Sorting On-line and On-time

Stefano Ban
Sekyu Choi
Benjamín Villena-Roldán

Discussion Paper 19/706

Original 22 January 2019
Updated 7 November 2020

School of Economics

University of Bristol
Priory Road Complex
Bristol
BS8 1TU
United Kingdom

University of
BRISTOL
School of Economics

Sorting On-line and On-time*

Stefano Banfi
National Economic
Prosecutor of Chile

Sekyu Choi
University of Bristol

Benjamín Villena-Roldán
Diego Portales University
& MIPP & SABE

November 7, 2020

Abstract

Using proprietary data from a Chilean online job board, we compute sorting between workers and job positions types at the application stage (ex ante) and predict sorting in the flow and stock of created matches (ex post) for different type measures. We find strong evidence for positive and procyclical correlations between workers and job types. Since ex ante and ex post sorting are very similar, we conclude that sorting is largely generated at the application stage. This suggests that theoretical models of sorting with directed search are a promising path for future research.

Keywords: Online search, assortative matching, labor markets, applications.

JEL Codes: E24, E32, J24, J60

*Email: stefano.banfi@gmail.com , sekyu.choi@bristol.ac.uk and benjamin@benjaminvillena.com. We thank Jan Eeckhout, Shouyong Shi, Yongsung Chang, Chao He, Ilse Lindenlaub, Chinhui Juhn, and Germán Cubas, the editor Florin O. Bilbiie and two anonymous referees for comments and discussions. We acknowledge the support of the SABE (Sistema de Análisis de Bolsas de Empleo) project team (Juan Velásquez, Rocío Ruiz and Felipe Vera) for online job board data curatory and financial support of SENCE and OTIC SOFOFA. We are also in debt to www.trabajando.com for data provision, and especially to Ramón Rodríguez, Jorge Vergara, and Ignacio Brunner for informative conversations and support. We also thank colleagues at University of Houston, 2018 SOLE Meeting, Cardiff University, University of Manchester, Diego Portales University, 2016 Midwest Macro Meetings, 2016 LACEA-LAMES, 2016 Workshop in Macroeconomic, Search & Matching at the University of Chile, 2016 Shanghai Workshop in Macroeconomics, and Alberto Hurtado University. Villena-Roldán thanks for financial support the FONDECYT grant projects 1151479 and 1191888, Proyecto CONICYT PIA SOC 1402, and the Millenium Institute for Research in Market Imperfections and Public Policy, ICM IS130002, Ministerio de Economía, Fomento y Turismo de Chile. All errors are ours.

1 Introduction

Sorting in labor markets has been extensively studied in economics: whether high wage workers work (or not) for high wage firms is relevant for questions of efficiency and inequality. However, little is known about the way these allocations are generated: do workers and firms meet randomly, with sorting arising ex post by selective hiring and separations? is the ex ante self selection of workers into different positions at the application stage relevant for sorting?

Using a decade of daily data from the Chilean online job board `www.trabajando.com` (along with representative household survey data) from 2010 to 2019, we study the importance of the directed search component in determining labor market sorting. Using information of job seekers and job ads linked through applications on the job board, we can estimate several notions of worker and job types as well as their correlations, something we label *ex ante sorting*. Using the longitudinal dimension of the household survey and a simple imputation procedure to match observed allocations of new hires, we predict the amount of ex ante sorting that is likely to prevail ex post .

As a first contribution, we find a large and robust correlation between worker and job types: around 0.6 using our preferred measure. When comparing ex ante and ex post correlations, we find that the two numbers are remarkably similar, leading us to conclude that most of labor market sorting is due to application decisions of workers (the ex ante sorting). We also provide evidence that sorting is significantly procyclical, thus economic downturns have a sullying (rather than cleansing) effect on matches.[1]

We provide a second contribution in terms of measurement of types. Since the seminal work of Abowd, Kramarz, and Margolis (1999) (AKM henceforth) economists have looked for ways to identify and estimate productivity types from realized wages. The key idea is that in expectation, high productivity workers earn high wages regardless of the specific match they are in and high productivity positions expect to pay systematically higher wages independent of the worker. Instead of relying on a particular theoretical/statistical/structural model to identify types,[2] we use the fact that our data contains these wage expectations directly: employers declare wages they expect to pay before hiring anyone in particular, and workers declare a wage they expect to receive before applying to any job. These expectations are an ideal measure of productivity types, following the insights of Abowd, Kramarz, and Margolis (1999) and Borovickova and Shimer (2020). On the other hand, we also show how we can approximately compute with our data the AKM fixed effects, one of the most used type measurements in the literature. We show that the sorting implied by these fixed effects is substantially lower than the one obtained from our preferred types, as expected in theory and in practice: see Eeckhout and Kircher (2011), Hagedorn, Law, and Manovskii (2017),

[1] Our evidence is related to the findings of Crane, Hyatt, and Murray (2018).

[2] We see our results as complementary to the structural frameworks in Lise and Robin (2017) and Bagger and Lentz (2018).

Lopes de Melo (2018), Bonhomme, Lamadon, and Manresa (2019), among others.

Our third and final contribution is a proposal to switch the focus of sorting from worker-firm to worker-job allocations, which are conceptually closer to the idea in the seminal paper by Becker (1973). While worker-firm assignment is interesting on its own regard because of its links to raising inequality in different contexts,[3] focusing on worker-job matches can provide a richer picture of the labor market in terms of skills mismatch, occupational mobility, and overall efficiency. This focus is closely related to a rapidly growing strand of research on job search in frictional environments where individuals match with jobs or occupations rather than with firms per se, as in Baley, Figueiredo, and Ulbricht (2020), Guvenen, Kuruscu, Tanaka, and Wiczer (2020), Taber and Vejlin (2020), Dvorkin and Monge-Naranjo (2019), Cubas and Silos (2020) among others. We provide further evidence of the relevance of worker-job assignments by estimating job types, both with and without controlling for the posting firm's fixed effects. Sorting patterns remain independent of the choice of fixed effects, which shows that jobs alone are a meaningful margin to study.

This paper is related to the growing literature using online job-posting and retail websites in order to study different aspects of frictional markets. This literature is already vast, so we refer to papers that provide evidence on how posted wages or prices affect behavior of applicants or buyers. Kudlyak, Lkhagvasuren, and Sysuyev (2013) study how job seekers direct their applications over the span of a job search. They find some evidence on positive sorting of job seekers to job postings based on education and how this sorting worsens the longer the job seeker spends looking for a job (the individual starts applying for worse matches). Marinescu and Wolthoff (2020) use the data from the `Careerbuilder.com` posting website to study the relationship between job titles and wages posted on job advertisements. Lewis (2011) and Banfi and Villena-Roldan (2019) show internet seekers significantly react to posted information for car and labor markets, respectively. Jolivet and Turon (2014) and Jolivet, Jullien, and Postel-Vinay (2016) use information from a major French e-commerce platform, `www.PriceMinister.com`, to study the effects of search costs and reputational issues (respectively) in product markets.

In sum, we offer a new approach to study sorting in labor markets. While there are some limitations to our exercise (mainly, we do not observe the final allocations of workers to jobs), our setting offers a transparent and straightforward way to measure sorting. The fact that sorting at the application stage is remarkably similar to ex post sorting, suggests that theoretical settings, along the lines of Shimer (2005) and Abowd, Kramarz, Pérez-Duarte, and Schmutte (2018), in which directed search is a key ingredient, are a desirable path for future research in this area.

The next section describes thoroughly our main data set which is needed to explain how we intend to rationalize our results given the analytical framework we present in section 3. In sections 4 and 5 we present our results for the cross section and cyclical conditions, respectively. Section 6 presents our description of the imputation procedure we use to predict ex post sorting and presents

[3] See for example Card, Heining, and Kline (2013) and Song, Price, Guvenen, Bloom, and Von Wachter (2019).

results from that exercise. The last section concludes.

2 The Data

We use data from `www.trabajando.com` (henceforth TC or the website). Our data covers daily job postings and job seeker activity in the Chilean labor market, between January 1st, 2010 and December 31st, 2019. We observe entire histories of applications (dates and identification numbers of jobs applied) from job seekers and dates of ad postings for employers. After some cleaning and applying restrictions, our dataset contains 50,467,629 applications linking 1,709,887 applicants and 1,165,885 job ads. We also use information from the *Encuesta Nacional de Empleo* (National Employment Survey in Spanish, ENE henceforth)[4] for some representativeness analysis and the ex post exercise later in the paper.

A novel feature of the dataset, is that the website asks job seekers to record their expected salary (a monthly amount, net of taxes), which they can then choose to show or hide from prospective employers. Recruiters are also asked to record the expected wage for the advertised job (also a net monthly number), and are also given the same choice of whether to make this information visible or not to applicants. Employers are required to fill this information to post a vacancy, but there is no such requirement for job seekers. While there is no unique rationalization as to why individuals and firms may choose to hide this information, Banfi and Villena-Roldan (2019) provide some suggestive evidence: hidden/implicit job ads usually post higher job requirements (education and/or experience), expect to pay higher wages and receive more applications (on average), all evidence of some form of strategy by firms to attract workers non-randomly.

Admittedly, our analysis relies on self-reported wage information that applicants or employers may choose to keep private, which may rise doubts about its accuracy. However, Banfi and Villena-Roldan (2019) find that the informational content of wages that are kept private (implicit wages) is high, given that they can be predicted quite accurately using observable characteristics and an estimated model with a sample of explicit wages only.

For each posting, besides offered wage (which can or cannot be visible by applicants) we observe its required level of experience (in years), required education (required college major, if applicable), indicators of required skills ("specific", "computing knowledge" and/or "other") how many positions must be filled, an occupational code (not necessarily aligned with international standardized codes), geographic information and some limited information on the firm offering the job: its size (brackets with number of employees) and an industry code.

For job seekers we observe date of birth, gender, nationality, place of residency ("comuna" and "región", akin to county and US state, respectively), marital status, years of experience and educational attainment. We have codes for a referential occupational area (also not aligned with

[4]See the Appendix for further details of this survey.

standardized taxonomies) of the current/last job of the individual as well as information on the monthly salary of that job and both its starting and ending dates (if unemployed).

We restrict our sample to consider individuals working under full-time contracts and those unemployed. We further restrict our sample to individuals aged 25 to 55, to avoid education and retirement decisions. We discard individuals reporting desired monthly net wages above 5 million pesos.[5] This amounts to approximately 8,520 USD per month. We also discard individuals who desire net wages below 150,000 CLP (around 256 USD) a month, somewhat below the minimum wage for full-time workers in Chile in 2010 (165,000 CLP). Consequently, we also restrict job postings to those offering monthly salaries in those bounds. Additionally, we restrict our sample to active individuals and job postings: we consider workers who make at least one application and job postings which receive at least one application during the span of our dataset.

As with many self-reported data sources, there are measurement issues. Individuals may misrepresent information in their CV's to look more appealing for employers, but this may be a dangerous strategy for job seekers in a concrete hiring situation if their credentials are likely to be verified. One caveat with our analysis is that the worker information is a snapshot of their last CV as of Dec 31st, 2019. For job seekers who have never updated their CV, this is not an issue. However, if job seekers change their CV in between applications, we may correlate information of a newer CV with information of the job ads they applied to. Since we focus on job seekers between 25 and 55 years old, most of them have already finished their education and are still sufficiently far from retirement so that we should not expect dramatic CV changes over the time we observe their applications, which is ten years at most. In addition, this measurement issue probably decreases the level of assortative matching we estimate, since it most likely makes job requirements and job seeker characteristics more *dissimilar* than what they actually are.

Table 1 shows some descriptive statistics for individuals in our sample. From the table we observe that males are a majority of all job seekers, especially among the unemployed. The sample is a young one with average age 33.69. Job seekers above 50 are scarce. Given the age group we consider, most individuals in the sample have some working experience, with mean number of years of experience hovering around eight. Job seekers in our sample are more educated than the average in Chile, with 34.7% of them claiming a college degree or more, compared to around 22% in a similar age group and time frame (2010-2019) according to the ENE.

We can also observe that most job seekers have studies related to management (around 17.3%) and technology (28.2%), but a significant fraction (around 26.7%) does not declare any specialization. In terms of salaries, average expected wages are (in Chilean Pesos) $1,040,000$ and $629,000$ for employed and unemployed seekers, respectively.

[5]A customary characteristic of the Chilean labor market is that wages are generally expressed in a monthly rate net of taxes, mandatory contributions to health insurance (7% of monthly wage), contributions to the fully-funded private pension system (10%), to disability insurance (1.2%), and mandatory contributions to unemployment accounts (0.6%).

Table 1: Characteristics of Job Seekers

	Unemployed	Employed	Total
Gender (%)			
Females	49.15	41.35	45.55
Males	50.84	58.65	54.44
Age (%)			
25 - 29	44.02	40.48	42.38
30 - 34	21.01	23.81	22.30
35 - 39	13.62	15.35	14.41
40 - 44	9.52	9.90	9.70
45 - 50	6.86	6.25	6.58
> 50	4.98	4.20	4.63
Applicants Age			
Age (Mean/(S.D.))	33.69	33.81	33.74
	(7.80)	(7.41)	(7.62)
Experience (%)			
0 - 3	33.31	23.43	28.73
4 - 7	25.38	29.11	27.08
8 - 12	21.18	24.69	22.78
13 - 20	14.62	17.16	15.78
> 21	5.31	5.55	5.42
Applicants Experience			
Years of Experience (Mean/(S.D.))	7.66	8.62	8.11
	(6.88)	(6.52)	(6.73)
Education level (%)			
Primary (1-8 years)	0.36	0.11	0.24
High School	39.63	27.97	34.25
Tech. Tertiary Educ.	24.02	23.06	23.57
College	27.07	44.25	35.00
Post-Graduate	8.83	4.57	6.86
Education area (%)			
Commerce and Management	16.64	18.09	17.30
Agropecuary	0.81	1.11	0.95
Art and Architecture	1.77	2.42	2.07
Natural Sciences	1.37	1.19	1.29
Social Sciences	4.18	5.69	4.88
Law	1.71	2.02	1.85
Education	4.11	3.86	3.99
Humanities	0.99	1.08	1.03
Health	6.62	5.72	6.21
Technology	24.87	32.19	28.24
Non-declared	33.90	18.35	26.73
Other	3.03	8.27	5.45
Expected wages, thousands of CLP (%)			
150 - 300	18.09	6.16	12.58
300 - 600	47.80	28.68	38.97
600 - 1,000	23.17	31.72	27.11
1,000 - 1,500	6.10	15.89	10.61
1,500 - 2,500	3.03	12.54	7.42
> 2,500	1.07	4.76	2.77
Not declared	0.07	0.24	0.16
Wages			
Expected wages (Mean/(SD))	629.1	1040.4	819.0
	(483.5)	(760.4)	(659.4)
Observations	919,436	788,704	1,709,887

Table 2: Characteristics of Job Ads

	Hidden wage	Explicit wage	Total
Required Experience (%)			
0	15.85	22.87	17.00
1	34.71	44.35	36.29
2 - 3	36.62	26.21	34.91
4 - 7	10.91	5.23	9.98
8 - 12	1.56	0.58	1.40
$>$ 12	0.19	0.11	0.18
Required Experience			
Experience (Mean/(SD))	2.09	1.39	1.98
	(81.71)	(1.61)	(74.76)
Required educ. level (%)			
Primary (1-8 years)	1.36	5.10	1.97
High School	32.97	51.62	36.02
Tech. Tertiary Educ.	29.08	27.17	28.76
College	35.98	15.91	32.69
Post-Graduate	0.62	0.20	0.55
Major study area (%)			
Commerce and Management	13.03	12.60	12.96
Agropecuary	0.53	0.71	0.56
Art and Architecture	0.35	0.26	0.33
Natural Sciences	0.59	0.91	0.64
Social Sciences	3.41	2.17	3.20
Law	0.66	0.44	0.63
Education	0.90	0.94	0.91
Humanities	0.20	0.29	0.22
Health	1.76	1.48	1.72
Technology	31.11	18.87	29.10
Non-declared	47.30	61.17	49.56
Other	0.16	0.17	0.16
Offered wages, thousands of CLP (%)			
150 - 300	18.67	29.47	20.43
300 - 600	36.22	49.65	38.42
600 - 1,000	24.10	12.71	22.23
1,000 - 1,500	11.10	3.90	9.92
1,500 - 2,500	6.74	1.75	5.92
$>$ 2,500	1.82	0.63	1.62
Offered Wages			
Wages (Mean/(SD))	754.6	494.0	712.0
	(606.8)	(404.9)	(586.6)
Observations	975,224	190,661	1,165,885

Table 2 shows sample statistics for job postings. We separate our sample between postings with hidden or implicit wages (which do not post information on salaries) and with explicit wages. From a total of $1,165,885$ active job postings in our sample period, only $190,661$ (16.4%) have an explicit wage.

Hidden wage postings are characterized for requiring higher levels of experience and higher levels of education. They also tend to concentrate more on technology related occupations: 31.1% of ads with hidden wages are related to technology, versus 18.8% of job postings with explicit wages. Job postings in our sample receive a mean of 22.2 applications, with a significant difference in the number received by implicit wage postings (23.2) versus those received by explicit ones (15.1). The average of posted wages is near $712,000$ CLP, being the explicit-wage subsample considerably lower. In any case, jobs posted in the website offer wages that are substantially higher than those measured by household surveys in Chile. In Appendix A1 we show further statistics for both applicants and job ads.

To assess the representativeness of our data when compared to the Chilean Economy in a more formal way, in table 3 we compare quintiles of salaries in the Chilean economy computed from the ENE.[6] We use the same sample period (2010 to 2019) and apply the same age and salary (minimum/maximum) restrictions as in our TC data. We also focus only on full-time workers not on short-term/temporary contracts.

Table 3: Quintiles of Monthly Salaries (after tax, thousand of CLP)

	p20	p40	p60	p80
All workers	250	339	452	703
New workers	225	298	379	538
New workers, NOT in Agriculture nor Construction	240	300	400	596
New college workers, NOT in Agriculture nor Construction	304	460	639	1,000
TC workers	360	500	700	1,100
TC ads	300	420	600	1,000

In the table we compute monthly, nominal (after tax) wage quintiles for workers in the Chilean Economy for several cases: all workers, new workers with job tenure of less than twelve months (which represent a sample who has recently performed a job search), new workers who do not work in agriculture nor construction sectors (poorly represented in TC, as seen in 2) and new workers, not working in those sectors, who have a college degree (or a higher attainment).

As discussed above, the sample of workers and job ads in TC is described by monthly wages which are significantly higher than the overall distribution observed in the Chilean data. This is explained both because of the educational attainment of individuals using the website and the occupational/industry composition of jobs (mostly management and technology) which pay higher

[6]See http://webanterior.ine.cl/estadisticas/laborales/ene and http://webanterior.ine.cl/estadisticas/ingresos-y-gastos/esi.

wages in general. Note that even when we consider only college graduates (and individuals with higher degrees) in the table, the wage distribution is still lower than the one in TC.

Besides this level differences between TC wages and the rest of the Chilean economy, wages in the website still present a significant spread, as seen by the different quintiles in the table and relatively in line with the dispersion observed in the Chilean economy. What is more important, there seem to be a rough alignment between wage expectations by workers and job positions in the website.

3 Analytical framework

In this section we study how the evidence in our paper relates to theories of sorting between heterogeneous agents. As discussed in the introduction, we are focused on *pre-match* allocations, i.e., the sorting between job applicants and job positions during the application stage. This is a new dimension of the data that has not been explored and that can provide guidance on the models frequently used to study sorting. For example, Shimer (2005) suggests a test of log-supermodularity using application decisions by unemployed workers, our exact empirical setup.

A useful benchmark: Borovickova and Shimer (2020)

In our data we have data on two key variables: (i) the wage employers expect to pay to a hired worker for the advertised position *before knowing any specific applicant*, and (ii) the wage job seekers expect to receive when hired by an employer *before applying to any specific job ad.* At face value, one could rationalize in different ways how wage expectations are related to heterogeneity and whether any empirical moment constructed using these, may be indicative of sorting in the labor market.

However, Borovickova and Shimer (2020) suggest measures of types which match *exactly* the wage expectations from our online job board data. In particular, they suggest to use as the worker's type "the expected log wage she receives in an employment relationship conditional on taking the job" and for the firm type, "the expected log wage that it pays to an employee conditional on hiring the worker". They also suggest that the degree of assortative matching can be inferred by the correlation of these expectations. The main goal for the rest of this section is to answer whether these variables have any sensible interpretation (in terms of any established sorting theory) and discuss how they are related to more applied approaches in the empirical literature.

Expected wages as types: in theory.

Search and matching activity inside an online job board can be rationalized naturally using the competitive search framework, which follows Moen (1997): firms post vacancies independently of each other, which can lead to coordination frictions of potential applicants. These online job advertisements, along the wealth of information they provide about working conditions, may be taken as particular *sub-markets* designed to attract appropriate applicants. At any point in time (after positions are posted), job seekers observe all relevant job ads and decide where to apply, choosing job ads which maximize the expected value provided to them.

There are theoretical reasons to believe that posted wages are a credible measurement for realized wages. In a competitive auction setting, akin to competitive search, Kim and Kircher (2015) show that auctioneers' posted prices (wages) truthfully reveal their types although their announcements are essentially cheap talk. On the other hand, there is empirical evidence that high-wage ads attract more applicants (all else equal), which clearly indicates competitive search behavior as noted by Banfi and Villena-Roldan (2019); Marinescu and Wolthoff (2020). A final remark on the reliability of expected wages as types: posted and expected wages are consequential for agents in the search process. The search engine of the website filters job ads by wage brackets even if employers choose to conceal wages from applicants. On the other side of the market, employers may observe the jobseeker's wage expectation if visible before contact them.

In this regard, the framework in Shimer (2005), where competitive search is paired with heterogeneity of the two sides of the market (workers and jobs), is useful to understand whether the notion of expected wages as types has any grounding. The model is composed of workers and firms who are risk neutral and heterogeneous. Production takes place when a single worker and a single job match. The timing of the market is standard: firms post type-contingent wage offers. Then, workers observe all offers and apply for one job. Firms who receive at least one application, hire a worker, produce and obtain profits given production minus the promised wage. Firms and workers who do not match remain iddle/unemployed. Shimer (2005) provides proofs of existence and uniqueness of a competitive search equilibrium. Moreover, under certain assumptions for the production function,[7] the competitive equilibrium which decentralizes the social optimum is described by one in which firms hire the most productive type.

Naturally, one can ask whether expected wages as defined in this model are related to firm and worker types. For jobs, it is straightforward to prove that the expected wage a firm is expecting to pay is increasing in job productivity: this is just a corollary that wages in Shimer (2005) increase with firm productivity.[8] However, the same cannot be said about expected wages for workers without imposing strong restrictions on the exogenous distribution of worker and job types.

Expected wages as types: in practice.

Although the framework in Shimer (2005) contains some useful predictions and testable implications for expected wages as proxy for types, it cannot fully account for several phenomena that define labor markets (something we observe in our data and document in the next section) and that potentially affect the theoretical results: First, worker's search strategies may rely on multiple

[7]Monotonicity and Supermodularity.

[8]See Proposition 4 in that paper.

simultaneous applications whereas the setting described above admits only one application per worker. Second, there is non-significant amount of workers who search while already employed who may compete directly with unemployed workers of the same type. Finally, the setting is static which precludes the study of business cycle effects. To the best of our knowledge, there is no theoretical nor quantitative model in the literature with all these ingredients.

Following the insight from Borovickova and Shimer (2020), we use log expected wages directly reported on both sides of the market, and refer to them thereafter as EW types:

$$EW_a = \log E[\omega(w,a)|a] \qquad \text{and} \qquad EW_w = \log E[\omega(w,a)|w].$$

On the other hand, the applied literature usually relies on a statistical framework as the one proposed by Abowd, Kramarz, and Margolis (1999), where types are identified as worker and firm fixed effects estimated using wage information from matched employer-employee data. In simple terms, we can relate to that type of statistical model (usually referred as the AKM setup), by assuming

$$\log W(w,a) = \tau_w + \tau_a + \epsilon_{wa} \tag{1}$$

where $W(w,a)$ is the wage of worker w at job position a. According to the AKM setup, this log wage depends on the type of the worker τ_w, the type of the position τ_a and is subject to some error ϵ_{wa}. Direct application of equation (1) to our setting is unfeasible, given that we don't observe realized wages but only application decisions. However, assuming that agents have unbiased expectations with respect to their potential partners, we can use (1) to compute expected wages in this setting. For a particular worker (w) we have that

$$E[\log W(w,a)|\tau_w] = \tau_w + E[\tau_a|\tau_w]$$

where the left hand side of the equation is the expected log wage of the worker, conditional on being hired (which applicants directly report in our data) while $E[\tau_a|\tau_w]$ is the expected value of the job type, conditional on the market faced by worker w. This latter term can be also thought as the average type of position that the worker has applied to, which in the *pre-match* stage is an unknown. We approximate the expected log wage $E[\log W(w,a)|\tau_w]$ by the log of the expected wage $\omega_w = \log E[W(w,a)|\tau_w]$, and project the term $E[\tau_a|\tau_w]$ onto a set of average characteristics $\overline{X}_{a|w}$ of the ads to which worker w applies to in order to obtain the following measurement equation

$$\omega_w \approx E[\log W(w,a)|\tau_w] = \tau_a + E[\tau_a|\tau_w] \approx \tau_a + \overline{X}_{a|w}\gamma_a$$

Estimating this equation by OLS, we obtain

$$\widehat{\tau}_a \approx \omega_a - \overline{X}_{a|w}\widehat{\gamma}_a. \tag{2}$$

We can apply the same logic to the other side of the market. We use a set of average worker characteristics that apply to ad a, $\overline{X}_{w|a}$ in order to obtain an estimate for the worker AKM-like type

$$\widehat{\tau}_w \approx \omega_w - \overline{X}_{w|a}\widehat{\gamma}_w. \tag{3}$$

In sum, the job type is the residual of a regression between the log expected salary to be paid at the position as a dependent variable and average characteristics of workers applying to the position as independent variables. The set $\overline{X}_{w|a}$ includes gender, nationality, marital status, region of residence, educational attainment, educational area, age and self reported experience of job seekers, as well as indicators for whether individual applicants have explicit wage expectations.

Analogously, we estimate the worker type by obtaining the residuals of a regression between each worker's expected wage and a set of (average) controls $\overline{X}_{a|w}$ of job/firm characteristics: indicators for explicit wage posting, firm size, firm's geographic region, educational requirements, contract types (long term, fixed term, etc), time arrangements (full- or part-time), average number of vacancies, required experience, and number of employed and unemployed applicants received by the job ad.

Our procedure has two advantages: first, given our data, it is straightforward to compute using simple linear regressions. Second, it can be estimated independently for each side of the market.

Since we are focusing on the concept of a worker-position match instead of the more traditional worker-firm match, we must consider how much controling for firm's fixed effects alters our results. If position productivity is firm-dependent, or if the firm is an attribute of the position in itself, we should not include firm fixed effects in our estimates above. Under this assumption, (2) and (3) correctly approximate AKM types. On the contrary, if we think that firm fixed effects are additively separable from the position type τ_a, then we should include a term τ_f in equation (1). We call this second measure AKM-f types. We remain agnostic and produce both estimates in the next section.

Expected wages as types: discussion.

How do the results presented in the next section relate to results in papers with different model environments? Two leading examples of quantitative frameworks where sorting is studied using structural models are Lise and Robin (2017) and Bagger and Lentz (2018).

The main distinction between our results and these papers is the mechanisms leading to sorting. In our setup, we study labor market sorting in a *pre-match* stage. Both in Lise and Robin (2017) and Bagger and Lentz (2018), sorting is driven both by initial sorting (mutual agreement to form a match) plus post-match forces: mismatched pairs endogenously separate when aggregate conditions change, as is the case in Lise and Robin (2017), or workers climb the job ladder to find better matches, a mechanism present in both papers. We see our evidence as complementary to these studies, in that *pre-match* sorting can be thought of as disciplining evidence for *acceptance* regions in those papers.

A further distinction exists in terms of the amount of information used to determine types empirically: Lise and Robin (2017) use aggregate information on labor productivity and value added to identify worker and firm types, without controlling for other observables. Bagger and Lentz (2018) use Danish administrative data which contains limited demographic information, education being the most salient one. As a point of comparison, French data in Abowd, Kramarz, and Margolis (1999) contains limited demographic information (only 10% of the sample has information on individuals' education). Our dataset contains a very rich set of covariates for individuals, job positions and firms, which we use in our main analysis. Thus, the comparability of our estimates with the rest of the literature needs to be taken with these discrepancies of data granularity in mind.

4 Sorting on-line

In this section, we report cross-sectional sorting between workers (w) and job ads (a) using the measures described above: expected wages (EW) and the two AKM like measures (with and without firm fixed effects).

In Figure 1 we show the joint frequency of observed applications given 30 quantiles of worker and ad types, that is, their empirical bivariate distribution. A darker area reflects high percentage of applications made by workers in a given quantile (projected onto the vertical axis) and received by job ads in a another quantile (projected onto the horizontal axis). The first subfigure shows a large alignment of EW types generated by applications. Lighter gaps can be seen due to the fact that most wages posted are round numbers. Some bunching of applications is observed for low worker and job ad EW types. As we move further away from the 45 degree line, more lighter tones appears, reflecting that applications combining low and high types are nearly 10 times more infrequent than the most abundant cases, depicted in black. Overall, the correlation of EW types is around 0.6, suggesting strong positive assortative matching. The number is highly significant[9].

Both AKM and AKM-f types show smoother joint densities exhibiting some bunching on the extremes of the 45 degree line. As we move away from the latter line, we observe progressively lighter tones that are consistent with positive sorting. However, the smother transition also implies

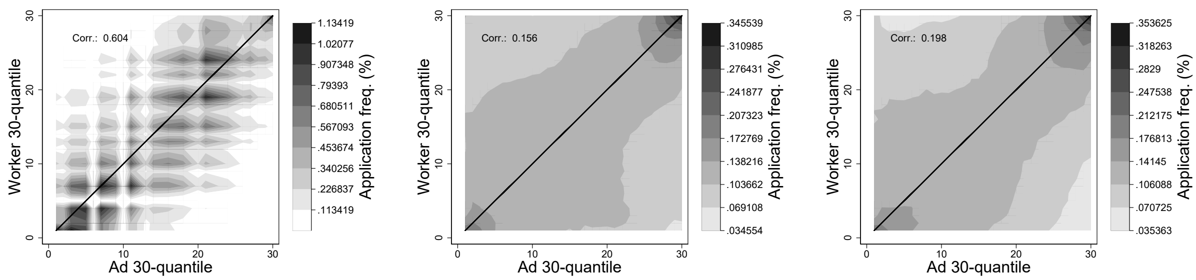
[9] We compute White (1980) robust standard errors in our regressions. Following the method of Cameron, Gelbach, and Miller (2011), we also compute standard errors using a two-way cluster (at the job ad and applicant levels). These standard errors are between 17% and 31% larger than White's, and therefore do not really affect p-values (given that our sample consists of nearly 40 million observations). Thus, we focus on economic rather than statistical significance.

a weaker assortative pattern than the one found for EW types. Indeed, the sorting correlations of AKM and AKM-f types are lower than that of EW types (0.156 and 0.198, respectively). Whether we define firm identity as a job's attribute or not (AKM vs AKM-f) seems to be hardly empirically relevant for the degree of sorting. This is evidence that sorting at the worker-job level represents a relevant margin, given that we find sorting even after controlling for the effect of the firm.

We notice a sizable difference in the strength of sorting between EW and AKM types. This is a natural consequence of controlling for average observable characteristics of agents that are linked through an application on the other side of the market. Taking away the variation of attributes that guide application decisions is likely to decrease the alignment of types. Moreover, Borovickova and Shimer (2020) show that the correlation of AKM types is able to recover the true strength of sorting only if the data generating process is truly AKM. In contrast, these authors find that the correlation of EW types captures the true sorting under a wider set of circumstances, including when AKM is the true underlying model. Therefore, the correlation gap between the two measures suggests that the AKM model may misrepresent the true sorting in the data.

Setting aside the discussion about whether AKM types allow to appropriately measure sorting, their use is a standard way to empirically quantify this feature of labor markets. Hence, our version of AKM types can be compared to other studies in the literature as a basic assessment of how our results on EW types would look like in other settings. The correlation of AKM types obtained here are on the upper bound of estimates found in the literature. For example, Lopes de Melo (2018) (see table 1 of the paper) summarizes results for estimates using data from the US, France, Germany, Italy, Denmark and Brazil: all estimates are either close to zero or even negative, as is the case for France. Bonhomme, Holzheu, Lamadon, Manresa, Mogstad, and Setzler (2020) in their Table D1 also report estimates ranging from -0.24 to 0.231. One of the issues in this literature is that low (job to job) mobility of workers biases types' correlations to zero, as explained in Andrews, Bradley, Stott, and Upward (2008), Lamadon, Mogstad, and Setzler (2019) and many others. To the extent that individuals manifest their interest in potential jobs in our data, our AKM results probably attenuate the low mobility problem. Therefore, obtaining an estimate in the upper part of the range in the literature should not be a surprise. However, a strict comparison between our setting and matched employer-employee data used in the literature is difficult because job seekers are already self-selected workers with relatively high likelihood of moving.

Robustness. We study sorting in different subsamples. Our results for the strength of the types correlation are robust and hover around 0.56 - 0.65. First, we divide our sample between applications made by unemployed and employed job seekers. By doing so, the correlation of the two subsamples goes down compared to the benchmark of 0.6 obtained for the whole sample. Results in Table A4 are expected: employed workers tend to be of higher types and apply to higher type job ads, and the opposite occurs for the unemployed. Some degree of positive correlation is lost



(a) EW (b) AKM (c) AKM-f

Figure 1: Frequencies of applications, by worker and ad percentiles of EW, AKM and AKM-f types. Sample consists of 39,480,855 applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`

when conditioning on employment status, as this variable is highly positively correlated with both types.

In addition, Table A4 also shows estimates for a subsample of applications that were made after the last CV update of the worker (set of most recent applications for each worker). As expected, the correlation between types is larger in this group (0.658) because of less measurement error (albeit, in sacrifice of sample size): our full sample considers old applications between job ads and workers for whom we might observe the *innacurate* information. The same result is observed with AKM and AKM-f types. Nevertheless, there are caveats about using only applications made after the online CV is updated as a baseline result. Newly CV updates may generate selection bias in terms of composition of job seekers, because some demographic groups may update CVs more frequently. Moreover, this sample contains relatively fewer older applications, which may be problematic to assess the cyclical properties of sorting.

Table A5 contains correlation estimates for samples with varying degrees of visibility for job seekers. As mentioned in Section 2, workers and employers may choose to hide wages to the other side of the market. Therefore, it is interesting to investigate whether releasing information spurs or dampen assortative applications. The upper panels show that the EW correlation mildly decreases when splitting applications between ads with and without explicit wages. As explicit wage posting is consistently related to lower wages and requirements (Banfi and Villena-Roldan, 2019), we should expect lower correlation when conditioning on such a variable. When we split applications by the visibility of worker wage expectations, we observe that correlations remain at roughly the same level as in the baseline scenario. Therefore, we conclude that either applicants have correct expectations or private information of wages regardless the fact they are explicitly posted online. As applications are formal requests for being considered to applied positions, we conclude that directed search is a good theoretical setting to make sense of the evidence.

5 Sorting on-time

In this section, we study how the correlation between characteristics of workers and jobs vary with aggregate business cycle conditions. To assess this, we run the following standard specification using all applications (pairs of workers w and job ads a) at time t:

$$y_{w,t} = \alpha + \rho \, y_{a,t} + \delta \, y_{a,t} z_t + \phi z_t + X_t \lambda + \epsilon_{a,w,t} \tag{4}$$

where $y_{w,t} = \{\omega_{w,t}, \tau_{w,t}\}$ is the statistic of interest of the worker at time t, $y_{a,t} = \{\omega_{a,t}, \tau_{a,t}\}$ is the statistic of the job posting and z_t is a variable capturing aggregate economic conditions at monthly frequency (time t). These variables are standardized, so their mean is zero and their standard deviation is one. The specification also includes monthly seasonal dummies and a linear trend to capture long run movements in both types and the cyclical variable.

In this specification, the estimate for ρ is the average correlation between $y_{w,t}$ and $y_{a,t}$ when the cyclical variable is at its sample mean value $\bar{z}$, and matches the notion of sorting in the previous section. The coefficient δ, in turn, measures how assortative matching is affected when the cyclical variable z_t increases in one sample standard deviation. In what follows, we use the average unemployment rate for the Chilean economy[10] as an aggregate indicator. During our sample period (2010-2019), the Chilean economy experienced an economic recovery from the 2009 global financial crisis, a period of low and stable unemployment roughly between 2014 and 2016 and an uptick in joblessness by the end of the sample, all of which produce a fair amount of variation in the considered aggregate conditions.

To properly interpret these regressions as evidence of sorting, the composition of job ads and workers should be unchanged over the business cycle. Otherwise, estimated changes in correlations during the cycle may be generated by cyclical composition changes of postings and/or job seekers. We address this possibility by controlling for compositional changes in our sample using the reweighing technique of DiNardo, Fortin, and Lemieux (1996) (DFL henceforth). We implement the method by first choosing the composition of jobs and workers in 2017, the year with an unemployment rate (6.93%) closest to the sample average (6.92%). We run a probit model estimating the probability of being at 2017 as a function of observables on the applicant side X_w (gender, gender cubic age profiles, marital status categories, region of residence, and educational level categories) and on the job ad side X_a (firm industry, firm size, region, required educational area and educational level categories, and required minimum experience). We compute a predicted probability and define a weight for a worker w and a job a in time t as

$$\varphi_{awt} = \frac{\Phi(\widehat{\pi}_w X_w + \widehat{\pi}_a X_a)}{1 - \Phi(\widehat{\pi}_w X_w + \widehat{\pi}_a X_a)},$$

[10]Our source is ENE. We use the official updated version frequency weights from the 2017 Census.

Table 4: Cyclical assortative matching, constant composition

	EW	AKM	AKM-f
ρ	0.615***	0.149***	0.200***
	(0.014)	(0.018)	(0.018)
δ	-0.007***	-0.004***	-0.000**
	(0.012)	(0.016)	(0.016)
Obs	39,062,160	39,052,247	39,052,095
R^2	0.396	0.053	0.053
Δ_c	-0.0530	-0.0323	-0.0027

Note: 100X Robust standard error in parenthesis. We report mean correlation ρ and cyclical sensitivity δ as defined by equation (4). The cyclical measure is the Chilean non-seasonally adjusted unemployment rate reweighed according to the 2017 Census, as published by Instituto Nacional de Estadísticas. Regressions use DFL weights (see in the main text)

where $\Phi(\cdot)$ stands for the cumulative density of a standard normal distribution and $\widehat{\pi}$ represent estimates. To appropriately implement this method, we consider that being in 2017 to be a treatment, and its probability to be a propensity score for treated (2017) and non-treated (not 2017) groups. In Appendix A3, Figure A2 we depict kernel estimates of these propensity scores densities. Due to the common support assumption (i.e, each observation must have a non-zero probability of being in both groups), we trim a very few number of observations that have extreme probabilities of being observed in 2017.

Table 4 shows the estimates for ρ and δ, when we consider EW, AKM and AKM-f types. The estimated EW type correlation when the unemployment rate is at its average is 0.615, very close to the cross-section correlation in Section 4. The table also shows that all estimates of ρ are negative, which indicates that sorting is *procyclical*: when aggregate business cycle conditions improve (and unemployment decreases), sorting at the application level increases, and vice versa.

To gauge the variability of sorting over the cycle, we compute $\Delta_c = \delta(z^*_{max} - z^*_{min})$ where z^* is the unemployment rate (cyclical measure, as mentioned above) which has been partialled out from time effects (linear trend and monthly seasonal dummies). From the table, we see that $\Delta_c = -0.053$ for EW types, which implies that the average correlation at the nadir of the cycle (maximum unemployment rate) is 0.053 lower than the average correlation at the peak of the cycle (minimum unemployment rate). As shown in the previous section, the average correlations of AKM and AKM-f types are substantially lower. On the other hand, cyclicality of sorting is relatively more importante for AKM compared to EW types, since the value of Δ_c relative to that of ρ is higher in the earlier case: $0.22 = 0.0323/0.149$ versus $0.09 = 0.0530/0.615$, respectively. Finally, we find that cyclicality of sorting according to the AKM-f measure is almost insignificant.

Robustness. We do some robustness exercises to show that procyclical sorting is not driven by some auxiliar assumptions we have made for our baseline results. In the Appendix A3, Table A6 we present a more flexible specification with year fixed effects. Although the latter may absorb some of the cyclical variation of z, the results barely change. In the same Table, we also report estimates without DFL weights, which shows that DFL by itself is not substantially driving the results. This also shows that the procyclicality of sorting is genuinely a behavioral change of agents, and not a compositional change of ads or workers.

In Table A7 we also report results for three subsamples of interest: unemployed job seekers, employed job seekers, and applications done after the last CV update. As in the previous section, the average EW correlation measure at the average unemployment rate is not far from the benchmark baseline of 0.6. Both employed and unemployed EW sorting is procyclical, but the pattern is greater for employed job seekers. To summarize this result, the statistic $\Delta_c = -0.079$ shows that a sizable change in sorting among the employed job seekers is due to cyclical fluctuations. These findings suggest that the observed procyclicality overall is particularly driven by the application behavior of on-the-job seekers. AKM types, separated by employment status, exhibit procyclical sorting as well.

Sorting is acyclical in practice for AKM-f types. To the extent that in AKM-f types we take away the firm specific component, the aforementioned finding means that the pure job title / occupational component barely changes in response to business cycle conditions. In constrast, AKM types are clearly procyclical. The two facts together suggest that the increase in positive assortative matching in response to expansions is mainly due to workers applying to ads with job titles similar to the ones they often have applied in the past, but in firms that are more aligned to their own types. In other words, within the same occupation or job title, high-wage workers tend apply more to high-wage firms in expansions.

In the last panel of Table A7 we show that average sorting for the post CV update sample is higher than the baseline, reaching 0.64 when the unemployment rate is at its average. While sorting remains procyclical, the business cycle is not as important as in the other samples to drive assortativeness of applications. The degree of procyclicality is higher for AKM types in this sample, while AKM-f sorting becomes acyclical.

In sum, studying how sorting evolves over the business cycle leads to two main results. First, as shown before, positive assortative matching can be regarded as high in comparison to the literature, around 0.6, and the cyclical behavior of the labor market may affect this figure but not overturn it. Second, EW, AKM and AKM-f sorting is procyclical, a very robust result as discussed above.

Note once more that this is procyclical sorting at the *pre-match* stage and may reflect sorting of the flow of new job positions rather than the stock (what the rest of the literature focuses on). We see our results as complement to current literature. For example, in Lise and Robin (2017) there are strong forces to produce pro-cyclical sorting. However, these forces apply mostly during the *post-match* stage: in their model, employed have an easier time finding better matches during an expansion due to on-the-job search. On the other hand, higher aggregate productivity during an expansion, increases the acceptance regions of workers and firms, posing an additional force against sorting.

6 From applications to realized matches

Our analysis so far cannot be directly compared to the existing sorting literature which provides evidence (usually) based on administrative matched employer-employee information. Considering that sorting in the stock of worker-job matches can be due to ex ante and ex post channels, it is an empirical question which is more important. Nevertheless, our previous results are suggestive since they show that a high degree of sorting exists ex ante (searching and pre-match stage).

In this section, we use the Chilean employment survey (*Encuesta Nacional de Empleo*, henceforth ENE) to forecast unemployment-to-employment (U2E) and job-to-job (J2J) transitions in the Chilean economy, conditional on a number of observable characteristics.[11] We use these estimates to assess the likelihood that a given application (worker-job pair observed in TC) ends up as a hire. We use these transition probabilities to create sample weights to recompute our sorting estimates from the previous sections. These weights give us an approximation of realized match sorting of the flow of new job creation, making them more comparable to other studies in the literature. Of course, this result hinges on the law of large numbers: we are unable to trace specific hirings out of applications (we don't observe who gets hired in TC), but we can infer the average degree of sorting generated after real matches form.

To be more concrete, we select a vector of covariates $\mathbf{X}$ which are present both in the online job board and in ENE. The probability of being hired is estimated from the survey using a probit model for both those hired from unemployment (U2E) and from another job (J2J).

For U2E transitions, we consider individuals who have exerted some job search effort in the past four weeks, the standard definition. For the J2J calculations, we identify job movers as those individuals who report being employed for two consecutive quarters, but that their most recent employment duration is less than three months. We want to compute transition probabilities of actual job seekers to correctly impute them probabilities of transition in the TC data. Fortunately, the ENE survey inquires about on-the-job search activity.

Variables in $\mathbf{X}$ related to the worker consist of gender, age, educational level, and region of residence.[12] For the firm side, we have information on industry and firm size. We also use year and month dummies to control for time, business cycle and seasonal effects. After estimating the different models using ENE data, we can apply them to worker-job pairs observed in TC, to forecast the likelihood of each application becoming an actual hire. In the Appendix, Figure A3

[11] The ENE is a quarterly rotating panel data, akin to the Current Population Survey (CPS) in the US. We report average transition rates for different groups in Table A8 in the appendix.

[12] Geographic administrative divisions of Chile, akin to US states

shows kernel estimates for these predicted probability for applicants in `trabajando.com`. Equation (5) shows the formula we use for the imputation procedure

$$\text{Prob}\,(\text{Hire}|\mathbf{X}) = \mathbb{I}[\text{Employed}]J2J(\mathbf{X}) + (1 - \mathbb{I}[\text{Employed}])U2E(\mathbf{X}) \tag{5}$$

where observables $\mathbf{X}$ on the right hand side of the equation are from TC applications, $\mathbb{I}$ represents an indicator function for being an employed user of TC and $J2J$ and $U2E$ represent the empirical models estimated from ENE.

We also estimate a model to forecast employment-to-non-employment (E2U) transitions using roughly the same covariates we described before, except for the omission of firm size.[13] With this information, we can forecast the likelihood of a job separation given covariates. Our ultimate goal here is to approximate a long-run employment probability under the assumption that transition rates remain stable over time. Using standard dynamic equations that are common in search models in steady state, we approximate the long-run probability of being employed by

$$\text{Prob}(\text{Long-run Employment}|\mathbf{X}) = \frac{U2E(\mathbf{X})}{U2E(\mathbf{X}) + E2U(\mathbf{X})\,(1 - J2J(\mathbf{X}))} \tag{6}$$

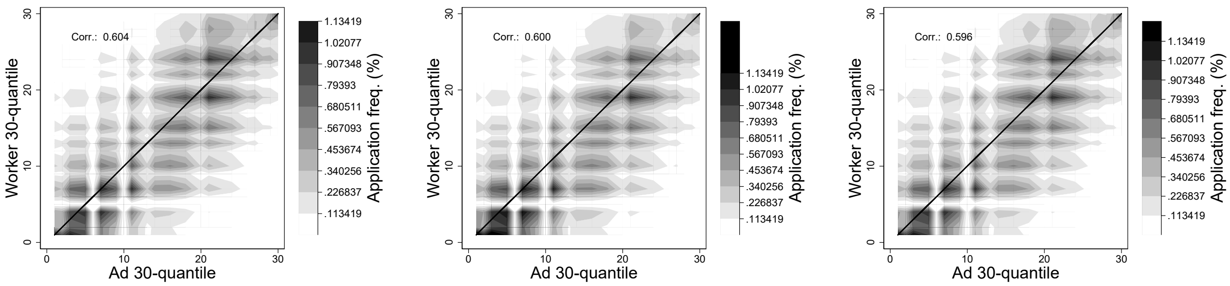
in which we have assumed that a job-to-job transition precludes a job separation (E2U transition) within a quarter. The distribution of predicted hiring and long run employment probabilities are shown in Figure A3 in the Appendix.

We recognize the limitation of this approach, as the hiring process probably depends on a large list of factors that are not observed in ENE data such as job requirements and their fit with the applicant's profile. These estimations just have a predictive purpose and do not try to uncover any causal relationship between applicant and firm traits on hiring probabilities.

With these predicted probabilities, we redo the sorting analysis in preceding sections to assess the degree of assortative matching in a sample that is closer to the actual flow of matches (using hiring probabilities in eq. 5) and in the employment stock (using long-run employment probabilities in eq. 6). As in previous sections, we also study how these features evolve over the business cycle.

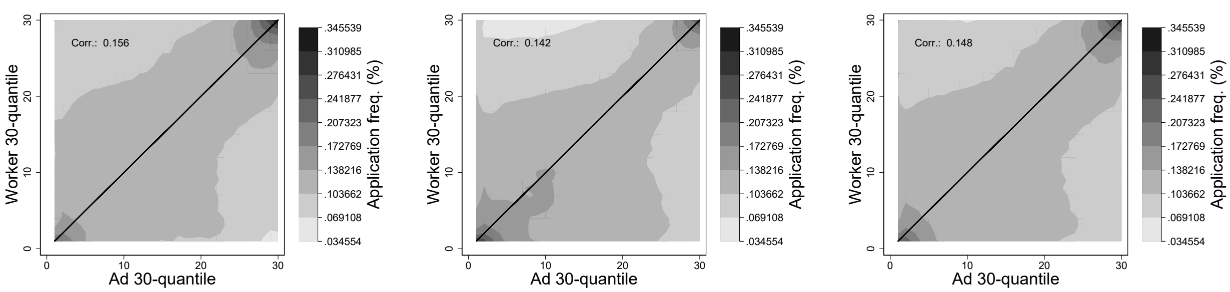
From Figure 2 we observe the original pre-match correlation, in subfigure (a). Subfigure (b) depicts the empirical joint probability distribution of types when weighted by the hiring probability. In subfigure (c) we observe the same distribution, but weighted by the long-run probability of employment. The shapes look remarkably similar with a slight variations in areas near the median of ad types and low and high types of workers. The computed correlations for EW types vary only marginally and remain around 0.6. As the overall sorting does not change when types are reweighed, we conclude that the application or search stage is, by far, the most important stage in the labor market to generate the observed sorting.

[13] For the unemployed we use information regarding the sector of their last employer. Nevertheless, the questionnaire in ENE do not include information regarding the size of the last employer.



(a) EW pre-match (b) EW hiring (c) EW long-run employment

Figure 2: Frequencies of applications, by worker and ad percentiles of EW types reweighed by the estimated probability given observables. Details of the computation of predicted probabilities are in the text. Sample consists on 39,480,855 applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`



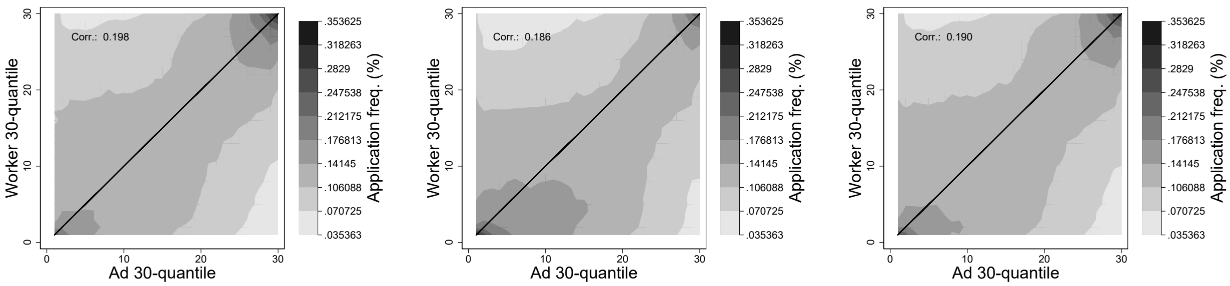
(a) AKM pre-match (b) AKM hiring (c) AKM long run employment

Figure 3: Frequencies of applications, by worker and ad percentiles of AKM types reweighed by the estimated probability given observables. Details of the computation of predicted probabilities are in the text. Sample consists on 39,480,855 applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`

Figure 3 shows a marginal decline in sorting of AKM types. While the correlation of types remain quite stable nearly 0.15, the shape of the joint distribution of types slightly changes between application and hiring stages as there is a larger alignment for low types, reflected in a darker area at the bottom-left of subfigure (b), and also an enlargement of the area for average joint frequencies. Subfigure (c) looks similar to (a) indicating a marginal increase of sorting in the long term. Despite these differences, there is no economically relevant variation between these subfigures.

Results in Figure 4 are qualitatively similar to those in Figures 2 and 3. Despite some changes in the shape of the joint distribution of types at the pre-match and hiring stages, especially for low types, the correlations remain very similar and close to 0.19. Again, combining evidence from EW, AKM and AKM-f types leads us to conclude that sorting is generated at the application stage to a great extent.

Table 5 shows the cyclical correlation obtained when weighting observations by the hiring and



(a) AKM-f pre-match (b) AKM-f hiring (c) AKM-f long run employment

Figure 4: Frequencies of applications, by worker and ad percentiles of AKM-f types reweighed by the estimated probability given observables. Details of the computation of predicted probabilities are in the text. Sample consists on 39,480,855 applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`

Table 5: Cyclical assortative matching, constant composition, ex post reweigh

	Hired prob weight			Employment prob weight		
	EW	AKM	AKM-f	EW	AKM	AKM-f
ρ	0.607***	0.134***	0.186***	0.609***	0.142***	0.194***
	(0.015)	(0.020)	(0.020)	(0.014)	(0.018)	(0.018)
δ	-0.005***	-0.006***	-0.000	-0.009***	-0.003***	-0.000
	(0.013)	(0.017)	(0.017)	(0.012)	(0.016)	(0.016)
Obs	39,039,567	39,029,655	39,029,503	39,039,567	39,029,655	39,029,503
R^2	0.388	0.049	0.049	0.391	0.054	0.053
Δ_c	-0.0377	-0.0440	-0.0017	-0.0592	-0.0219	-0.0017

Note: 100X robust standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$

the long-run employment probabilities. For EW types, the hiring probability weighting implies somewhat lower procyclicality of sorting although it clearly remains significant. This is reflected in the comparison of Δ_c in both estimations: -0.0530 in the pre-match versus -0.0377 at the hiring stage, which refers to the maximum variation of types correlation attributable to business cycle fluctuations in the sample. In contrast, the long-term weighted correlation shows a slightly higher procyclicality of sorting than the ex ante case discussed in Section 5. For the AKM types, the results show the reverse pattern. The sorting weighted by the hiring probability exhibits larger procyclicality, while long-term employment exhibits a somewhat lower value. Both remain significant. In the case of AKM-f types, sorting turns out to be acyclical when weighted as it is in the ex ante stage.

Robustness: We estimate ex post sorting on additional subsamples of interest: unemployed and on-the-job seekers. In Table A9 and Figures A4 and A5 we report the weighted correlations and bivariate empirical distributions of types weighted to approximate the sorting patterns of the flow and stock of realized matches. These estimates ought to be compared to those in Table A4. For the unemployed, the ex ante sorting of EW types (0.588) slightly decreases when weighted by hiring (0.576) and long-term employment probabilities (0.566). For on-the-job seekers, the sorting of EW types for the just hired increases with respect to the the pre-match stage (0.568 vs 0.581), but decreases to 0.558 when correlation is computed with long-term employment probabilities. For AKM and AKM-f types both types of ex post weights marginally decrease sorting. These estimates point in the same direction as the ones obtained for the whole sample. From the application stage no much additional sorting is gained, if any, in realized matches. As Faberman, Şahin, Ayşegül, and Topa (2019) have shown, on-the-job search may be quite different from unemployed search in terms of effort and outcomes. Thus, we should expect notorious differences in realized matches for these two groups. Although we observe little effect of hiring and long-term employment probability weighting in sorting for the whole sample, it cold be possible to see different behavior because transition probabilities for J2J and U2E flows are estimated separately. Even so, we find that splitting the sample by labor status barely affects the sorting of realized matches. Hence, we conclude that our evidence for the whole sample is not driven by labor force status.

7 Conclusions

In this paper we revisit the question of whether workers and job positions are sorted in any meaningful way in the job market. By analyzing information from an online job board along household survey data from the Chilean economy, we are able to contribute to the debate in three dimensions.

First, we provide sharp and robust evidence of positive assortative matching in the labor market. The correlation between workers and job positions, using our preferred measure, hovers around 0.6. We show that sorting exists in the cross section and also exhibits cyclical variation: it is procyclical albeit its economic significance is muted.

Second, we advance the discussion on measurement of types and show how expected wages (to be paid at the position and expected wages by workers) as types, the exact suggestion by Abowd, Kramarz, and Margolis (1999) and Borovickova and Shimer (2020), are directly related to the so called AKM fixed effects. We compute both measures from our data and show how they are related.

Third, we use data from a Chilean employment survey to impute probabilities of actually forming a match in the short and long run. We compute correlations with these weights to generate a measure of sorting of realized matches, a measure more comparable to estimates from matched employer-employee datasets. Pre-match and ex post correlations are very similar for three types, leading us to conclude that application decisions are the most important margin behind the strong positive assortative matching observed.

Finally, we argue that worker-job assignments are a meaningful dimension to consider, as opposed to the more traditional worker-firm sorting, given new literature that considers the match of workers to either jobs and occupations, which by definition represent a finer dimension than firms.

Our results suggest that theoretical settings, along the lines of Shimer (2005) and Abowd, Kramarz, Pérez-Duarte, and Schmutte (2018), in which directed search is a key ingredient, are a desirable path for future research in this area.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 2(67), 251–333.

Abowd, J. M., F. Kramarz, S. Pérez-Duarte, and I. M. Schmutte (2018): "Sorting Between and Within Industries: A Testable Model of Assortative Matching," *Annals of Economics and Statistics*, (129), 1–32.

Andrews, M. J., S. Bradley, D. Stott, and R. Upward (2008): "Successful Employer Search? An Empirical Analysis of Vacancy Duration Using Micro Data," *Economica*, 75(299), 455–480.

Bagger, J., and R. Lentz (2018): "An Empirical Model of Wage Dispersion with Sorting," *The Review of Economic Studies*, 86(1), 153–190.

Baley, I., A. Figueiredo, and R. Ulbricht (2020): "Mismatch Cycles," Working paper.

Banfi, S., and B. Villena-Roldan (2019): "Do High-Wage Jobs Attract more Applicants? Directed Search Evidence from the Online Labor Market," *Journal of Labor Economics*, 37(3), 715–746.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813–846.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020): "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?," Working Paper 27368, National Bureau of Economic Research.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87(3), 699–739.

Borovickova, K., and R. Shimer (2020): "High Wage Workers Work for High Wage Firms," Unpublished manuscript.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29(2), 238–249.

Card, D., J. Heining, and P. Kline (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 128(3), 967–1015.

Crane, L., H. Hyatt, and S. Murray (2018): "Cyclical Labor Market Sorting," Discussion paper.

Cubas, G., and P. Silos (2020): "Social Insurance and Occupational Mobility," *International Economic Review*, 61(1), 219–240.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–1044.

Dvorkin, M. A., and A. Monge-Naranjo (2019): "Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations," *FRB St. Louis Working Paper*, (2019-13).

Eeckhout, J., and P. Kircher (2011): "Identifying Sorting–In Theory," *The Review of Economic Studies*, 78(3), 872–906.

Faberman, J., A. Şahin, Mueller, Ayşegül, and G. Topa (2019): "Job Search Behavior among the Employed and Non-Employed," Working paper, Federal Reserve Bank of New York.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer (2020): "Multidimensional Skill Mismatch," *American Economic Journal: Macroeconomics*, 12(1), 210–44.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85(1), 29–65.

Jolivet, G., B. Jullien, and F. Postel-Vinay (2016): "Reputation and Prices on the e-Market: Evidence from a Major French Platform," *International Journal of Industrial Organization, forthcoming.*

Jolivet, G., and H. Turon (2014): "Consumer Search Costs and Preferences on the Internet," Mimeo, University of Bristol.

Kim, K., and P. Kircher (2015): "Efficient competition through cheap talk: the case of competing auctions," *Econometrica*, 83(5), 1849–1875.

Kudlyak, M., D. Lkhagvasuren, and R. Sysuyev (2013): "Systematic Job Search: New Evidence from Individual Job Application Data," mimeo, Federal Reserve Bank of Richmond.

Lamadon, T., M. Mogstad, and B. Setzler (2019): "Imperfect competition, compensating differentials and rent sharing in the US labor market," Discussion paper, National Bureau of Economic Research.

Lewis, G. (2011): "Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors," *The American Economic Review*, 101(4), 1535–1546.

Lise, J., and J.-M. Robin (2017): "The macrodynamics of sorting between workers and firms," *American Economic Review*, 107(4), 1104–35.

Lopes de Melo, R. (2018): "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*, 126(1), 313–346.

Marinescu, I., and R. Wolthoff (2020): "Opening the Black Box of the Matching Function: The Power of Words," *Journal of Labor Economics*, 38(2), 535–568.

Moen, E. R. (1997): "Competitive Search Equilibrium," *The Journal of Political Economy*, 105(2), 385–411.

Shimer, R. (2005): "The assignment of workers to jobs in an economy with coordination frictions," *Journal of Political Economy*, 113(5), 996–1025.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019): "Firming up inequality," *The Quarterly journal of economics*, 134(1), 1–50.

Taber, C., and R. Vejlin (2020): "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 88(3), 1031–1069.

White, H. (1980): "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, pp. 817–838.

Appendix: Sorting on-line and on-time

A1 Additional Descriptive Statistics

Table A1: Characteristics of Job Seekers

	Employed	Unemployed	Total
Marital status (%)			
Married	25.44	30.79	27.90
Partner	2.23	2.06	2.15
Divorced	2.54	2.50	2.52
Separated	2.46	2.29	2.38
Single	67.04	62.17	64.79
Widow	0.29	0.19	0.25
Nationality (%)			
Argentina	0.38	0.49	0.43
Bolivia	0.18	0.14	0.16
Brazil	0.17	0.17	0.17
Chile	87.70	89.26	88.41
Colombia	1.20	0.83	1.03
Ecuator	0.18	0.16	0.17
U.S.A.	0.03	0.05	0.04
Spain	0.30	0.32	0.31
France	0.04	0.05	0.05
Italy	0.04	0.08	0.06
Mexico	0.06	0.07	0.07
Other	2.06	1.91	1.99
Peru	0.68	0.54	0.61
Uruguay	0.06	0.08	0.07
Venezuela	6.92	5.86	6.43
Region of Residence (%)			
I Tarapaca	1.56	1.47	1.52
II Antogafasta	4.07	4.08	4.08
III Atacama	1.33	1.45	1.39
IV Coquimbo	3.23	2.77	3.02
V Valparaiso	10.02	8.55	9.34
VI OHiggins	3.28	3.10	3.20
VII Maule	2.90	2.74	2.83
VIII Bio Bio	6.77	6.08	6.45
IX Araucania	2.84	2.32	2.60
X Los Lagos	2.69	2.44	2.57
XI Aysen	0.25	0.26	0.25
XII Magallanes	0.48	0.53	0.50
RM Metropolitana	56.10	60.08	57.94
XIV Los Rios	1.30	1.10	1.21
XV Arica y Parinacota	1.11	0.87	1.00
Foreigners	1.19	0.97	1.08
Observations	919,436	788,704	1,710,007

Table A2: Characteristics of Job Seekers

	Employed	Unemployed	Total
Days job searching (%)			
1 day	26.52	22.45	24.67
2-10 days	4.96	2.86	3.99
11 - 25 days	4.15	2.45	3.36
26 - 60 days	5.70	3.58	4.72
61 - 90 days	3.21	2.21	2.75
91 - 140 days	4.14	3.00	3.61
141 - 210 days	4.47	3.58	4.06
¿ 210 days	46.86	59.86	52.83
Job searching			
Days (Mean/(S.D.))	614.16	882.04	737.34
	(886.43)	(1017.37)	(958.31)
Ads applied (%)			
1	17.70	16.49	17.16
2	10.44	9.29	9.91
3	7.46	6.52	7.03
4 - 6	14.32	12.68	13.56
7 - 10	10.91	9.98	10.48
11 - 20	13.69	13.44	13.56
21 - 30	6.86	7.30	7.06
Ads Applied			
Applications per worker (Mean/(S.D.))	25.60	34.07	29.50
	(87.07)	(92.65)	(89.76)
Applications to explicit wages per worker (Mean/(S.D.))	3.02	2.64	2.84
	(10.84)	(7.92)	(9.61)
Explicit Wages			
% Appl. WP/Total Appl. (Mean/(S.D.))	13.602	9.095	11.522
	(21.232)	(17.225)	(19.621)
Observations	919,436	788,704	1,710,007

Table A3: Characteristics of Job Ads

	Implicit wage	Explicit wage	Total
Applications per ad (%)			
1	8.80	13.31	9.54
2	6.29	9.60	6.83
3	4.99	7.44	5.39
4 - 5	7.51	10.92	8.07
6 - 10	12.25	15.94	12.85
11 - 20	14.12	14.92	14.25
21 - 50	20.17	15.26	19.36
$>$ 50	25.88	12.61	23.70
Applications per ad			
Applications (Mean/(SD))	46.736	25.500	43.264
	(92.391)	(63.595)	(88.674)
Days Ad available			
Days (Mean/(SD))	55.168	47.885	53.977
	(3022.577)	(793.486)	(2782.967)
Observations	975,224	190,661	1,165,885

A2 Sorting on line: additional analysis

Table A4: Cross-sectional assortative matching: Additional samples I

	All			Post CV update		
	EW	AKM-like	AKM-firm	EW	AKM	AKM-firm
ρ	0.604***	0.156***	0.198***	0.658***	0.173***	0.222***
	(0.013)	(0.017)	(0.017)	(0.045)	(0.060)	(0.061)
Obs	39,480,855	39,469,362	39,469,181	3,517,879	3,511,752	3,511,684
R^2	0.365	0.024	0.039	0.433	0.030	0.049

	Unemployed			Employed		
	EW	AKM	AKM-firm	EW	AKM	AKM-firm
ρ	0.588***	0.122***	0.166***	0.568***	0.147***	0.186***
	(0.020)	(0.025)	(0.025)	(0.018)	(0.023)	(0.023)
Obs	18,923,584	18,917,132	18,916,995	20,538,765	20,533,768	20,533,724
R^2	0.346	0.015	0.027	0.323	0.022	0.035

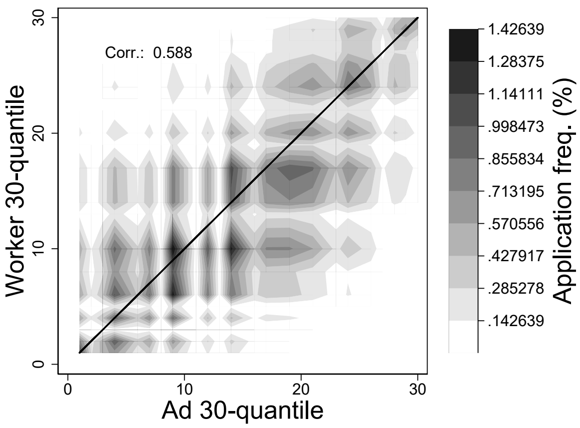
Note: 100X robust standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Table A5: Cross-sectional assortative matching: Additional samples II

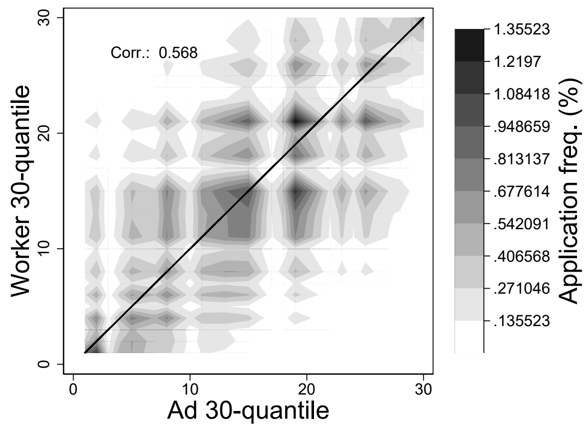
	Explicit wage ad			Hidden wage ad		
	EW	AKM	AKM-f	EW	AKM	AKM-f
ρ	0.570***	0.150***	0.178***	0.598***	0.154***	0.198***
	(0.051)	(0.057)	(0.056)	(0.014)	(0.018)	(0.018)
Obs	4,011,133	4,007,920	4,007,854	35,469,722	35,461,442	35,461,327
R^2	0.325	0.022	0.032	0.357	0.024	0.039

	Explicit wage worker			Hidden wage worker		
	EW	AKM	AKM-f	EW	AKM	AKM-f
ρ	0.617***	0.177***	0.201***	0.595***	0.143***	0.192***
	(0.027)	(0.036)	(0.035)	(0.015)	(0.020)	(0.019)
Obs	9,282,586	9,277,607	9,277,518	30,198,269	30,191,755	30,191,663
R^2	0.381	0.031	0.040	0.355	0.020	0.037

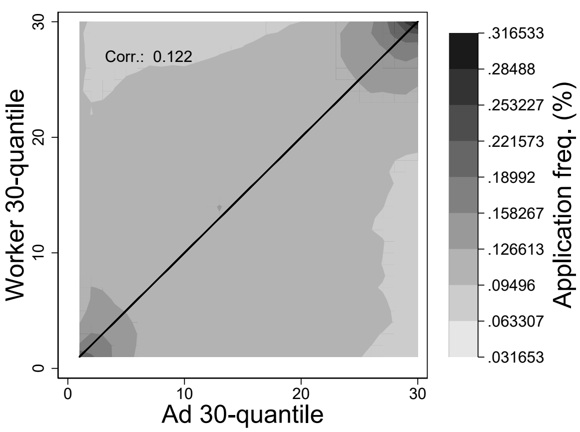
Note: 100X robust standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$



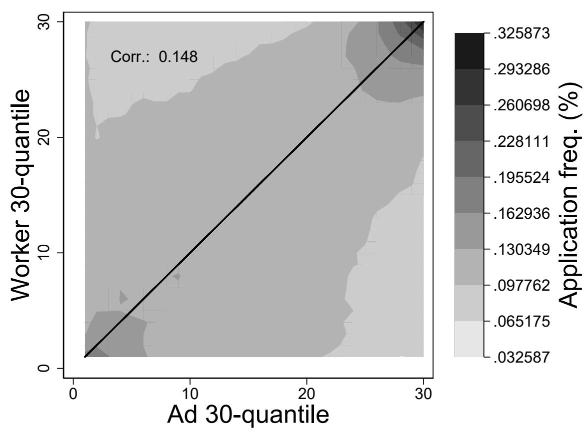
(a) Unemployed, EW types



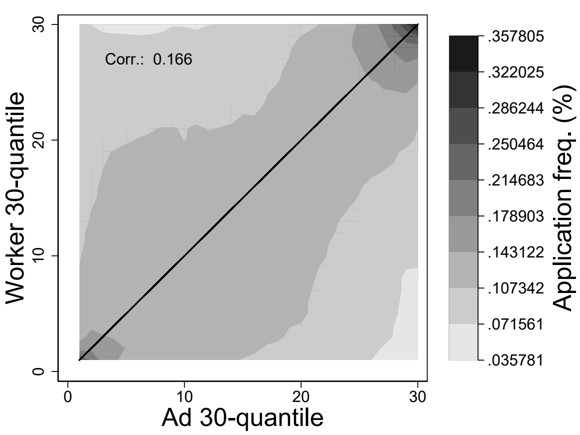
(b) Employed, EW types



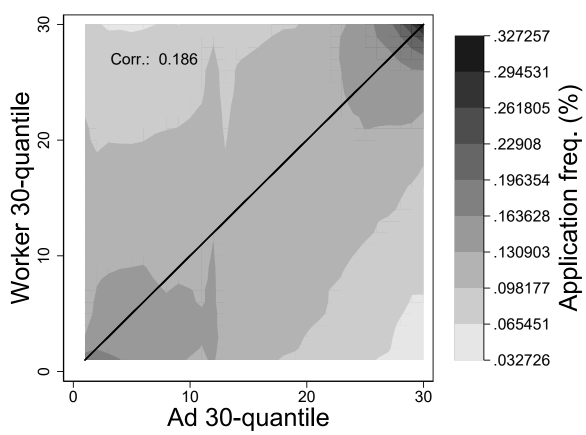
(c) Unemployed, AKM types



(d) Employed, AKM



(e) Unemployed, AKM-f types



(f) Employed, AKM-f types

Figure A1: Frequencies of applications, by worker and ad percentiles of EW, AKM and AKM-f types by employment status. All applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`.

A3 Sorting on time: additional analysis

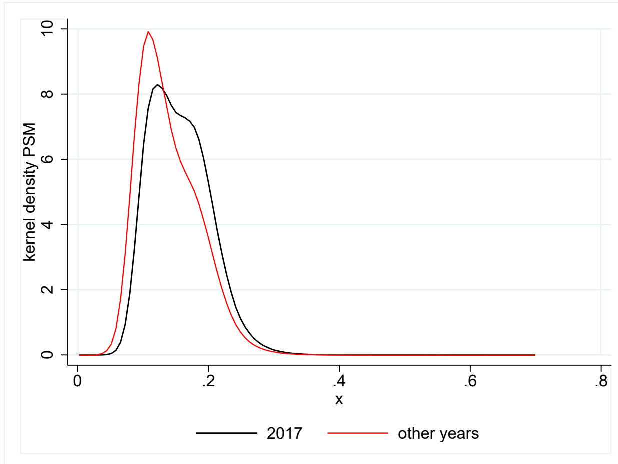


Figure A2: Estimated propensity score matching for the DFL reweighing compositional adjustment. Observations outsuide of the common support of the distrution are trimmed.

Table A6: Cyclical assortative matching: Additional samples I

	Year FE / DFL			no DFL		
	EW	AKM	AKM-f	EW	AKM	AKM-f
ρ	0.614***	0.148***	0.200***	0.623***	0.159***	0.205***
	(0.014)	(0.018)	(0.018)	(0.013)	(0.017)	(0.017)
δ	-0.007***	-0.004***	-0.000	-0.008***	-0.004***	-0.001***
	(0.012)	(0.016)	(0.016)	(0.012)	(0.016)	(0.016)
Obs	39,062,160	39,052,247	39,052,095	39,175,504	39,164,399	39,164,225
R^2	0.396	0.054	0.053	0.400	0.056	0.054
Δ_c	-0.0574	-0.0325	-0.0011	-0.0509	-0.0260	-0.0086

Notes: 100X Standard error in parenthesis. We report mean correlation ρ and cyclical sensitivity δ as defined by equation (4). The cyclical measure is the Chilean non-seasonally adjusted unemployment rate reweighed according to the 2017 Census, as published by Instituto Nacional de Estadisticas. Regressions use DFL weights (see main text), which are computed using a probit model in which the dependent variable is an indicator for 2017, and independent variables are for applicants: age, age squared, gender, gender interacted with age terms, and a full array of indicators of nationality, marital status, region, educational mayor. Independent variables for job ads are indicators for region, industry, economic activity (job board classification), educational area required, and firm size category.

Table A7: Cyclical assortative matching: Additional samples II

	Unemployed		
	EW	AKM	AKM-f
ρ	0.578***	0.102***	0.157***
	(0.021)	(0.026)	(0.027)
δ	-0.006***	-0.006***	0.001***
	(0.019)	(0.023)	(0.023)
Obs	18,764,034	18,758,117	18,757,991
R^2	0.368	0.049	0.040
Δ_c	-0.0473	-0.0458	0.0074

	Employed		
	EW	AKM	AKM-f
ρ	0.587***	0.147***	0.194***
	(0.019)	(0.025)	(0.025)
δ	-0.010***	-0.003***	-0.001***
	(0.017)	(0.022)	(0.022)
Obs	20,284,699	20,280,714	20,280,688
R^2	0.367	0.054	0.053
Δ_c	-0.0709	-0.0182	-0.0084

	Post CV update		
	EW	AKM	AKM-f
ρ	0.640***	0.138***	0.203***
	(0.050)	(0.065)	(0.066)
δ	-0.002***	-0.006***	0.001
	(0.046)	(0.059)	(0.061)
Obs	3,488,828	3,483,492	3,483,429
R^2	0.424	0.041	0.049
Δ_c	-0.0180	-0.0582	0.0065

Notes: 100X Standard error in parenthesis. We report mean correlation ρ and cyclical sensitivity δ as defined by equation (4). The cyclical measure is the Chilean non-seasonally adjusted unemployment rate reweighed according to the 2017 Census, as published by Instituto Nacional de Estadísticas. Regressions use DFL weights (see main text), which are computed using a probit model in which the dependent variable is an indicator for 2017, and independent variables are for applicants: age, age squared, gender, gender interacted with age terms, and a full array of indicators of nationality, marital status, region, educational mayor. Independent variables for job ads are indicators for region, industry, economic activity (job board classification), educational area required, and firm size category.

A4 Sorting ex post: additional analysis

Description of Encuesta Nacional de Empleo (ENE): The ENE is the official employment survey in Chile, conducyed by the *Instituto Nacional de Estadísticas* (INE) to produce official labor force statistics. It is a quarterly rotating panel survey in which urban households remain up to 6 quarters in sample and rural ones, up to 12 quarters. There are unique person identifiers which allows any researcher to trace labor market transitions on a quarterly basis. Table A8 reports the average transition rates by gender, age groups, educational attainment, and region. The total number of transitions is reported at the end.

Table A8: Descriptive statistics *Encuesta Nacional de Empleo*

	J2J	U2E	E2U
Gender			
Females	3.9%	5.6%	5.3%
Males	7.8%	6.5%	4.0%
Age			
25 - 29	6.6%	9.5%	5.4%
30 - 34	6.6%	6.8%	4.7%
35 - 39	6.2%	5.8%	4.3%
40 - 44	5.8%	5.1%	4.4%
45 - 49	5.3%	4.5%	4.5%
>50	4.5%	3.9%	4.4%
Education level			
Primary (1-8 years)	8.9%	7.4%	6.3%
High School	6.6%	6.8%	5.1%
Tech. Tertiary Educ.	4.3%	5.0%	4.0%
College	3.7%	5.2%	3.4%
Post-Graduate	2.9%	2.9%	2.4%
Region			
I Tarapaca	5.0%	6.1%	4.7%
II Antofagasta	4.1%	5.0%	4.1%
III Atacama	5.7%	6.6%	4.7%
IV Coquimbo	7.2%	6.8%	4.7%
V Valparaiso	6.2%	5.9%	4.6%
VI O'Higgins	8.6%	7.4%	5.5%
VII Maule	8.7%	8.2%	5.5%
VIII Biobio	6.1%	6.3%	4.8%
IX La Araucania	7.0%	6.4%	5.6%
X Los Lagos	5.6%	5.2%	4.7%
XI Aysen	6.1%	6.3%	4.0%
XII Magallanes	4.2%	5.1%	3.4%
RM Metropolitana	4.9%	5.5%	4.3%
XIV Los Rios	6.4%	6.6%	5.1%
XV Arica y Parinacota	4.9%	5.3%	4.9%
Number of transitions	108,071	107,175	86,655
Observations		1,391,479	

Notes: Reported numbers are average quarterly transition rates by different categories. The time span is March 2010 to December 2019.

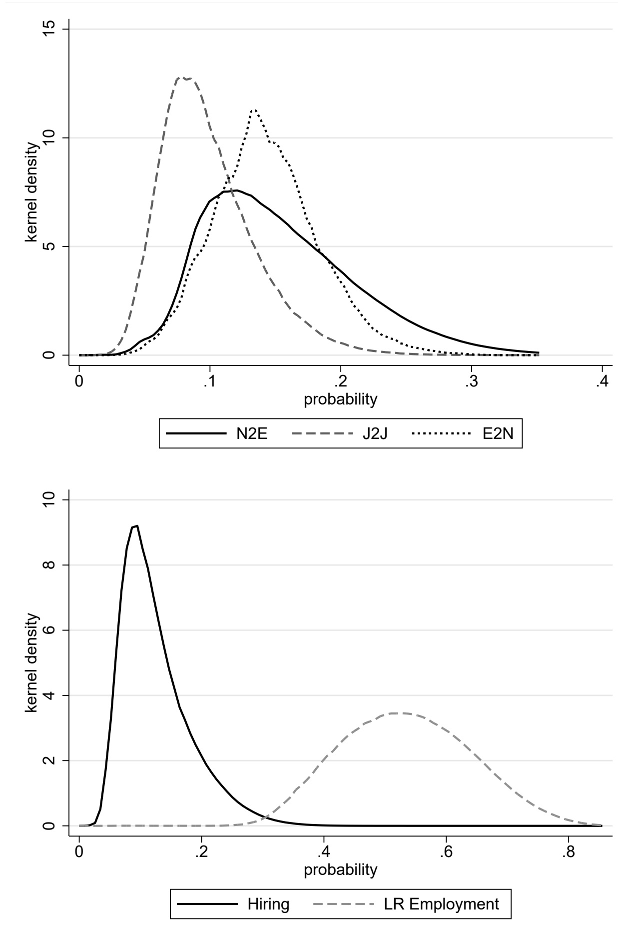


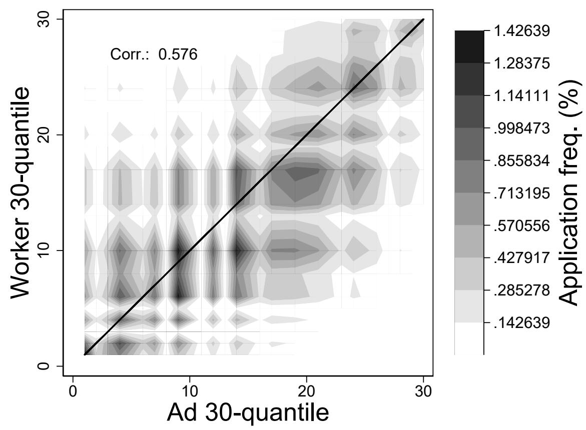
Figure A3: Predicted transition probabilities using ENE data given observables in the website (upper panel) and predicted hiring and long term employment probabilities, i.e. ex post weights (bottom panel)

Table A9: Ex post reweighed cross-section sorting estimates: additional samples

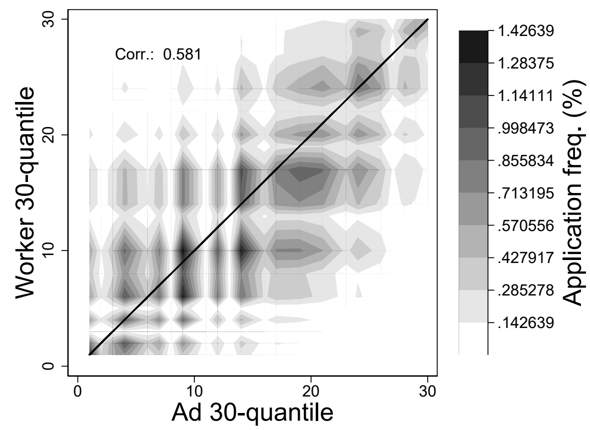
Unemployed						
	Hired prob weight			Employment prob weight		
	EW	AKM	AKM-f	EW	AKM	AKM-f
ρ	0.576***	0.102***	0.153***	0.566***	0.145***	0.182***
	(0.022)	(0.027)	(0.027)	(0.019)	(0.025)	(0.024)
Obs	18,912,751	18,906,299	18,906,162	20,526,727	20,521,731	20,521,687
R^2	0.331	0.010	0.023	0.320	0.021	0.033

Employed						
	Hired prob weight			Employment prob weight		
	EW	AKM	AKM-f	EW	AKM	AKM-f
ρ	0.581***	0.116***	0.160***	0.558***	0.138***	0.177***
	(0.021)	(0.026)	(0.026)	(0.019)	(0.024)	(0.023)
Obs	18,912,751	18,906,299	18,906,162	20,526,727	20,521,731	20,521,687
R^2	0.338	0.013	0.026	0.311	0.019	0.031

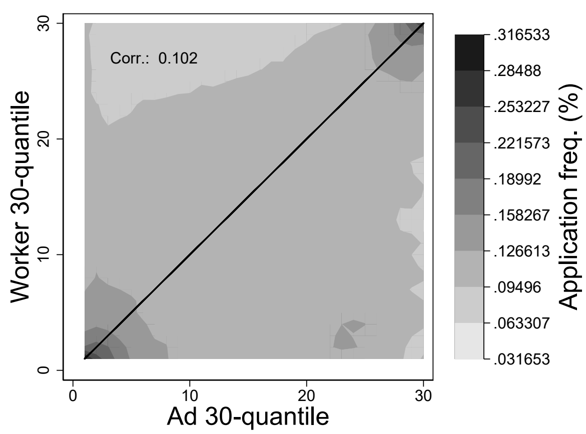
Note: 100X robust standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$



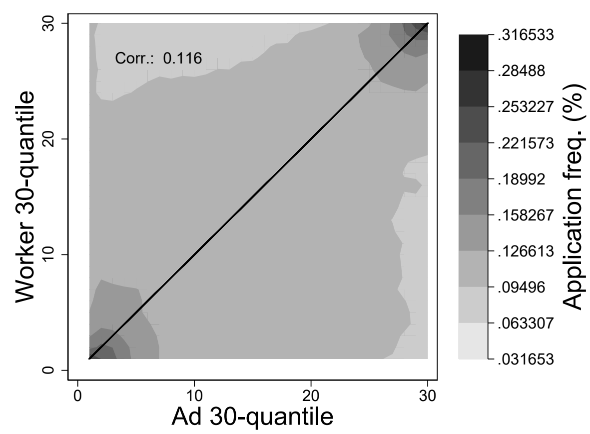
(a) Unemployed, EW types, hirings



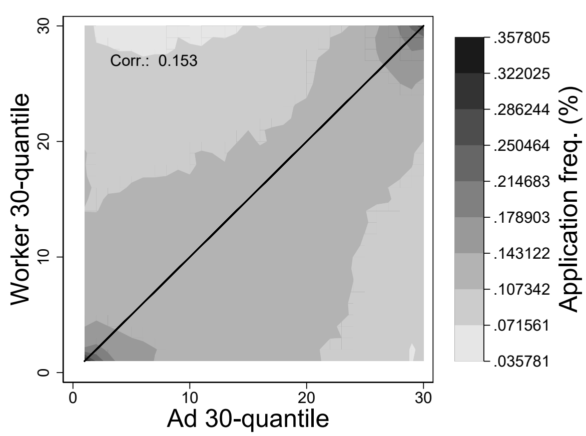
(b) Unemployed, EW types, long-term employment



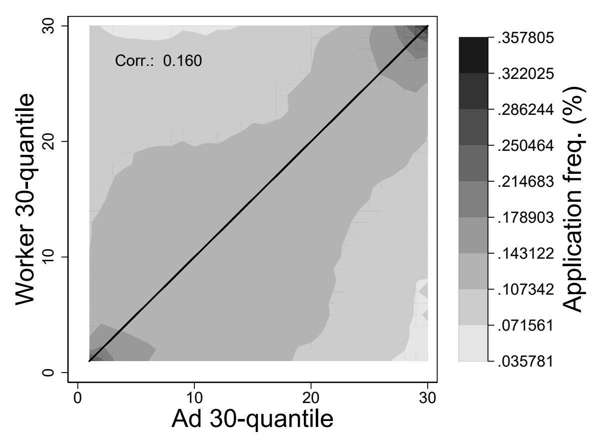
(c) Unemployed, AKM types, hirings



(d) Unemployed, AKM types, long-term employment

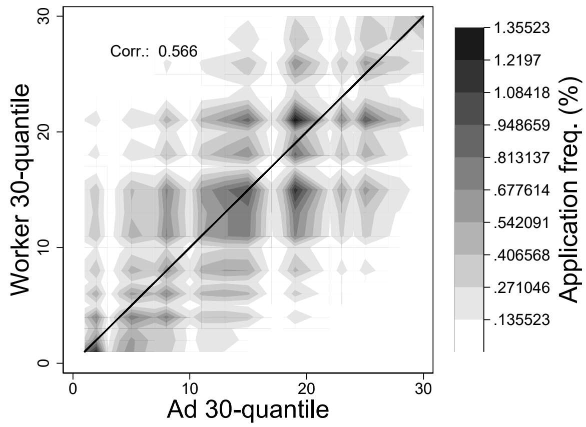


(e) Unemployed, AKM-f types, hirings

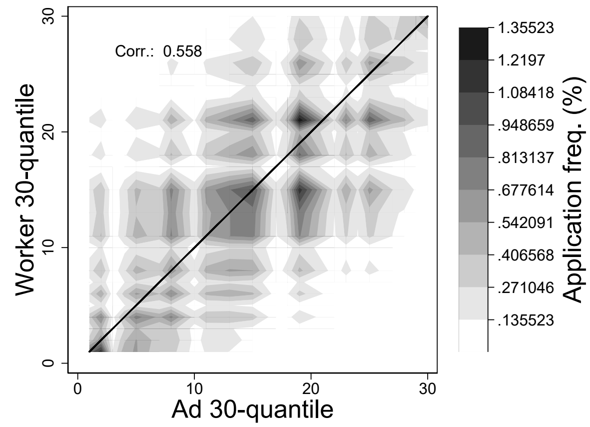


(f) Unemployed, AKM-f types, long-term employment

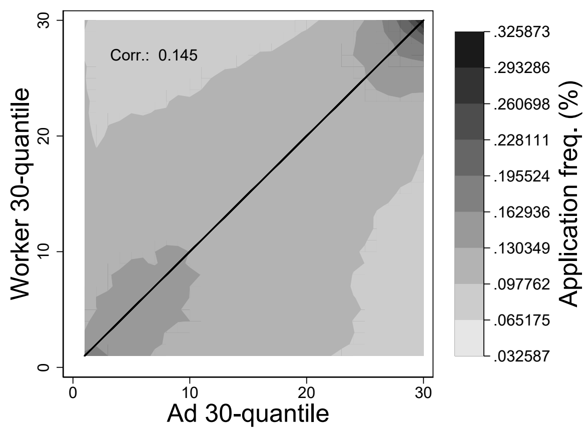
Figure A4: Note: Frequencies of applications, by worker and ad percentiles of EW, AKM, and AKM-f types reweighed by the estimated probabilities of being hired or being employed on the long term, given observables. All applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`.



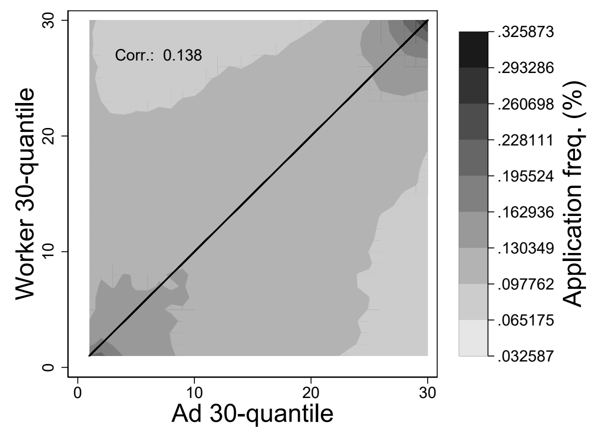
(a) Employed, EW types, hirings



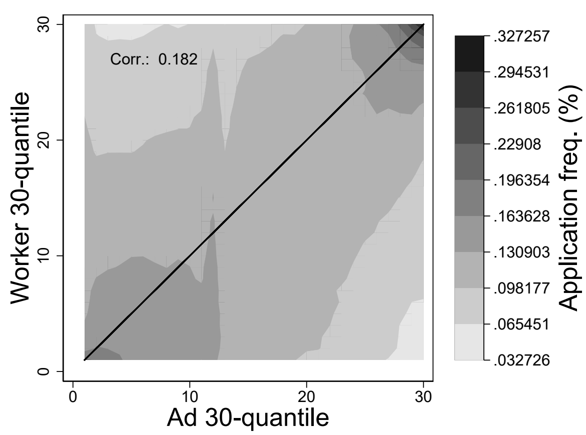
(b) Employed, EW types, long-term employment



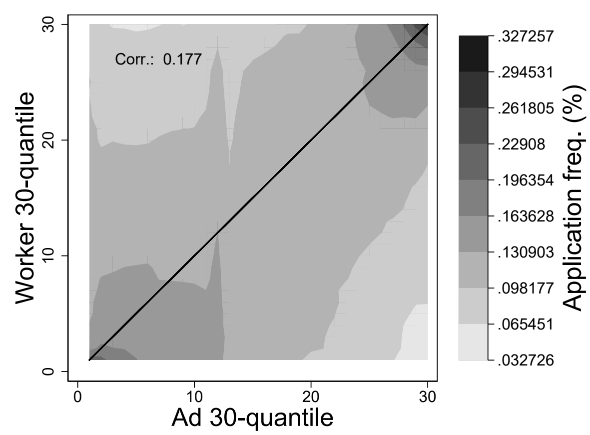
(c) Employed, AKM types, hirings



(d) Employed, AKM types, long-term employment



(e) Employed, AKM-f types, hirings



(f) Employed, AKM-f types, long-term employment

Figure A5: Note: Frequencies of applications, by worker and ad percentiles of EW, AKM, and AKM-f types reweighed by the estimated probabilities of being hired or being employed on the long term, given observables. All applications between 1-Jan-2010 and 31-Dec-2019 in `www.trabajando.com`