy and the job seeker profile.

[20]The platform also scrapes vacancies from the world wide web, which are also shown to job seekers. Job seekers cannot apply directly through the platform to these vacancies. I thus cannot track applications to these vacancies and therefore exclude them from my sample.

[21]To crate a vacancy recruiters have to provide the job title and a job description.

[22]The same matching algorithm is used to select job seekers that are listed to vacancies as to select vacancies that are listed to job seekers.

[23]The vacancy duration is within the range of 4.4-7.1 weeks found in Mueller et al. (2018, Table 1) for Austria.

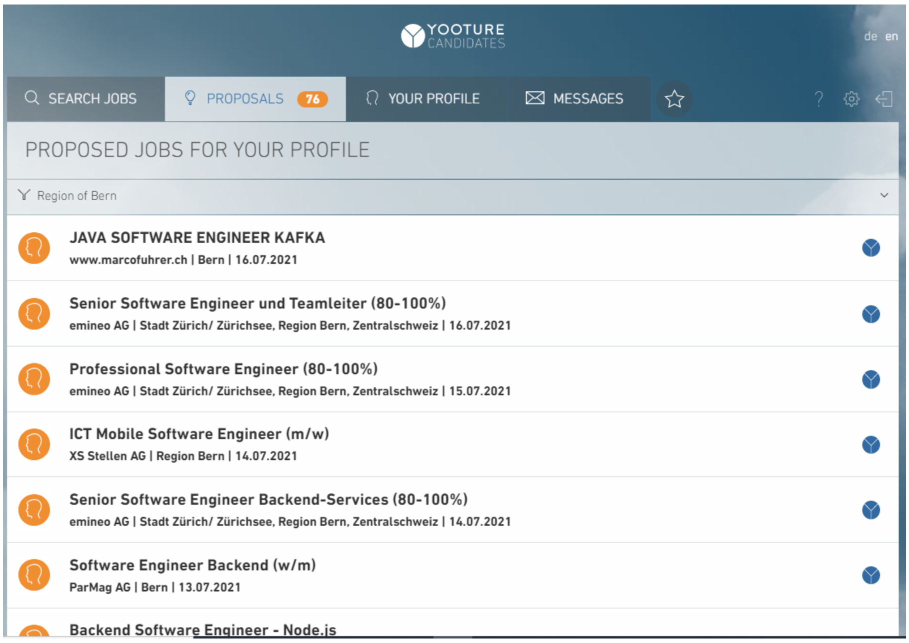
Table 1: Descriptive Statistics

	Job seekers (1)	**Vacancies** (2)
A. Activity		
Nb of applications	9.4	11.3
Nb of weeks active	28.0	6.2
Nb of logins	59.4	
B. Industry		
Construction	0.04	0.10
Manufacturing	0.13	0.15
Public admin. & education	0.02	0.01
Arts & entertainment	0.01	0.04
Information technology	0.05	0.22
Hotel and restaurant	0.02	0.03
Retail & cars	0.05	0.05
Finance, legal & insurance	0.08	0.25
Transportation	0.03	0.02
Health	0.03	0.03
Advertising & communication	0.02	0.05
Other	0.10	0.04
Unknown	0.42	0.00
C. Region		
Whole Switzerland	0.63	0.00
Lake Geneva Region	0.02	0.01
Espace Mittelland	0.10	0.13
Northwestern Switzerland	0.08	0.13
Central Switzerland	0.08	0.14
Ticino	0.01	0.00
Zurich	0.15	0.50
Eastern Switzerland	0.08	0.14
D. Education & work experience		
Highest education		
Vocational traning	0.36	
Bachelor	0.13	
Master	0.14	
Other	0.12	
Not reported	0.26	
Years of education	7.5	
Years of education not reported	0.84	
Nb of jobs held	3.9	
Years of work experience	11.4	
Work history not reported	0.05	
Observations	7,314	3,910

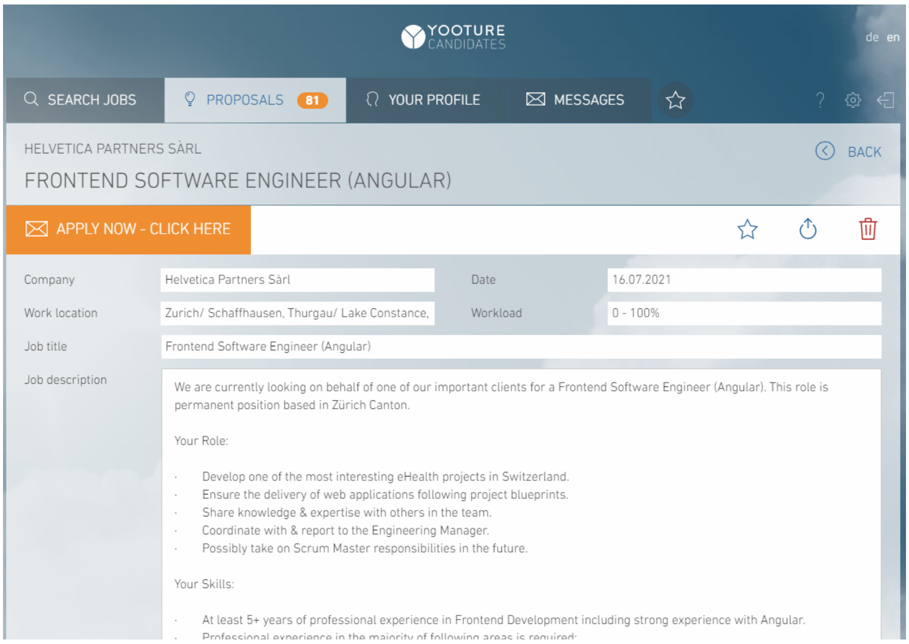
Notes: This table shows descriptive statistics on vacancies and job seekers in the sample, according to Section 4. *Panel A.*: Logins not tracked for vacancies. Vacancy duration measured for vacancies created before October 1st 2018. *Panel B.*: Industry of job seeker according to job seekers' current or last job. *Panel C.*: For job seekers: Shares as share of job seeker who restrict geographic search filter to corresponding region. For vacancies: Shares by vacancies' firm location.

Figure 2: Job seekers' view on yooture

(a) List of proposed vacancies



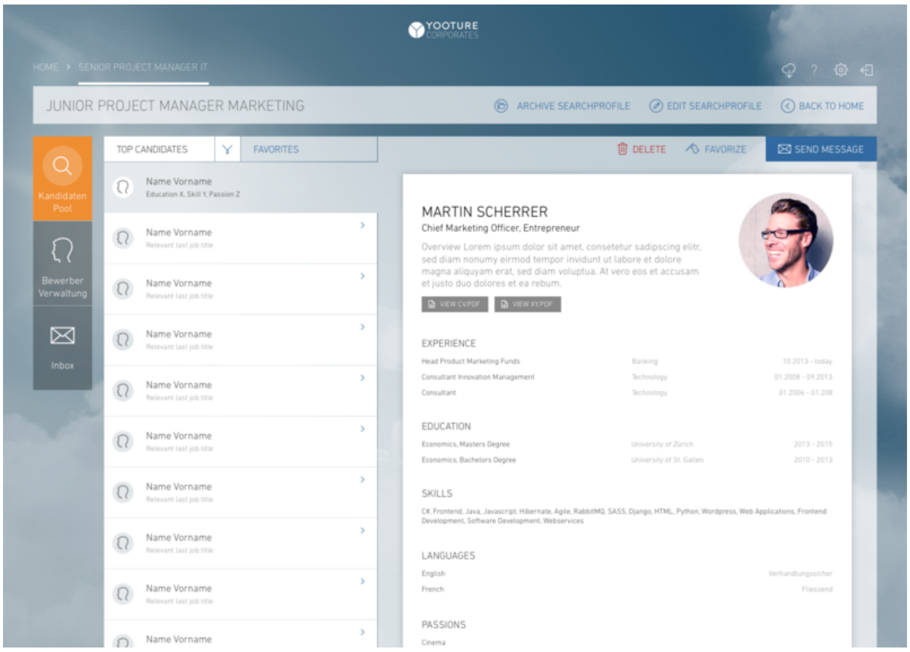
(b) Vacancy detail view



Notes: Panel a shows an example of the list of vacancies proposed to a job seeker. By clicking on a vacancy, the job seeker reaches the page shown in panel b, where he can apply for that vacancy.

their educational path and attainment. The broad absence of key indicators of quality, such as education for job seekers, working conditions for vacancies, and wages for both, makes it

Figure 3: Recruiters' view on yooture



Notes: This figure shows an example of the list of job seekers that are proposed to a vacancy.

difficult to estimate job seeker and vacancy quality through hedonic regressions.[24] In the next section I will thus introduce a method to estimate quality that solely relies on information on applications made between job seekers and vacancies.

4 Estimating Job Seeker and Vacancy Quality Using Revealed Preferences

In order to estimate the quality of vacancies and job seekers, I exploit the feature of the platform that not only job seekers can apply for vacancies, but also vacancies can apply to job seekers. I use the network of applications to estimate the ranking for two-sided markets developed in Lehmann et al. (2022). In the following I will explain the intuition behind the ranking, while technical details can be found in Appendix B.

The ranking defines the quality of a vacancy recursively depending on both the number

[24]Scarcity of micro data also applies to many other online job boards. As education information on this platform is sourced from Linkedin, any lack of information on it in my data means that this information is also missing in the Linkedin profile of job seekers. Moreover, many job boards, for instance Glassdoor and CareerBuilder, allow job seekers to search and apply for jobs without providing any information apart from a job title and geographic preferences.

of applications the vacancy receives and the quality of job seekers from which the vacancy receives applications. Thus, a vacancy is ranked high if it receives applications from numerous highly ranked job seekers. The quality of job seekers is defined analogously, meaning that a job seeker is ranked high if he receives applications from numerous highly ranked vacancies.

This approach gives rise to a system of equations which has a unique solution. As I show in Appendix B.2, this solution corresponds to the ranking of job seeker and vacancy quality under both directed and competitive search. The ranking has no such interpretation if random search is the data generating process. However, it has a testable prediction: if random search is the data generating process, then the ranking must not have any predictive power over actions taken by vacancies and job seekers.

Like any other labor market, the platform I consider changes every day due to the entry and exit of job seekers and vacancies. To account for these dynamics, I estimate a ranking every day using applications from the past 90 days. I thus end up estimating 275 rankings, one for every day between April 1st 2018 and December 31st 2018. I then standardize every daily ranking on a scale from 0 (lowest quality rank) to 1 (highest quality rank).

Table 2: Descriptive statistics on job seeker and vacancy Popularity Ranking

	Job Seekers	**Vacancies**
Panel A: Statistics on all rankings		
Nb of interactions used to compute the rankings	20,067	19,074
Nb ranked at least once	7,314	3,910
Share ranked	39%	42%
Panel B: Statistics on ranking per day		
Nb of applications used to compute the ranking	3,740	3,992
Nb ranked	2,060	942
Panel C: Average rankings		
Nb of ranking days	77	66

Notes: This table shows descriptive statistics on the quality rankings for every day from April 1st 2018 to December 31st 2018. Panel A shows descriptive statistics on all rankings valid between April 1st 2018 and December 31st 2018. *Nb of interactions used to compute the rankings* refers to the number of applications made or accepted by job seekers and vacancies that are used in at least one of the 275 rankings. *Share ranked* shows the share of vacancies and job seekers that are ranked at least once, out of all vacancies created and job seekers active in the year 2018. Panel B shows how many applications a ranking is based on, on average, and how many job seekers and vacancies are ranked, on average. *Nb of ranking days* shows the average number of ranking days over which the average rank of a vacancy or a job seeker is calculated.

Table 2 Panel A shows descriptive statistics on the ranking. To compute the ranking, I use

20,067 applications from job seekers and 19,074 applications from vacancies. The ranking interprets an application as a revelation of interest in a partner. I therefore treat the acceptance of an application like an application to estimate the ranking. I can estimate the ranking for 7,314 job seekers and 3,910 vacancies. Each one of the 275 daily rankings on average covers 2,060 job seekers and 942 vacancies (Table 2 Panel B). I estimate the quality of job seekers and vacancies as their average rank over all days on which they are ranked. Job seekers on average are ranked 77 days and vacancies 66 days (Table 2 Panel C).

I can only estimate the ranking for job seekers and vacancies that are strongly connected by applications. The restriction concerns the network of applications between job seekers and vacancies. A job seeker is in the strongly connected set if he receives at least one application from a vacancy in the strongly connected set, and applies to at least one vacancy in the strongly connected set (and vice versa for vacancies).[25] Table 2 shows that this allows me to estimate the ranking for 39% of job seekers and 42% of vacancies.

Job seekers and vacancies in the strongly connected set are very similar to those not in the strongly connected set (Table A.1). The only major difference is that job seekers and vacancies in the stronlgy connected set send more applications. This is, however, not surprising given that sending many applications increases the chance of being in the strongly connected set.

5 Is Search Random, Directed, or Competitive?

I will now present a sequence of tests to determine whether search is random, directed with respect to partner quality, or competitive.

5.1 Predicting application success using the quality ranking

Figure 4 shows that the quality ranking has quite some predictive power about whether a vacancy or a job seeker will accept or reject an application. The green dashed line shows the probability a job seeker accepts an application from a vacancy (y-axis) as a function of his relative quality (x-axis), that is, as a function of the difference between the quality rank of the

[25] Having an own application accepted is treated like receiving an application to determine the strongly connected set. To increase the number of job seekers and vacancies for which I can estimate the ranking, I also include job seekers (vacancies) that only receive applications from vacancies (job seekers) in the strongly connected set. See Appendix B for details.

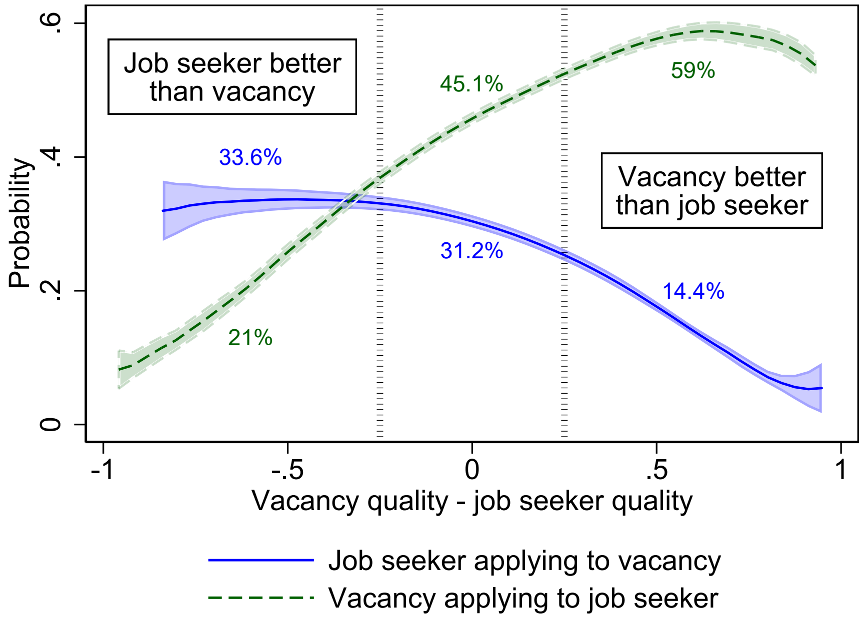
vacancy and the job seeker. The left side of the graph corresponds to cases where the vacancy is of lower quality than the job seeker, while the right side corresponds to the opposite.

We see that job seekers accept applications from vacancies that are of much lower quality than themselves (at least .25 quality ranks better) in 21% of the cases, while they accept applications of much higher quality vacancies in 59% of the cases. While job seekers prefer high quality vacancies, vacancies prefer high quality job seekers. The latter is depicted by the blue line in Figure 4. When a job seeker is of much higher quality than a vacancy, the vacancy will accept the application in 33.6% of the cases. On the other hand, when the job seeker is of much lower quality, the vacancy will accept the application in only 14.4% of the cases.[26]

Rejecting random search The figure thus shows that the quality ranking is a strong predictor of whether vacancies and job seekers accept applications. This provides strong evidence against random search being the data generating process. Intuitively, if random search were the data generating process, any ranking derived from application patterns would itself be random, and thus vacancy and job seeker rank should not predict acceptance of applications.

[26] Figure 4 shows probabilities of acceptance by *relative* quality, thus combining job seekers and vacancies of very different absolute quality (e.g., a job seeker with quality rank 1 and a vacancy with quality rank .8 have the same relative quality as a job seeker with quality rank .2 and a vacancy with quality rank 0). Figure A.2 shows acceptance probabilities for various combinations of absolute quality. Probabilities of acceptance are indeed similar for absolute quality combinations reflecting the same relative quality (i.e., along the diagonals in Figure A.2).

Figure 4: Probability an application is accepted by relative quality



Notes: This figure shows a local polynomial regression of a dummy variable indicating whether an application is accepted on the difference between the quality rank of the vacancy and the job seeker, separately for applications made by job seekers to vacancies (blue line) and by vacancies to job seekers (green line). Shaded areas depict 95% confidence intervals. Percentages shown correspond to share of applications that are accepted in corresponding quality rank difference bin ($< -.25$;$[-.25, .25]$;$> .25$). Based on 12,813 applications made by job seekers and 19,605 applications made by vacancies. The local polynomial uses an Epanechnikov kernel with bandwidth .2.

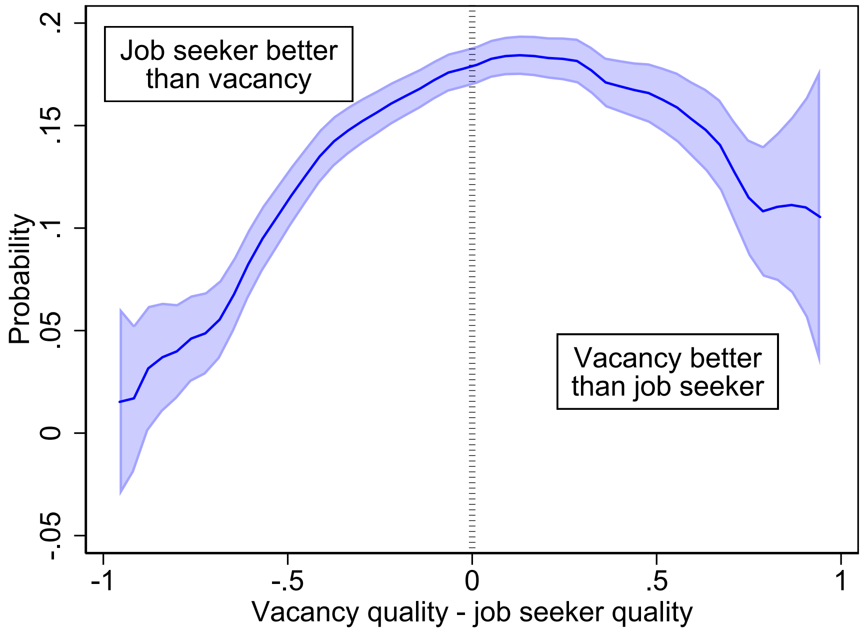
5.2 Providing evidence for directed and competitive search

Evidence from probability a job seeker applies For all vacancies job seekers see in their proposal list (Figure 6a) the platform registers when job seekers click on a vacancy to see the details of the vacancy (Figure 6b).[27] This allows me to calculate the probability with which viewing a vacancy leads to an application. The blue line in Figure 5 shows this probability, as a function of the difference in quality between the vacancy and the job seeker. As before, the left side of the graph corresponds to cases where the job seeker is of lower quality than the vacancy, while the right side corresponds to the opposite. The positive slope of the blue line in the left half of the graph reflects that the probability a job seeker applies to a vacancy increases when the vacancy has a lower quality rank than the job seeker. This is consistent with both directed and competitive search.

[27] Apart from the proposal list, job seekers can also apply for vacancies that they are alerted to via a push message, an email, or vacancies they find through the search mask. In 2018 job seekers found 72% of vacancies through the proposal list. Comparing observable characteristics of these vacancies does not suggest that they systematically differ from the vacancies job seekers find thorough these other channels.

Where the vacancy and the job seeker are of about the same quality rank, which is at the value 0 on the horizontal axis, this probability becomes flat. There is even evidence, although imprecisely estimated, that the probability of applying is lower when the vacancy is of a much higher quality rank than the job seeker, which is towards the right end of the graph.[28] Recalling the theoretical considerations in Section 2, this is evidence that job seeker search is competitive, rather than purely directed with respect to vacancy quality. The intuition is that job seekers refrain from applying to vacancies that are of a much higher quality than themselves because they believe the chances of eventually getting offered the job are low.

Figure 5: Probability job seeker applies after viewing vacancy



Notes: This figure shows a local polynomial regression of a dummy variable indicating whether a job seeker applies to a vacancy after he looks at it in the detail view (Figure 2b) on the difference between the quality rank of the vacancy and the job seeker. Shaded areas depict 95% confidence intervals. Based on 18,317 vacancies viewed and 2,991 applications made. The local polynomial uses an Epanechnikov kernel with Stata's default bandwidth.

Evidence from the distribution of vacancy to job seeker applications While the platform registers when a job seeker clicks on a vacancy in his proposal list, it unfortunately does not do so on the other side of the market, that is, when a vacancy clicks on a job seeker in its proposal list. This means that I cannot apply the test shown in Figure 5 to the vacancy side. However, as the platform still registers all applications made by vacancies to job seekers, I can compare

[28]Figure A.3 shows that this holds across different levels of job seeker quality.

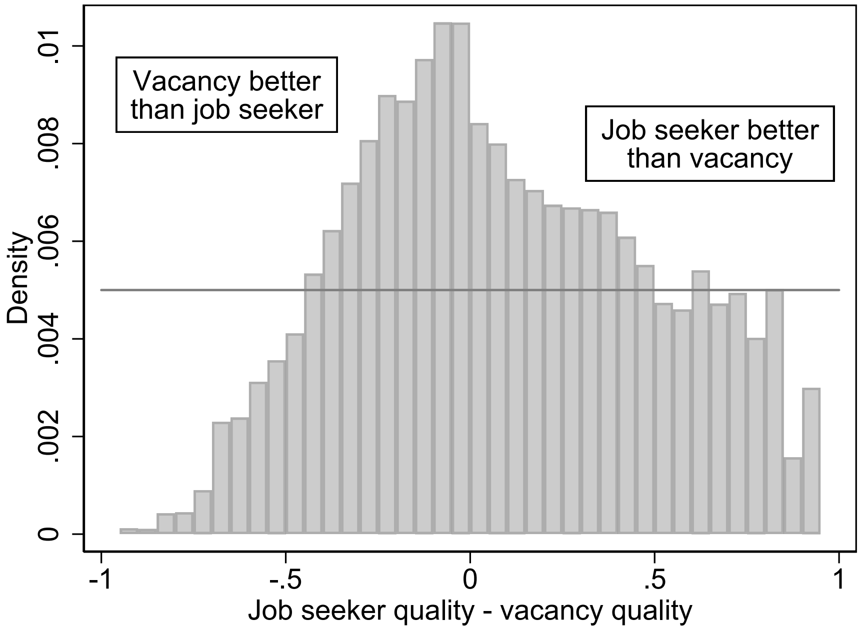
the distribution of applications made by vacancies to the distribution that would result from directed or competitive search.

To see the intuition behind this approach, consider the median quality vacancy, that is, the vacancy with quality rank .5. If search was random, the applications made by this vacancy should be evenly distributed among job seekers with any quality rank, meaning that the vacancy should make as many applications to job seekers with quality rank 0 as to job seekers with quality rank .5 and quality rank 1. If search was directed, then the vacancy should make more applications to job seekers with quality rank .5 than to job seekers with quality rank 0, and even more applications to job seekers with quality rank 1. If search was competitive, the vacancy should make more applications to job seekers with quality rank .5 than to job seekers with quality rank 0, but fewer applications to job seekers with quality rank 1 than to job seekers with quality rank .5. This means that the distribution of applications under random, directed and competitive search should have exactly the same shape as a function of quality difference as the probability of applying depicted in Figure 1.[29]

Figure 6 shows the density of applications made by vacancies to job seekers. Note that compared to the previous graphs the horizonal axis now represents job seeker quality − vacancy quality, which implies that the left side of the graph corresponds to cases where the vacancy is better than the job seeker, and the right side to the opposite. We see that the density is lowest at the very left end of the graph, and increases towards the middle. This means that vacancies more frequently apply to job seekers that are of a similar quality as themselves than to job seekers that are of a much lower quality, which is consistent with both directed and competitive search. The density then declines again towards the right end of the graph. This means that vacancies are less likely to apply to job seekers that are of a much higher quality, which is evidence for competitive search.

[29] I do the same analysis for applications made by job seekers (Figure A.4), which confirms the conclusion drawn from Figure 5.

Figure 6: Distribution of applications made by vacancies



Notes: This figure shows the distribution of applications made by vacancies by difference between the quality rank of the job seeker and the vacancy. The gray line represents the density that would result from random search. Based on 19,605 applications by vacancies to job seekers. Applications in this figure are weighted as follows to ensure the density represented by the gray line would result from random search: first, applications are binned by 0.01 quality rank of the vacancy making the application, and each application is weighted by $\frac{19{,}605}{\text{Nb of applications made by vacancies in quality rank bin}}$. This corrects for the differential intensity with which vacancies with a different quality rank apply to job seekers *independent of job seeker quality rank*. These weighted applications are then weighted by $\frac{1}{1-|\text{Job seeker quality - vacancy quality}|}$. This weight corrects for the fact that different values of $|\text{Job seeker quality - vacancy quality}|$ have a different ex ante probability of existing. To see this, consider $|\text{Job seeker quality - vacancy quality}| = 0$. This difference exists for any vacany quality in $[0, 1]$. $|\text{Job seeker quality - vacancy quality}| = .5$, however, only exists for vacancy quality in $[0, .5]$, and for that reason the density weight put on applications with $|\text{Job seeker quality - vacancy quality}| = .5$ is twice as high as that put on applications with $|\text{Job seeker quality - vacancy quality}| = 0$.

5.3 Discussion

My results strongly reject random search on both the job seeker and the vacancy side. This echoes findings from previous studies showing that vacancies offering higher wages attract more applicants (Holzer et al., 1991; Dal Bó et al., 2013; Faberman and Menzio, 2018; Banfi and Villena-Roldan, 2019; Marinescu and Wolthoff, 2020; Belot et al., 2022). One novelty of my study is that I directly measure quality from revealed preferences, while previous studies have measured vacancy quality through wage. This means that I avoid the problem that wage is an imperfect proxy for quality, for example, because of compensating differentials for amenities.[30] Another novelty relative to the existing literature is that I am, to the best of my knowledge, the first to study search behavior in the context of random, directed and competitive search on the firm side.

Probably the most important contribution of this paper is that it provides evidence that strategic considerations from competitive search theory play a role for both job seekers and vacancies. Previous studies have only provided very limited evidence on this. Faberman and Menzio (2018) provide survey-based evidence that the wage of a vacancy is negatively related to the number of applicants the vacancy gets, even within three-digit occupations. Belot et al. (2022) assign wages randomly to pairs of otherwise similar vacancies. They then show these vacancies to unemployed job seekers. They find that some job seekers only apply for the vacancy with the lower wage even when they were exposed to both vacancies. The findings of Faberman and Menzio (2018) and Belot et al. (2022) can be rationalized by job seekers trading off the wage of a vacancy with the probability of being hired. My finding that job seekers are less likely to apply for vacancies that are of a much higher quality is in line with the finding of these two studies. The novelty of my findings is that I also consider heterogeneity on the job seeker side, and show that job seekers are much less likely to apply if they are themselves of a much lower quality than the vacancy.[31]

Limitations One caveat related to my framework is that I only consider one systematic dimension in which vacancies and job seekers differ, which is their quality type. While this allows me to study search in the overall market, it means that I disregard characteristics of job seekers and vacancies other than quality type that might guide their search and application behavior. For example, one might be concerned that my finding that job seekers are most likely to apply for vacancies of a similar quality type arises because these vacancies also happen to be suitable in terms of characteristics such as job title and geographic location. Recall, however, that the sample in Figure 5 is only based on vacancies which the job seeker has clicked on, meaning that the job seeker has selected these vacancies based on job title, employer, and the location

[30] I refer to quality as a measure of the value a vacancy has for a job seeker. Apart from compensating differentials for amenities, the quality of a vacancy can also be imperfectly correlated with the current wage because of expectations about future wage flows (Jarosch, 2021).

[31] While Faberman and Menzio (2018) do not observe applicant characteristic, except for the applicant eventually hired, they show theoretically that their evidence can result from a directed search model with two-sided heterogeneity.

of the workplace. Figure 5 therefore shows that the clear pattern of application probabilities as a function of relative quality occurs *within* the set of vacancies a job seeker has considered suitable.

Another caveat is related to the way I calculate my quality ranking. Because it is based on applications between vacancies and job seekers, when a job seeker applies to a vacancy, or vice versa, this creates a mechanical relationship between the vacancy quality and the job seeker quality. One way to rule out any mechanical relationship between the job seeker and vacancy quality and search outcomes is to ensure the ranking of every vacancy-job seeker pair for which an outcome enters the sample is not based on any interaction between this vacancy-job seeker pair. This approach drastically reduces the number of applications each vacancy and job seeker's quality ranking is based on, thus making the ranking less precise.[32] In addition, it reduces the number of job seeker-vacancy interactions for which a quality rank exists for both the vacancy and the job seeker. Together, this makes the results much more noisy. Nevertheless, results obtained with this alternative approach are consistent with the findings above, thus rejecting random search and suggesting that strategic considerations from competitive search play a role (Figures A.5-A.8).

6 Conclusion

This paper studies search behavior using data from an online job board. I derive predictions of job seeker and firm search behavior under different assumptions on the information they possess about the quality of matching partners and the probability they will match with a given partner. The predicted search behaviors are random search, directed search with respect to partner quality, and competitive search. I estimate a ranking of job seeker and vacancy quality from the network of applications between them. Using the ranking, I test whether search is random, directed with respect to partner quality, or competitive. I find that competitive search matches best observed search behavior of firms and job seekers, which I infer from finding that firms and job seekers are most likely to apply to partners that are of a similar quality rank.

By clearly rejecting random search, this paper lends support to the literature emphasiz-

[32] I implement this approach by only considering the ranking of day $t-1$ to determine the quality of job seekers and vacancies applying or receiving an application on day t.

ing the importance of strategic considerations in the process leading to meetings between job seekers and firms (summarized in Wright et al. (2021)). This paper's approach to estimate the quality of job seekers and vacancies may also be useful in future studies on how partner quality and matching chances shape application behavior in the labor market. Fruitful avenues could be comparisons between women and men, or the study of changes in search behavior over the unemployment spell.

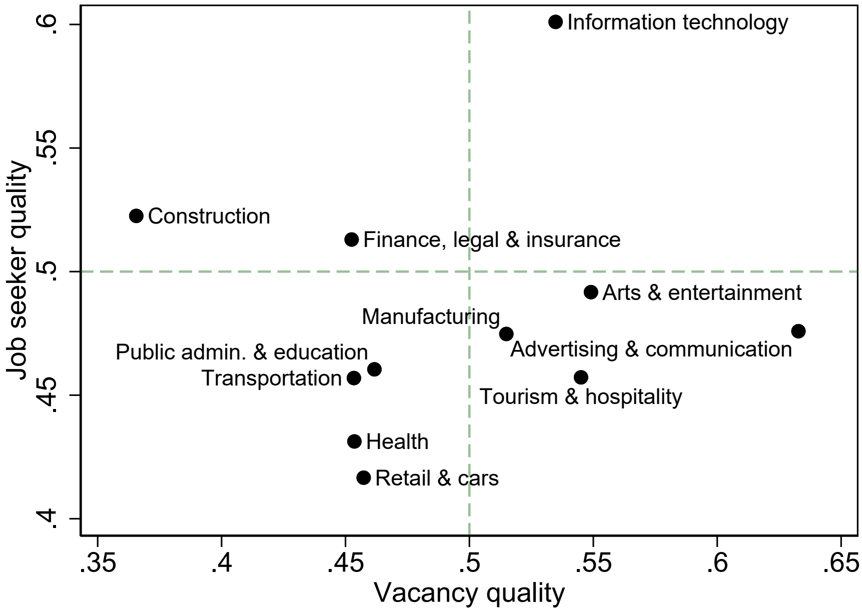
A Additional Tables and Figures

Table A.1: Descriptive Statistics

	Job seekers		Vacancies	
	All (1)	Sample (2)	All (3)	Sample (4)
A. Activity				
Nb of applications	5.3	9.4	5.8	11.3
Nb of weeks active	24.9	28.0	5.8	6.2
Nb of logins	48.7	59.4		
B. Industry shares				
Construction	0.05	0.04	0.08	0.10
Manufacturing	0.12	0.13	0.15	0.15
Public admin. & education	0.02	0.02	0.02	0.01
Arts & entertainment	0.01	0.01	0.04	0.04
Information technology	0.04	0.05	0.18	0.22
Hotel and restaurant	0.04	0.02	0.05	0.03
Retail & cars	0.05	0.05	0.03	0.05
Finance, legal & insurance	0.08	0.08	0.29	0.25
Transportation	0.03	0.03	0.02	0.02
Health	0.03	0.03	0.03	0.03
Advertising & communication	0.02	0.02	0.04	0.05
Other	0.10	0.10	0.06	0.04
Unknown	0.42	0.42	0.00	0.00
C. Region				
Whole Switzerland	0.64	0.63	0.00	0.00
Lake Geneva Region	0.02	0.02	0.01	0.01
Espace Mittelland	0.09	0.10	0.14	0.13
Northwestern Switzerland	0.08	0.08	0.15	0.13
Central Switzerland	0.08	0.08	0.14	0.14
Ticino	0.01	0.01	0.00	0.00
Zurich	0.15	0.15	0.47	0.50
Eastern Switzerland	0.08	0.08	0.13	0.14
D. Education & work experience				
Highest education				
Vocational training	0.35	0.36		
Bachelor	0.13	0.13		
Master	0.15	0.14		
Other	0.10	0.12		
Not reported	0.27	0.26		
Years of education	7.6	7.5		
Years of education not reported	0.86	0.84		
Nb of jobs held	3.7	3.9		
Years of work experience	11.2	11.4		
Work history not reported	0.06	0.05		
Observations	18,871	7,314	9,355	3,910

Notes: This table shows descriptive statistics on all vacancies and job seekers active in 2018 (column (1) and (3)) and of the sample of strongly connected vacancies and job seekers (column (2) and (4)), according to Section 4. *Panel A.*: Logins not tracked for vacancies. Vacancy duration measured for vacancies created before October 1st 2018. *Panel B.*: Industry of job seeker according to job seekers' current or last job. *Panel C.*: For job seekers (column (1) and (2)): Shares as share of job seekers who restrict geographic search filter to corresponding region. For vacancies (column (3) and (4)): Shares by vacancy firm location.

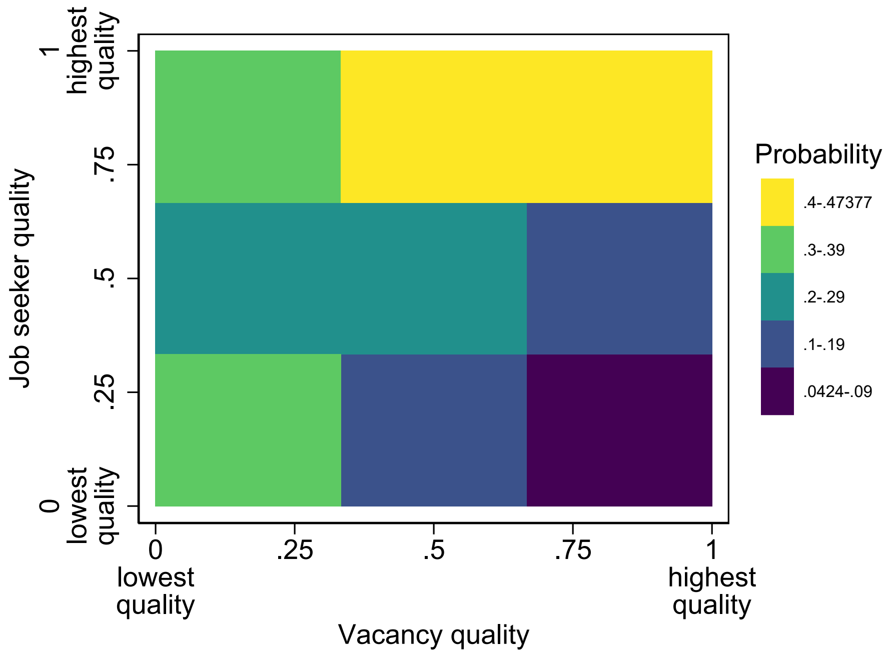
Figure A.1: By industry – average quality of vacancies and job seekers



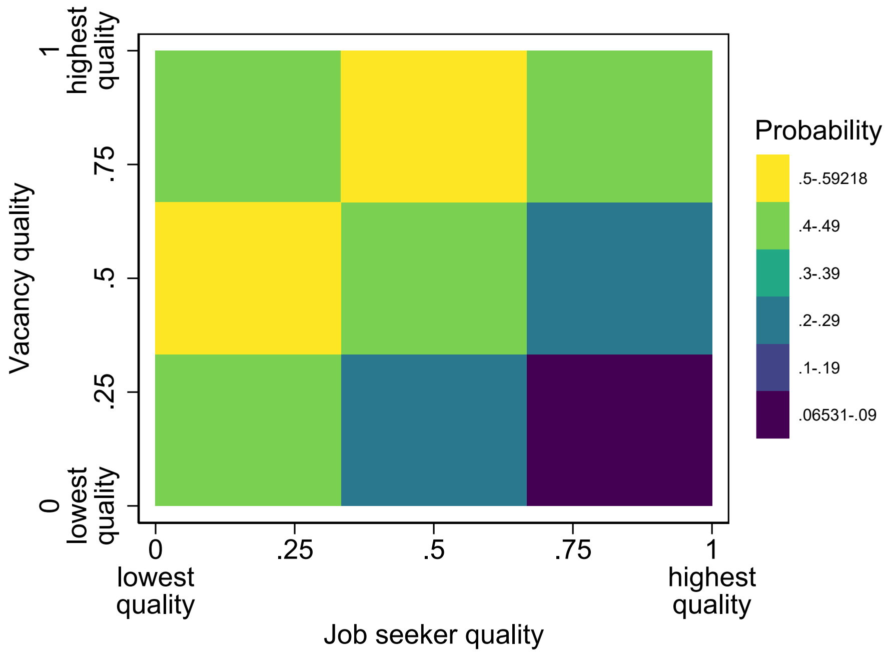
Notes: This figure shows the average quality rank of job seekers and vacancies by industry. Industry of job seeker is the industry of his current or last employer. Based on 268,027 ranking-day observations of 3,440 job seekers and 76,178 ranking-day observations of 1,190 vacancies.

Figure A.2: Share applications accepted by job seeker and vacancy quality

(a) Job seeker applying for vacancy

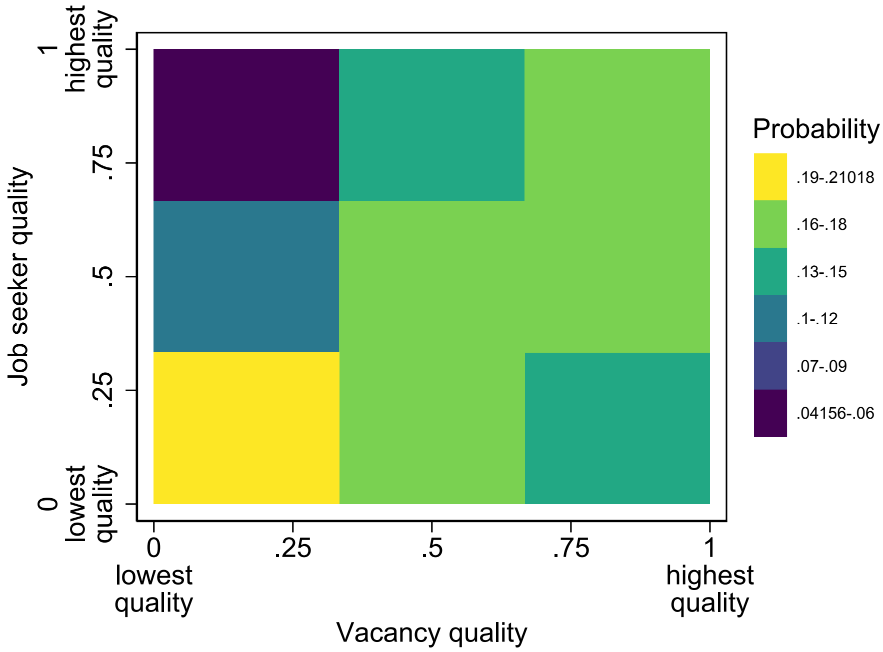


(b) Vacancy applying to job seeker



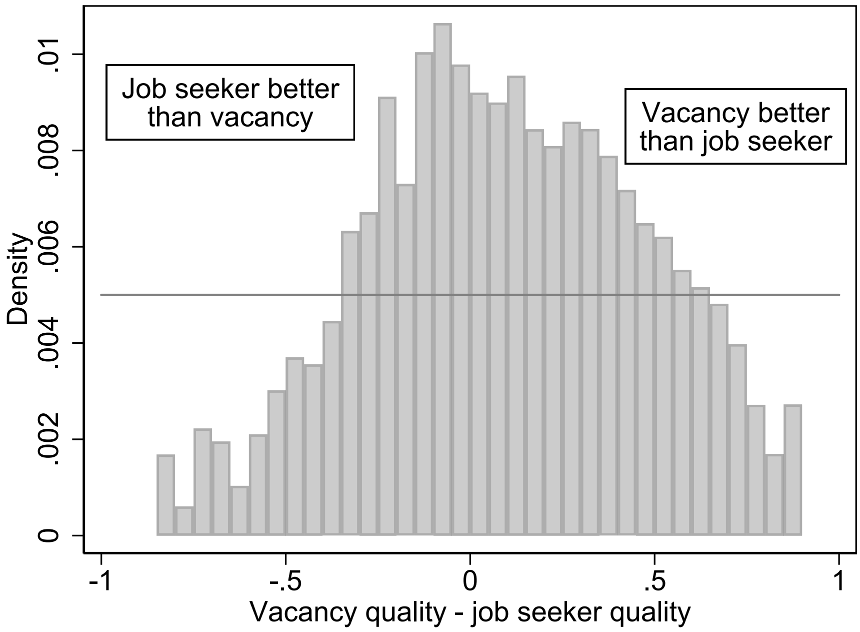
Notes: Figure (a) shows the probability a vacancy accepts an application from a job seeker. Figure (b) shows the probability a job seeker accepts an application from a vacancy. Figure (a) is based on 12,813 applications. Figure (b) is based on 19,605 applications.

Figure A.3: Probability job seeker applies after viewing vacancy



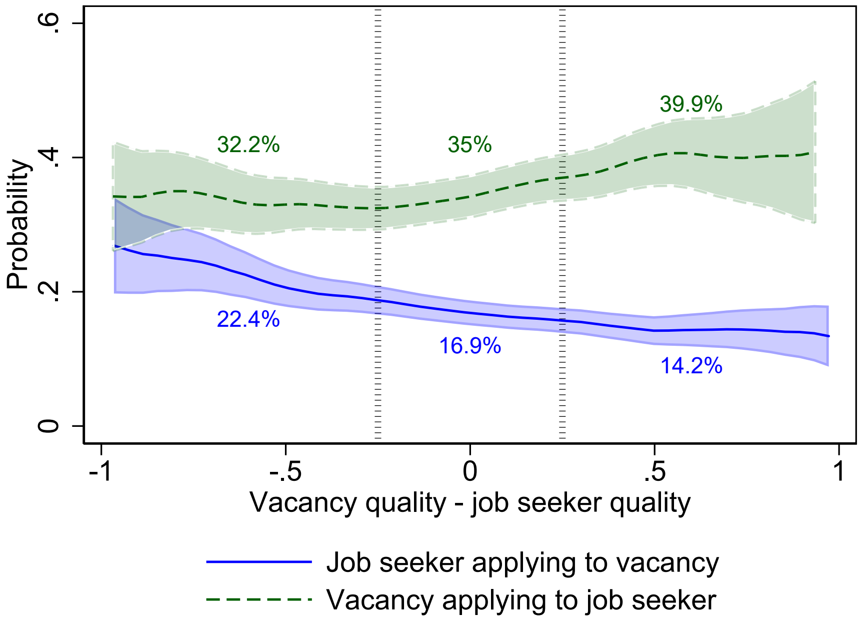
Notes: This figure shows the probability a job seeker applies to a vacancy after he looks at it in the detail view (Figure 2b). Based on 18,317 vacancies viewed and 2,991 applications made.

Figure A.4: Distribution of applications made by job seekers



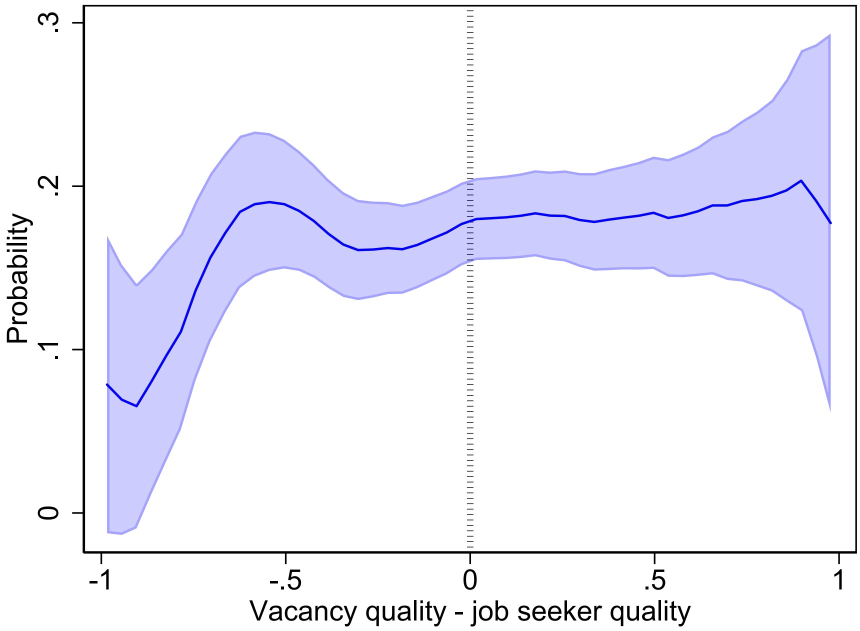
Notes: This figure shows the distribution of applications made by job seekers by difference between the quality rank of the vacancy and the job seeker. The gray line represents the density that would result from random search. Based on 12,813 applications by job seekers to vacancies. Applications in this figure are weighted as follows to ensure the density represented by the gray line would result from random search: first, applications are binned by 0.01 quality rank of the job seeker making the application, and each application is weighted by $\frac{12{,}813}{\text{Nb of applications made by job seekers in quality rank bin}}$. This corrects for the differential intensity with which job seekers with a different quality rank apply for vacancies *independent of vacancy quality rank*. These weighted applications are then weighted by $\frac{1}{1-|\text{vacancy quality - job seeker quality}|}$. This weight corrects for the fact that different values of $|\text{vacancy quality - job seeker quality}|$ have a different ex ante probability of existing. To see this, consider $|\text{vacancy quality - job seeker quality}| = 0$. This difference exists for any vacany quality in $[0, 1]$. $|\text{vacancy quality - job seeker quality}| = .5$, however, only exists for vacancy quality in $[0, .5]$, and for that reason the density weight put on applications with $|\text{vacancy quality - job seeker quality}| = .5$ is twice as high as that put on applications with $|\text{vacancy quality - job seeker quality}| = 0$.

Figure A.5: Out of sample ranking only: Probability of accepting an application by relative quality



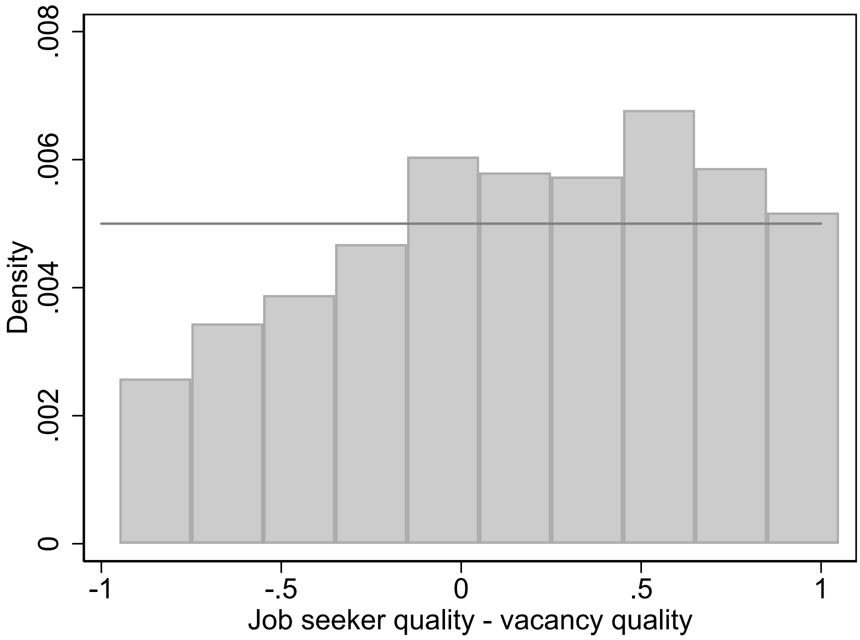
Notes: This figure shows a local polynomial regression of a dummy variable indicating whether an application is accepted on the difference between the quality rank of the vacancy and the job seeker, separately for applications made by job seekers to vacancies (blue line) and by vacancies to job seekers (green line). Shaded areas depict 95% confidence intervals. Percentages shown correspond to share of applications that are accepted in corresponding quality rank difference bin ($< -.25$;$[-.25, .25]$;$> .25$). The quality ranking used is the out of sample ranking as described in section 5.3. The figure is based on 2,440 applications made by job seekers and 1,406 applications made by vacancies. The local polynomial uses an Epanechnikov kernel with bandwidth .2.

Figure A.6: Out of sample ranking only: Probability job seeker applies after viewing vacancy



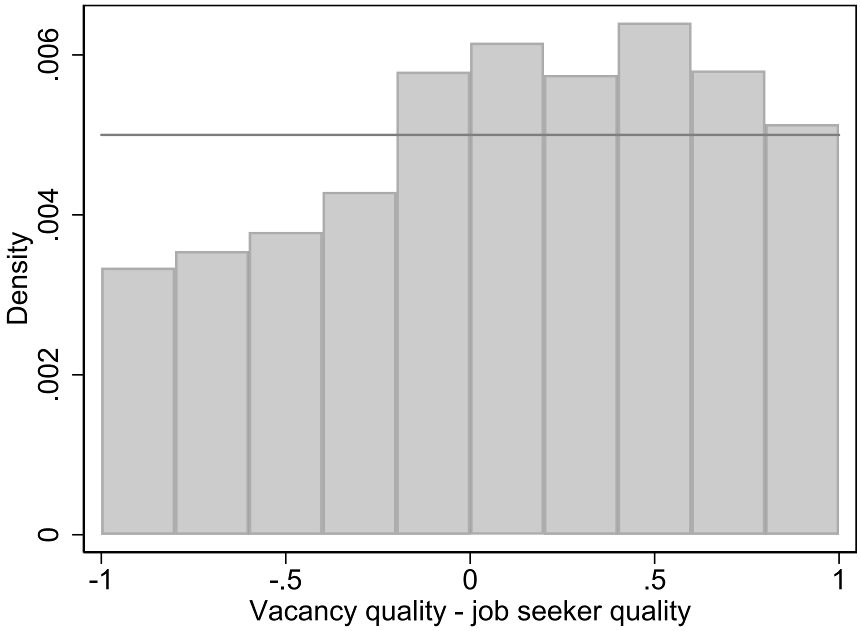
Notes: This figure shows a local polynomial regression of a dummy variable indicating whether a job seeker applies to a vacancy after he looks at it in the detail view (Figure 2b) on the difference between the quality rank of the vacancy and the job seeker. Shaded areas depict 95% confidence intervals. The quality ranking used is the out of sample ranking as described in section 5.3. The figure is based on 2,666 vacancies viewed and 467 applications made. The local polynomial uses an Epanechnikov kernel with Stata's default bandwidth.

Figure A.7: Out of sample ranking only: Distribution of applications made by vacancies



Notes: This figure shows the distribution of applications made by vacancies by difference between the quality rank of the job seeker and the vacancy. The gray line represents the density that would result from random search. The quality ranking used is the out of sample ranking as described in section 5.3. The figure is based on 19,605 applications by vacancies to job seekers. Applications in this figure are weighted as follows to ensure the density represented by the gray line would result from random search: first, applications are binned by 0.01 quality rank of the vacancy making the application, and each application is weighted by $\frac{\text{19,605}}{\text{Nb of applications made by vacancies in quality rank bin}}$. This corrects for the differential intensity with which vacancies with a different quality rank apply to job seekers *independent of job seeker quality rank*. These weighted applications are then weighted by $\frac{1}{1-|\text{Job seeker quality - vacancy quality}|}$. This weight corrects for the fact that different values of $|\text{Job seeker quality - vacancy quality}|$ have a different ex ante probability of existing. To see this, consider $|\text{Job seeker quality - vacancy quality}| = 0$. This difference exists for any vacany quality in $[0, 1]$. $|\text{Job seeker quality - vacancy quality}| = .5$, however, only exists for vacancy quality in $[0, .5]$, and for that reason the density weight put on applications with $|\text{Job seeker quality - vacancy quality}| = .5$ is twice as high as that put on applications with $|\text{Job seeker quality - vacancy quality}| = 0$.

Figure A.8: Out of sample ranking only: Distribution of applications made by job seekers



Notes: This figure shows the distribution of applications made by job seekers by difference between the quality rank of the vacancy and the job seeker. The gray line represents the density that would result from random search. The quality ranking used is the out of sample ranking as described in section 5.3. The figure is based on 12,813 applications by job seekers to vacancies. Applications in this figure are weighted as follows to ensure the density represented by the gray line would result from random search: first, applications are binned by 0.01 quality rank of the job seeker making the application, and each application is weighted by $\frac{12{,}813}{\text{Nb of applications made by job seekers in quality rank bin}}$. This corrects for the differential intensity with which job seekers with a different quality rank apply for vacancies *independent of vacancy quality rank*. These weighted applications are then weighted by $\frac{1}{1-|\text{vacancy quality - job seeker quality}|}$. This weight corrects for the fact that different values of $|\text{vacancy quality - job seeker quality}|$ have a different ex ante probability of existing. To see this, consider $|\text{vacancy quality - job seeker quality}| = 0$. This difference exists for any vacany quality in $[0, 1]$. $|\text{vacancy quality - job seeker quality}| = .5$, however, only exists for vacancy quality in $[0, .5]$, and for that reason the density weight put on applications with $|\text{vacancy quality - job seeker quality}| = .5$ is twice as high as that put on applications with $|\text{vacancy quality - job seeker quality}| = 0$.

B Revealed Preference Quality Ranking for Job Seekers and Vacancies

B.1 Estimating the ranking from the network of applications

The following illustration follows Lehmann et al. (2022, Section 5). I write the quality of a job seeker s_i as a function of the vacancies that apply to it, or more precisely, as a function of the share of applications the job seeker receives from every vacancy:

$$s_i = \frac{y_{i,1} * v_1}{\sum_{\iota \in I} y_{\iota,1}} + \frac{y_{i,2} * v_2}{\sum_{\iota \in I} y_{\iota,2}} + ... \frac{y_{i,K} * v_K}{\sum_{\iota \in I} y_{\iota,K}} \quad \forall i \in I \tag{1}$$

The quality s_i of a job seeker $i \in I$ is a function of the sum of the (unknown) qualities v_k of the vacancies $k \in K$ that apply to this job seeker $y_{i,k} = 1$, each divided by the total number of job seekers vacancy k applies to $\sum_{\iota \in I} y_{\iota,k}$. As this holds for every job seeker, we can write this as a linear system of equations:

$$\begin{pmatrix} s_i \\ \vdots \\ s_I \end{pmatrix} = \underbrace{\begin{pmatrix} y_{1,1} & \dots & y_{1,K} \\ \vdots & \ddots & \vdots \\ y_{I,1} & \dots & y_{I,K} \end{pmatrix}}_{Y} \underbrace{\begin{pmatrix} \frac{1}{\sum_{\iota \in I} y_{\iota,1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sum_{\iota \in I} y_{\iota,K}} \end{pmatrix}}_{Y_w} \begin{pmatrix} v_1 \\ \vdots \\ v_K \end{pmatrix} \tag{2}$$

Following the same idea, I write the quality of a vacancy as a function of the share of applications the vacancy receives from every job seeker:

$$v_k = \frac{x_{k,1} * s_1}{\sum_{\kappa \in K} x_{\kappa,1}} + \frac{x_{k,2} * s_2}{\sum_{\kappa \in K} x_{\kappa,2}} + ... \frac{x_{k,I} * s_I}{\sum_{\kappa \in K} x_{\kappa,I}} \quad \forall k \in K \tag{3}$$

The quality v_k of a vacancy $k \in K$ is a function of the sum of the (unknown) qualities s_i of the job seekers $i \in I$ who apply for this vacancy $x_{k,i} = 1$, each divided by the total number of vacancies job seeker i applies for $\sum_{\kappa \in I} x_{\kappa,i}$. As this holds for every vacancy, I can write this as a linear system of equations:

$$\begin{pmatrix} v_k \\ \vdots \\ v_K \end{pmatrix} = \underbrace{\begin{pmatrix} x_{1,1} & \dots & x_{1,I} \\ \vdots & \ddots & \vdots \\ x_{K,1} & \dots & x_{K,I} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \frac{1}{\sum_{\kappa \in K} x_{\kappa,1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sum_{\kappa \in K} x_{\kappa,I}} \end{pmatrix}}_{X_w} \begin{pmatrix} s_1 \\ \vdots \\ s_I \end{pmatrix} \tag{4}$$

Observe that the quality of job seekers is a (linear) function of the quality of vacancies, and vice versa. I can thus write the quality of job seekers as a recursive function of their own quality:

$$s = YY_w XX_w s$$

In other words, the quality of a job seeker (resp. vacancy) is defined recursively and depends on both the number of vacancies who apply to the job seeker, and their quality.

YY_w and XX_w are both row-stochastic matrices, meaning that their rows sum up to one. Hence, the product of the two $YY_w XX_w$ is itself row stochastic. In addition, if we restrict the sample of vacancies and job seekers to the subset of strongly connected vacancies and job seekers, the matrix $YY_w XX_w$ is an irreducible non-negative matrix. A job seeker (a vacancy) is in the strongly connected set if he (it) has been liked by at least one vacancy (job seeker) in the strongly connected set and if he (it) likes at least one vacancy (job seeker) in the strongly connected set. By the Perron-Frobenius Theorem for irreducible non-negative stochastic matrices, the eigenvector associated with $YY_w XX_w$ is unique up to scaling and corresponds to an eigenvalue of one (Minc, 1988, theorem 3.2 and theorem 4.1). The quality of job seekers s therefore corresponds to this eigenvector, and because it is unique up to scaling, the ranking of the job seekers in s is unique. Following the same reasoning, the quality of vacancies v corresponds to the eigenvector associated with $XX_w YY_w$.

The eigenvector s is the fixed point of the function $YY_w XX_w : \mathbb{R}^{|I|} \to \mathbb{R}^{|I|}$. To reach this fixed point, I start with an initial guess and repeatedly apply the function $YY_w XX_w$ until the vector s has converged.

Extending the ranking to weakly connected agents After estimating the ranking using strongly connected job seekers, I extend the set of ranked vacancies and job seekers by also calculating the rank of vacancies and job seekers only connected by an incoming application. Consider $i \in I'$ the set of job seekers that are connected to vacancies $k \in K$ by receiving an application from at least one vacancy in K.[33] I can estimate the quality rank of every job seeker in I' stacked in the vector s' using $s' = Y'Y'_w v$, where Y' represents the matrix containing all applications by vacancies in K made to job seekers in I', and Y'_w represents the corresponding weighting matrix. I proceed the same way to estimate the quality rank of any vacancy that receives at least one application from a job seeker in I. 365 of all 7,314 job seekers in the ranking belong to the set I' but not to the set I, and 568 of all 3,910 vacancies in the ranking belong to the set K' but not to the set K. In the main text of the paper, I refer to the set I' with the notation I and to the set K' with the notation K.

B.2 Interpreting the ranking with regard to quality type

I will now establish a necessary condition under which the ranking estimator described in Section B.1 measures job seeker and vacancy quality type in the framework set out in Section 2. I will then show that this condition is satisfied under directed search, likely to be satisfied under competitive search, and that testable predictions are implied by random search.

[33] This means that the set I is a subset of I', as by the strong connectedtedness requirement on I every job seeker in I receives an application from at least one vacancy in K.

The necessary condition for an estimator $\tilde{x} = \mathbb{R}^{|I|}$ to consistently estimate the rank-order of $x = \mathbb{R}^{|I|}$ is that in the limit $F(\tilde{x}_i) = G(x_i)$, where F denotes the cumulative distribution function (CDF) of $\tilde{x}$ and G denotes the CDF of x. Adapted to my framework, this means that

$$\tilde{s}_i = \sum_{k \in K} \frac{y_{i,k} * \tilde{v}_k}{\sum_{\iota \in I} y_{\iota,k}} \; \forall \, i \in I$$

and

$$\tilde{v}_k = \sum_{i \in I} \frac{x_{k,i} * \tilde{s}_i}{\sum_{\kappa \in K} x_{\kappa,i}} \; \forall \, k \in K$$

are consistent estimators for the quality rank of s and v if in the limit $F(\tilde{s}_i) = G(s_i)$ and $F(\tilde{v}_k) = G(v_k)$.

Note that under both, directed search with respect to partners' quality and competitive search

$$P(y_{i,k} = 1) = P(\hat{P}(k \text{ will match with } i) * (s_i + \epsilon_{ki}) > c)$$

and

$$P(x_{k,i} = 1) = P(\hat{P}(i \text{ will match with } k) * (v_k + \epsilon_{ik}) > c)$$

which implies that $\tilde{s}_i$ is a function of s_i and $\tilde{v}_k$ is a function of v_k. Thus, a sufficient condition for $\tilde{s}_i$ ($\tilde{v}_k$) to identify the rank-order of s_i (v_k) is that $\frac{\partial P(y_{i,k}=1)}{\partial s_i} > 0$ and $\frac{\partial P(x_{k,i}=1)}{\partial v_k} > 0$.

Identification of true ranking under directed search Under directed search with respect to partner quality, $\hat{P}$ is constant and thus $P(x_{k,i} = 1)$ ($P(y_{i,k} = 1)$) is strictly increasing in v_k (s_i). Therefore, $\tilde{s}_i$ ($\tilde{v}_k$) is strictly increasing in s_i (v_k), and hence $F(\tilde{s}_i) = G(s_i) \; \forall i \in I$ ($F(\tilde{v}_k) = G(v_k) \; \forall k \in K$).

Identification of true ranking under competitive search Under competitive search, things are more complicated as $\hat{P}(i \text{ will match with } k)$ is decreasing in the partner's quality, and thus the probability a job seeker (vacancy) applies is not necessarily strictly increasing in v_k (s_i). A sufficient condition for the rank-order of $\tilde{s}_i$ ($\tilde{v}_k$) to identify the rank-order of s_i (v_k) is that the rate of decline of $\hat{P}$ is smaller than one, i.e.,

$$\frac{\partial P(i \text{ will match with } k)}{\partial v_k} | s_i < -1$$

and

$$\frac{\partial P(i \text{ will match with } k)}{\partial s_i}|v_k < -1$$

as then it follows immediately that $\frac{\partial P(x_{k,i}=1)}{\partial v_k} > 0$ and $\frac{\partial P(y_{i,k}=1)}{\partial s_i} > 0$.[34]

Identification of true ranking under random search Under fully random search, $P(y_{i,k} = 1)$ and $P(x_{k,i} = 1)$ are constant. It follows immediately that $\tilde{s}_i$ ($\tilde{v}_k$) does not identify the rank-order of s_i (v_k). The intuition is straight-forward: Because applications are fully random, an estimator cannot learn anything from patterns of applications. As a result, under fully random search, the rank-order implied by $\tilde{s}_i$ ($\tilde{v}_k$) is random too. This implies a testable prediction: The rank-order of $\tilde{s}_i$ and $\tilde{v}_k$ should not have any predictive power about whether applications are made or accepted. This is clearly rejected by the evidence presented in Section 5.

[34]Because under competitive search vacancies and job seekers have full information and rational expectations, $\hat{P}(k \text{ will match with } i) = P(i \text{ will match with } k) = \hat{P}(i \text{ will match with } k)$.

Bibliography

Abdulkadiroğlu, Atila, Parag A Pathak, and Alvin E Roth, "The new york city high school match," *American Economic Review*, 2005, *95* (2), 364–367.

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.

Albrecht, James, Pieter A Gautier, and Susan Vroman, "Equilibrium directed search with multiple applications," *The Review of Economic Studies*, 2006, *73* (4), 869–891.

Anbarci, Nejat and Nick Feltovich, "How sensitive are bargaining outcomes to changes in disagreement payoffs?," *Experimental economics*, 2013, *16* (4), 560–596.

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.

Arnosti, Nick, Ramesh Johari, and Yash Kanoria, "Managing congestion in matching markets," *Manufacturing & Service Operations Management*, 2021, *23* (3), 620–636.

Arteaga, Felipe, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman, "Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice," *The Quarterly Journal of Economics*, 03 2022. qjac013.

Avery, Christopher, Christine Jolls, Richard A Posner, and Alvin E Roth, "The Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, 2001, *68* (3), 793–902.

Banfi, Stefano and Benjamen Villena-Roldan, "Do High-Wage Jobs Attract More Applicants? Directed Search Evidence from the Online Labor Market," *Journal of Labor Economics*, 2019, *37* (3), 715–746.

_ **and Benjamin Villena-Roldan**, "Do high-wage jobs attract more applicants? Directed search evidence from the online labor market," *Journal of Labor Economics*, 2019, *37* (3), 715–746.

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

Belot, Michele, Philipp Kircher, and Paul Muller, "Providing advice to jobseekers at low cost: An experimental study on online advice," *The Review of Economic Studies*, 2019, *86* (4), 1411–1447.

Belot, Michèle, Philipp Kircher, and Paul Muller, "How Wage Announcements Affect Job Search—A Field Experiment," *American Economic Journal: Macroeconomics*, 2022, *forthcoming.*

Berger, David W, Kyle F Herkenhoff, and Simon Mongey, "Labor market power," *American Economic Review*, 2022, *forthcoming.*

Bhole, Monica, Andrey Fradkin, and John Horton, "Information About Vacancy Competition Redirects Job Search," Technical Report, Center for Open Science 2021.

_ , _ , _ et al., "Information About Vacancy Competition Redirects Job Search," 2021.

Bó, Ernesto Dal, Frederico Finan, and Martín A Rossi, "Strengthening state capabilities: The role of financial incentives in the call to public service," *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

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

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler, "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?," Technical Report, National Bureau of Economic Research 2020.

Borjas, George J, *Immigration economics*, Harvard University Press, 2014.

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 March 3rd 2021.

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

Camerer, Colin F, *Behavioral game theory: Experiments in strategic interaction*, Princeton university press, 2011.

_ , Teck-Hua Ho, and Juin-Kuan Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, 2004, *119* (3), 861–898.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

_ , _ , 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.

Cason, Timothy N and Charles Noussair, "A market with frictions in the matching process: an experimental study," *International Economic Review*, 2007, *48* (2), 665–691.

Chan, Jason and Jing Wang, "Hiring preferences in online labor markets: Evidence of a female hiring bias," *Management Science*, 2018, *64* (7), 2973–2994.

Choo, Eugene and Aloysius Siow, "Who marries whom and why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

Coles, Peter, Alexey Kushnir, and Muriel Niederle, "Preference signaling in matching markets," *American Economic Journal: Microeconomics*, 2013, *5* (2), 99–134.

_ , John Cawley, Phillip B. Levine, Muriel Niederle, Alvin E. Roth, and John J. Siegfried, "The job market for new economists: A market design perspective," *Journal of Economic Perspectives*, 2010, *24* (4), 187–206.

Correia, Sergio, "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator," 2017. Working Paper. `http://scorreia.com/research/hdfe.pdf`.

Dhia, Aïcha Ben, Bruno Crépon, Esther Mbih, Louise Paul-Delvaux, Bertille Picard, and Vincent Pons, "Can a Website Bring Unemployment Down? Experimental Evidence from France," Technical Report, National Bureau of Economic Research 2022.

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.

Eeckhout, Jan and Philipp Kircher, "Sorting and decentralized price competition," *Econometrica*, 2010, *78* (2), 539–574.

Faberman, R Jason and Guido Menzio, "Evidence on the Relationship between Recruiting and the Starting Wage," *Labour Economics*, 2018, *50*, 67–79.

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.

Fradkin, Andrey, "Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb," 2017.

Gale, David and Lloyd S Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, 1962, *69* (1), 9–15.

Galenianos, Manolis and Philipp Kircher, "Directed search with multiple job applications," *Journal of economic theory*, 2009, *144* (2), 445–471.

Gee, Laura K, "The more you know: Information effects on job application rates in a large field experiment," *Management Science*, 2019, *65* (5), 2077–2094.

_ **, Jason Jones, and Moira Burke**, "Social networks and labor markets: How strong ties relate to job finding on Facebook's social network," *Journal of Labor Economics*, 2017, *35* (2), 485–518.

Glassdoor, "Glassdoor's Best Places to Work 2021," Technical Report 2021. https://www.oeffentlicherdienst.gv.at/fakten/bundespersonal/bedienstete/index.html, Accessed on August 9th 2021.

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, "Econometric analysis 4th edition," *International edition, New Jersey: Prentice Hall*, 2000, pp. 201–215.

_ , "The bias of the fixed effects estimator in nonlinear models," *Unpublished Manuscript, Stern School of Business, NYU*, 2002.

_ , "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.

Halaburda, Hanna, Mikołaj Jan Piskorski, and Pınar Yıldırım, "Competing by restricting choice: The case of matching platforms," *Management Science*, 2018, *64* (8), 3574–3594.

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.

He, Yinghua and Thierry Magnac, "Application Costs and Congestion in Matching Markets," Technical Report 2020.

Heckman, James J., "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process," in D. McFadden C.F. Manski, ed., *Structural analysis of discrete data with econometric applications*, Cambridge: MIT Press, 1981, chapter 4, p. 179–195.

Hensvik, Lena, Thomas Le Barbanchon, and Roland Rathelot, "Do algorithmic job recommendations improve search and matching? Evidence from a large-scale randomised field experiment in Sweden," 2020.

Hitsch, Gunter J, Ali Hortaçsu, and Dan Ariely, "Matching and sorting in online dating," *American Economic Review*, 2010, *100* (1), 130–63.

Holzer, Harry J., Lawrence F. Katz, and Alan B. Krueger, "Job Queues and Wages*," *The Quarterly Journal of Economics*, 08 1991, *106* (3), 739–768.

Hornstein, Andreas, Per Krusell, and Giovanni L Violante, "Frictional wage dispersion in search models: A quantitative assessment," *American Economic Review*, 2011, *101* (7), 2873–98.

Horton, John, "The effects of algorithmic labor market recommendations: Evidence from a field experiment," *Journal of Labor Economics*, 2017, *35* (2), 345–385.

Horton, John J, "Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution," *Management Science*, 2019, *65* (8), 3518–3540.

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.

Ibañez, Marcela and Gerhard Riener, "Sorting through affirmative action: three field experiments in Colombia," *Journal of Labor Economics*, 2018, *36* (2), 437–478.

Jäger, Simon, Benjamin Schoefer, Samuel Young, and Josef Zweimüller, "Wages and the Value of Nonemployment," *The Quarterly Journal of Economics*, 2020, *135* (4), 1905–1963.

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

Kagel, John H and Alvin E Roth, "The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment," *The Quarterly Journal of Economics*, 2000, *115* (1), 201–235.

Kanoria, Yash and Daniela Saban, "Facilitating the search for partners on matching platforms," *Management Science*, 2021, *67* (10), 5990–6029.

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, "LeaveOutTwoWay," 2019. https://github.com/rsaggio87/LeaveOutTwoWay.

_ , _ , **and** _ , "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.

Kuhn, Peter and Kailing Shen, " Gender Discrimination in Job Ads: Evidence from China *," *The Quarterly Journal of Economics*, 11 2012, *128* (1), 287–336.

_ , _ , **and Shuo Zhang**, "Gender-targeted job ads in the recruitment process: Facts from a Chinese job board," *Journal of Development Economics*, 2020, *147*, 102531.

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.

_ **, Jan Van Ours, and Josef Zweimüller**, "How changes in financial incentives affect the duration of unemployment," *The Review of Economic Studies*, 2006, *73* (4), 1009–1038.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials and rent sharing in the US labor market," Technical Report, University of Chicago WP 2021.

Lang, Kevin and Sumon Majumdar, "The pricing of job characteristics when markets do not clear: theory and policy implications," *International Economic Review*, 2004, *45* (4), 1111–1128.

Lange, Fabian, "The speed of employer learning," *Journal of Labor Economics*, 2007, *25* (1), 1–35.

Lazear, Edward P, Kathryn L Shaw, and Christopher T Stanton, "Who gets hired? The importance of competition among applicants," *Journal of Labor Economics*, 2018, *36* (S1), S133–S181.

Lee, Soohyung and Muriel Niederle, "Propose with a rose? Signaling in internet dating markets," *Experimental Economics*, 2015, *18* (4), 731–755.

Lehmann, Tobias, "A Quality Check of the Part-Time Employment Indicator in the Austrian Wage Tax Data," Technical Report, University of Lausanne 2019.

_ **, Camille Terrier, and Rafael Lalive**, "Improving Matching Efficiency in Two-Sided Markets: A Mutual Popularity Ranking Approach," mimeo, University of Lausanne June 2022.

Leibbrandt, Andreas and John A List, "Do equal employment opportunity statements backfire? Evidence from a natural field experiment on job-entry decisions," Technical Report, National Bureau of Economic Research 2018.

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 motion*, Princeton University Press, 2013.

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

Marinescu, Ioana and Ronald Wolthoff, "Opening the black box of the matching function: The power of words," *Journal of Labor Economics*, 2020, *38* (2), 535–568.

_ , 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.

McCall, John Joseph, "Economics of information and job search," *The Quarterly Journal of Economics*, 1970, pp. 113–126.

Menzel, Konrad, "Large matching markets as two-sided demand systems," *Econometrica*, 2015, *83* (3), 897–941.

Minc, Hendryk, *Nonnegative Matrices*, Wiley, 1988.

Morris, Carl N, "Parametric empirical Bayes inference: theory and applications," *Journal of the American statistical Association*, 1983, *78* (381), 47–55.

Mortensen, Dale, *Wage dispersion: why are similar workers paid differently?*, MIT press, 2003.

Mueller, Andreas I, Damian Osterwalder, Josef Zweimüller, and Andreas Kettemann, "Vacancy durations and entry wages: Evidence from linked vacancy-employer-employee data," Technical Report, National Bureau of Economic Research 2018.

Nagel, Rosemarie, "Unraveling in guessing games: An experimental study," *The American economic review*, 1995, *85* (5), 1313–1326.

Narin, Francis, Gabriel Pinski, and Helen Hofer Gee, "Structure of the biomedical literature," *Journal of the American society for Information Science*, 1976, *27* (1), 25–45.

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

Niederle, Muriel and Alvin E Roth, "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match," *Journal of political Economy*, 2003, *111* (6), 1342–1352.

Nimczik, Jan Sebastian, "Job Mobility Networks and Data-driven Labor Markets," *mimeo*, 2020.

OECD, "Gross earnings: decile ratios," 2013.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd, "The PageRank Citation Ranking: Bringing Order to the Web," 1999. Technical Report. Stanford University.

Palacios-Huerta, Ignacio and Oscar Volij, "The measurement of intellectual influence," *Econometrica*, 2004, *72* (3), 963–977.

Parcoursup, 2021. https://dossier.parcoursup.fr/Candidat/carte, Accessed on April 12 2022.

Peters, Michael, "Limits of exact equilibria for capacity constrained sellers with costly search," *Journal of Economic Theory*, 2000, *95* (2), 139–168.

_ , "Noncontractible heterogeneity in directed search," *Econometrica*, 2010, *78* (4), 1173–1200.

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

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

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

Rogerson, Richard, Robert Shimer, and Randall Wright, "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, December 2005, *43* (4), 959–988.

Rosen, Sherwin, "The economics of superstars," *The American economic review*, 1981, *71* (5), 845–858.

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

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

Roth, Alvin E, "A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom," *The American economic review*, 1991, pp. 415–440.

_ , "Deferred acceptance algorithms: History, theory, practice, and open questions," *international Journal of game Theory*, 2008, *36* (3), 537–569.

_ , "Marketplaces, markets, and market design," *American Economic Review*, 2018, *108* (7), 1609–58.

_ **and Xiaolin Xing**, "Jumping the gun: Imperfections and institutions related to the timing of market transactions," *The American Economic Review*, 1994, pp. 992–1044.

_ **and** _ , "Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists," *Journal of political Economy*, 1997, *105* (2), 284–329.

Schaffner, Florian, "Swipe right: Preferences and outcomes in online mate search," mimeo, University of Zurich November 2016.

Schmieder, Johannes F, Till Von Wachter, and Stefan Bender, "The effects of extended unemployment insurance over the business cycle: Evidence from regression discontinuity estimates over 20 years," *The Quarterly Journal of Economics*, 2012, *127* (2), 701–752.

Shi, Shouyong, "A directed search model of inequality with heterogeneous skills and skill-biased technology," *The Review of Economic Studies*, 2002, *69* (2), 467–491.

Shimer, Robert, "The assignment of workers to jobs in an economy with coordination frictions," *Journal of political Economy*, 2005, *113* (5), 996–1025.

shin Hwang, Hae, 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.

Skandalis, Daphné, "Breaking news: information about firms' hiring needs affects the direction of job search," 2018.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *The Quarterly journal of economics*, 2019, *134* (1), 1–50.

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

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.

The Atlantic, "What Tinder and Halo Have in Common," 2016. https://www.theatlantic.com/entertainment/archive/2016/01/how-tinder-matchmaking-is-like-warcraft/424350/, Accessed on April 12 2022.

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.

Wong, Linda Y, "Structural estimation of marriage models," *Journal of Labor Economics*, 2003, *21* (3), 699–727.

Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri, "Directed search and competitive search equilibrium: A guided tour," *Journal of Economic Literature*, 2021, *59* (1), 90–148.

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.