Deconstructing Job Search Behavior*

Stefano Banfi
Fiscalía Nacional Económica, Chile

Sekyu Choi
University of Bristol

Benjamín Villena-Roldán
Universidad Andres Bello, LM^{2}C^2, and MIPP.

April 12, 2025

Abstract

Job search outcomes often differ for employed versus unemployed individuals. Using online job board data, we study the key factors driving preceding application decisions. We identify relevant job consideration sets using a network approach based on co-application patterns. We document how demographics and ad timing affect applications, finding evidence consistent with stock-flow matching for the unemployed. Furthermore, we show seekers respond strongly to misalignment in education, experience, wages, and location, generally applying where observable alignment is good, although employed seekers seem more ambitious, showing greater tolerance for underqualification in education and a tendency to apply for jobs above their declared wage expectation. Methodologically, we propose this network approach for defining consideration sets, helping address potential biases in standard market definitions. This evidence contributes to understanding search behavior and differences between seeker types.

Keywords: Online job search, Applications, Unemployment, On-the-job search, Networks.

JEL Codes: E24, J40, J64

*Email: sbanfi@fne.gob.cl, sekyu.choi@bristol.ac.uk and benjamin.villena@unab.cl. We thank Guido Menzio, Brenda Samaniego de la Parra, Jan Eeckhout, Shouyong Shi, Yongsung Chang, Francois Langot, Barış Kaymak, and seminar participants at Cardiff University, University of Manchester, Diego Portales University, the 2016 Midwest Macro Meetings, the 2016 LACEA-LAMES Workshop in Macroeconomic, the Search & Matching workshop at the University of Chile, 2017 SAM Meetings in Barcelona, Spain, 2017 SECHI Meeting, 2017 SM3 Workshop, and Universidad de Los Andes Bogotá, Université de Montréal, Universidad de Chile, the IDSC of IZA Workshop: Matching Workers and Jobs Online, and the 2022 RIDGE Labor Workshop for insightful discussion and comments. Villena-Roldán thanks for financial support the FONDECYT grant project 1191888 and 1251885; ANID Nucleo Milenio Labor Market Mismatch - Causes and Consequences, LM^{2}C^2 (NCS2022_045); and the Institute for Research in Market Imperfections and Public Policy, MIPP (ICS13_002 ANID). The data underlying this article were provided by www.trabajando.com through a legal agreement. Data will be shared on request to the corresponding author with the permission of the legal representative of www.trabajando.com. We are indebted to Ignacio Brunner, Álvaro Vargas, and Ramón Rodríguez for valuable practical insights regarding the data structure and labor market behavior. All errors are ours.

1 Introduction

Job search behavior is a fundamental determinant of wages and job (re)allocation, and empirical evidence consistently shows significant performance differentials between employed and unemployed job seekers (Faberman, Mueller, Şahin, and Topa, 2022). However, much of the existing research focuses on the outcomes of search, such as realized hires, employment spells, or accepted wages. To shed light on the process preceding these outcomes, particularly the choices made by individuals actively seeking employment, we use detailed data from the online job posting website `www.trabajando.com` in Chile and a network-based definition of individual labor markets, leveraging the linkages of job seekers through applications to the same job ads. Using these methods and a dataset containing rich information on both job ads and the job seekers applying to them, we analyze key factors behind observed applications.

The main focus of our paper is the selective component of job search. From the employer's perspective, jobs are complex objects with several required dimensions such as educational level and experience. Workers, on the other side of the market, possess a set of qualifications that may potentially match job requirements. A natural question is how the misalignment between a job seeker's qualifications and the requirements posted in job ads affects the probability of an application. Furthermore, does this application behavior differ systematically based on the current employment status of the seeker? Addressing these questions requires not only detailed data on applications but also a meaningful way to define the set of potential jobs a seeker realistically considers. Usual definitions of local labor markets, often based on fixed geographic or occupational cells (Şahin, Song, Topa, and Violante, 2014; Lamadon, Mogstad, and Setzler, 2022), may inadequately capture the fluid nature of actual search behavior since individuals frequently apply across such boundaries, a pattern consistent with mobility findings in other contexts (Carrillo-Tudela and Visschers, 2023; Jarosch, Nimczik, and Sorkin, 2024).

We contribute to the literature in three primary ways. First, we provide an in-depth analysis of job search behavior by studying application decisions, which are concrete actions seekers take before any match is realized, thus revealing potential labor market allocations. To conduct our empirical analysis, we estimate linear probability models using the consideration sets generated by our network algorithm and keeping the market composition constant. We incorporate flexible polynomial controls for misalignment dimensions and their interactions, alongside controls for worker and ad characteristics. We document how various worker and job characteristics influence the likelihood of an application. We find males apply more frequently, particularly if unemployed, while employed married individuals apply less than their single counterparts. The evidence is also consistent with stock-flow matching behavior (Gregg and Petrongolo, 2005; Coles and Petrongolo, 2008)

as the unemployed apply significantly more often to newer job ads. This focus complements studies examining search intensity or duration (Mukoyama, Patterson, and Şahin, 2018; Faberman and Kudlyak, 2019).

Second, we report systematic evidence on how job seekers respond to misalignment across multiple dimensions. A key aspect of the selective component is how individuals react to discrepancies between their own attributes and those required by an ad. We define and measure misalignment in terms of educational level, years of experience, expected wages, geographical distance, and occupation. Our results show that application probability is quite sensitive to this fit. Generally, workers avoid high misalignment; application probabilities tend to decrease as the gap in any dimension grows (except for experience), suggesting seekers target a certain level of misalignment they tolerate. While the overall patterns are similar for employed and unemployed seekers regarding misalignment , we observe differences: employed seekers seem slightly less deterred by being underqualified in education or wage expectations, potentially reflecting greater ambition or better outside options, whereas unemployed seekers are somewhat more likely to apply when overqualified in education or when the offered wage is below their expectation. This contributes directly to understanding sorting (Banfi, Choi, and Villena-Roldán, 2022) and the role of distance in search (Marinescu and Rathelot, 2018).

Third, methodologically, we use and advocate for a network-based definition of labor markets, derived from observed application patterns. We construct individual consideration sets by leveraging the interconnectedness revealed when different seekers apply to the same job ads. In essence, the consideration set for a given seeker includes not only the jobs they applied to, but also jobs pursued by linked co-applicants. Hence, the consideration set for each individual implicitly takes into account geographic and occupational dimensions to the extent they affect application behavior rather than imposing strict boundaries. This contrasts with cell-based methods and avoids the computational burden and potential biases associated with the "agnostic" view that assumes all contemporaneous jobs are considered. Our network approach, based on actual choices rather than predetermined dimensions, is conceptually akin to Nimczik (2023); Jarosch, Nimczik, and Sorkin (2024). Moreover, ignoring the heterogeneity in realistic consideration sets can lead to biased estimates, particularly if the factors influencing inclusion in the set correlate with application determinants (Tenn and Yun, 2008).

By deconstructing job search into application decisions and employing a behaviorally grounded definition of the relevant market, this paper provides novel insights into the selective component of search, the nuanced ways workers respond to job characteristics and potential mismatch, and the subtle but important differences between employed and unemployed search strategies.

1.1 Related Literature

Our work is related to a growing literature that uses data from online job posting and search websites in order to study different aspects of job search. Matsuda, Ahmed, and Nomura (2019) show that employers prefer applicants who are whose qualifications align with the job specifications in Pakistan. Kudlyak, Lkhagvasuren, and Sysuyev (2013) study how job seekers direct their applications over the span of a job search. They find some evidence on the 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). Faberman and Kudlyak (2019) use online job board data to study the intensive margin of job search. Marinescu and Rathelot (2018) use information from `www.careerbuilder.com` and find that job seekers are less likely to apply to jobs that are farther away geographically. Banfi and Villena-Roldán (2019) and Banfi, Choi, and Villena-Roldán (2022) use data from `www.trabajando.com` to find substantial evidence of directed search and assortative matching, providing complementary evidence related to the selective component. Fluchtmann, Glenny, Harmon, and Maibom (2024) merge administrative data and online job board applications to study the dynamics of applied-for wages for the unemployed rather than search intensity as we do.

Our paper also contributes to a literature showing compliance to job requirements or characteristics in different settings. For instance, Brenčič and Pahor (2019) examine the upgraded skill requirement and worker compliance after a firm becomes exporter, and Clemens, Kahn, and Meer (2021) show the effect of minimum wage raises in the changes in educational requirements and workers' compliance. Fabel and Pascalau (2013) take another angle and explore the experience-education substitution from between insiders and outsiders of the firm. Fredriksson, Hensvik, and Skans (2018) show that mismatch of abilities defined with respect to the average of experienced workers, decay over tenure. Our work is complementary to theirs since we focus on the ex ante misalignment , potentially generated by job search patterns, instead of outcomes from realized matches. Therefore, we study the process leading to observed allocations.

Finally, our paper is also related to a strand of the literature comparing the job search behavior of employed and unemployed seekers. This body of work has that on-the-job search typically yield infrequent but beneficial transitions. Belzil (1996) finds this fact holds for older workers in Canada. In the same vein, Holzer (1987) reports higher transition rates for unemployed individuals, albeit often into lower-wage positions. Furthermore, Longhi and Taylor (2011, 2013) indicate that employed job seekers in the UK exhibit greater selectivity and a higher propensity to transition to high-wage occupations. Our findings are more clearly related with those of Faberman, Mueller, Şahin, and Topa (2022) because they use retrospective questions in the NY Fed survey to

elicit search behavior. They find that employed jobseekers are more effective in obtaining more and high-wage job offers compared to the unemployed counterparts. We contribute to this literature by providing a systematic way to use online job search data to obtain a deeper analysis of the search process along several potentially misaligned dimensions, not just wages.

2 The data

We use data from `www.trabajando.com` (henceforth the website), a job search engine operating in Chile, covering a sample of job postings and job seekers between January 1st 2008 and December 24th, 2016. The raw information in the dataset contains more than 14 million single applications, from around 1.5 million job seekers to around 270 thousand job ads.

Our dataset has detailed information on both applicants and recruiters. First, we observe entire histories of applications from job seekers and dates of ad postings (and repostings) for recruiters. Second, we have detailed information for both sides of the market. 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, years of education, college major, and name of the granting institution of the major, for individuals with post-high school education. We have codes for the occupational area of the current or last job of individuals: We observe a one-digit classification, created by the website administrators,[1] and information on individual's salary and both their starting and ending dates.

In terms of the website's platform, job seekers can use the site for free, while firms are charged for posting ads. Job advertisements are posted for a minimum of 60 days, but firms can pay additional fees to extend this term.

For each posting, we observe its required level of experience (in years), required college major (if any), indicators on required skills (specific, computing knowledge and/or "other"), how many positions must be filled, the same occupational code applied to workers, geographic information ("región" only) and some limited information on the firm offering the job: its size (number of employees in brackets) and industry (1 digit code).[2] Educational categories are *primary* (one to eight years of schooling), *high school* (completed high school diploma, 12 years), *technical tertiary education* (professional training after high school, usually 2-4 years), *college* (completed university degree, usually 5-6 years) and *post-graduate* (any schooling higher than

[1] Categories are: Business administration, Agriculture, Art and Arquitecture, Basic Sciences, Social Sciences, Law, Education, Humanities, Health, Technology, Other and N/A.

[2] We observe an industry classification created by the website administrators that does not match formal taxonomies such as NAICS or ISIC.

a college degree).

A novel feature of the dataset, compared to the rest of the literature, is that the website asks job seekers to record their expected salary, which they can then choose to show or hide from prospective employers. Recruiters are also asked to record the expected pay for the job posting and are given the same choice as to whether to make this information visible to the applicants. Naturally, the reliability of wage information could be questionable, which will ultimately be hidden from the other side of the market. Banfi and Villena-Roldán (2019) address the potential issue of "nonsensical" wage information in job ads by comparing the sample of explicit vs implicit (job ads without any salary information) postings by firms and find that observable characteristics predict fairly well implicit wages and vice versa. Moreover, even if employers choose to hide wage offers, they are used in filters of the website for applicant search. Hence, employers are likely to report accurately even if their wage offers are not shown because misreporting may generate potential bad matches. On the other hand, a major caveat of our dataset is the absence of information on activities performed outside the website, such as individuals seeking jobs through other means and, more importantly, the outcomes of job applications.

For the remainder of the paper, we restrict our sample to individuals working under full-time contracts and those who are unemployed. We further restrict our sample to individuals aged 23 to 60. We discard individuals reporting desired net wages above 5 million pesos.[3] This amounts to approximately 8,347 USD per month[4], which is higher than the 99th percentile of the Chilean wage distribution, according to the 2013 CASEN survey.[5] We also discard individuals who desire net wages below 159 thousand pesos (around 350 USD) a month (the legal minimum wage at the start of our considered sample). Consequently, we also restrict job postings to those offering monthly salaries within those bounds.

Our unit of analysis are individual *applications*. We restrict our sample to active individuals and job postings during the sample period: those that made/received at least one application. While we observe long histories of job search for a significant fraction of workers (some workers have used the website for several years), we consider only applications pertaining to their last job search "spell", which we define as the time window between the last modification/creation of their online curriculum vitae (cv) on the website and the time of their last submitted application or the one year mark, whichever happens first. Since individuals maintain information

[3] In the Chilean labor market, wages are usually expressed in a monthly rate net of taxes and mandatory contributions to health (7% of monthly wage), to fully funded private pension system (10%), disability insurance (1.2%), and mandatory contributions to unemployment accounts (0.6%)

[4] Using the average nominal exchange rate between 2013-16 at the Central Bank of Chile: https://si3.bcentral.cl/Siete/en.

[5] CASEN stands for "Caracterización Socio Económica" (Social and Economic Characterization), and aims to capture a representative picture of Chilean households. For data and information in Spanish, visit http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen.

about their last job in their online profile, as well as contact information and salary expectations, we assume that any modification of this information is done primarily when individuals who are currently working or who have already used the website in the past are ready to search in the labor market again. We cannot infer any labor transitions based on application behavior because employed individuals may keep searching for jobs, and unemployed individuals may search outside of the website. We further drop individuals who apply to more than the 99-th percentile of job applicants in terms of number of submitted applications in the defined window.

Table 1 shows descriptive statistics for the job seekers in our sample. From the table, we observe that the average age is 33.5 and that job seekers are comprised mostly single males, with 59.71% being unemployed ($128,482$ unemployed seekers from a total of $215,169$ individuals). Average experience hovers around eight years. Job seekers in our sample are more educated than the average in Chile, with 41.84% of them having a college degree, compared to 25% for the rest of the country in the comparable age group (30 to 44 years of age), according to the 2013 CASEN survey. There is also a big discrepancy by labor force status: unemployed seekers are significantly less educated on the website.

From the table, we can also observe that most job seekers claim occupations related to management (around 20%) and technology (around 25%) and that average expected wages are approximately (in thousands) CLP$ $1,087$ and CLP$ 592 for employed and unemployed seekers, respectively. For comparison, the 2013-16 average minimum monthly salary in Chile was around CLP $ 226 thousand.

In terms of search activity, the average search spell amounts to around five weeks (37.49 days). The amount of time spent searching for a job is higher for those employed than for the unemployed (33.78 vs. 42.99 days). In terms of applications, in the table we show medians and means to display the skewed distribution of applications, with the majority sending few applications (total median of 3) while a few seekers concentrate large numbers, making the mean significantly higher (7.81 overall).

3 Application probabilities and job seeker preferences

In this section, we develop key ideas to determine which set of job ads is relevant for each individual in our sample. This is the first step towards empirically analyze how the match between attributes of job seekers and requirements of job ads translate into application decisions. The primary challenge is that we observe only realized applications, lacking information on the broader set of job ads actively considered by individuals but ultimately not pursued. Specifically, we do not observe the number of searches or *clicks* on job postings by individuals. Out of thousands available jobs for applicants, we only observe those that individuals choose to

Table 1: Characteristics of Job Seekers

	Employed	Unemployed	Total
Demographics (%)			
Male	62.03	53.97	57.21
Married	33.80	27.50	30.03
Demographics (Avg)			
Age	33.77	33.25	33.46
Experience (years)	8.28	7.64	7.90
Wages (thousand CLP)	1,087	592	792
Tenure (weeks)	179.29	–	179.29
Unemployment duration (weeks)	–	60.17	60.17
Education level (%)			
Primary (1-8 years)	0.12	0.25	0.2
High School	17.94	36.89	29.25
Technical Tertiary	26.56	28.82	27.91
College	54.22	33.48	41.84
Post-graduate	1.17	0.55	0.8
Occupation (%)			
Management	23.5	17.85	20.12
Technology	31.59	21.21	25.39
Not declared	20.29	42.54	33.57
Rest	24.62	18.4	20.92
Search Activity			
Days searching on website	42.99	33.78	37.49
Number of applications (median/mean)	4 / 9.19	3 / 6.87	3 / 7.81
Observations	86,687	128,482	215,169

apply to, not those that are observed but then discarded by seekers. This problem of "consideration sets", i.e. the domain of options effectively under evaluation by economic agents, is similar to the one addressed by marketing and industrial organization literatures (Van Nierop, Bronnenberg, Paap, Wedel, and Franses, 2010; Abaluck and Adams-Prassl, 2021).

3.1 Different approaches to local labor markets

While other papers have typically understood a local labor market as non-overlapping cells defined by occupation and location (Şahin, Song, Topa, and Violante, 2014; Lamadon, Mogstad, and Setzler, 2022; Azar and Marinescu, 2024, , among many others) or overlapping locations (Manning and Petrongolo, 2017), we advocate an approach based on realized applications as revealed preferences of workers. Using this information, we construct individual consideration sets using coincidental choices made by other applicants. Since workers apply to jobs (potentially) considering a large number of characteristics, many of which we observe, our approach takes an agnostic view as to the way seekers process information. Thus, our methodology incor-

porates geographic and occupational/major dimensions through network weighting, indirectly accounting for applicant behavior driven by these factors. However, as detailed in Table A5 in the appendix, while occupation and location are influential, imposing strict occupational/regional boundaries on job search is inconsistent with empirical observations. Consideration of jobs outside of fixed cells is common in our sample: nearly half of applications are submitted for positions outside the applicant's region or occupation/major category. Kambourov and Manovskii (2009); Carrillo-Tudela and Visschers (2023); Jarosch, Nimczik, and Sorkin (2024) also find plenty of transitions across strict cell markets in other countries.

Another potential approach is the fully unrestricted, or "agnostic" view, which would allow all time-feasible job ads into the consideration set of applicants, that is, the cross-product of all job seekers and all job ads in our sample, i.e. the exploded dataset. This is hardly realistic, as a typical job seeker may encounter more than 20,000 available job ads to screen and choose from, implying an unrealistic effort for workers. Moreover, the implied computational burden is substantial. Given our sample constraints, we have upwards of 200,000 workers who could potentially apply to more than 20,000 job ads, resulting in approximately 4 billion individual-job ad combinations.

Besides these considerations, we show that introducing some relatively straightforward structure of preferences and choices, renders the agnostic view biased. We demonstrate this by building upon the demand model of Tenn and Yun (2008), in which products are not available at every store, akin to the idea that not all jobs are truly present in every individual's consideration set. The model can accommodate a multiplicity of observed factors to explain observed choices within a multinomial logit framework. Crucially, ignoring the availability heterogeneity in Tenn and Yun (2008) —for example, assuming that all products are available in every retailer— leads to a significant estimation bias if the likelihood of a product being in the consideration set and the likelihood of an actual purchase decision are correlated with the same characteristics. This is exactly the case for online job search: factors such as educational level, experience, major, and wage are likely used as pre-screening filter variables before workers actually apply (indeed, the website has several of these features in its search engine). Moreover, given the importance of directed search, where job ad characteristics influence applications (Banfi and Villena-Roldán, 2019; Marinescu and Wolthoff, 2020; Banfi, Choi, and Villena-Roldán, 2022), the presence of irrelevant job ads would lead to an inconsistent estimation, akin to that arising in an omitted variable problem. In appendix A.1 we elaborate on these arguments in greater detail.

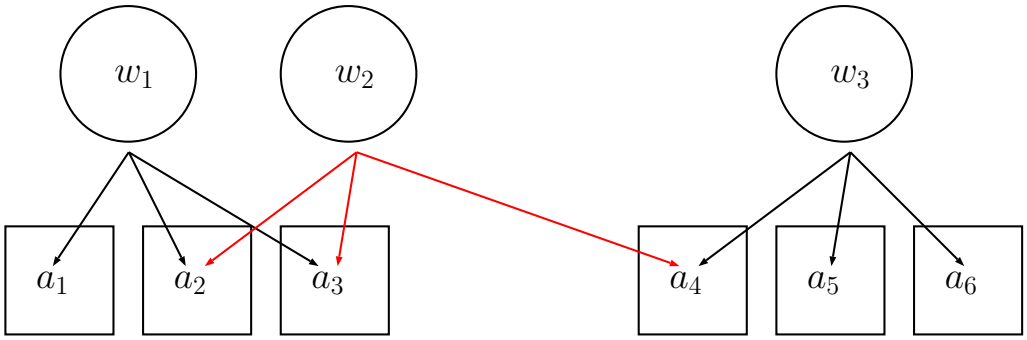

Figure 1: Example of a network formed by workers $\{w_1, w_2, w_3\}$. Worker w_1 is linked to worker w_2 by common applications to ads a_2 and a_3 but is not linked with w_3 in the network of degree 1. All workers are linked in the network of degree 2.

3.2 Network consideration sets

Our approach builds upon the intuition that job seekers who apply to the same jobs are likely to have considered jobs that their co-applicants applied to. To formalize this notion, we use the network formed by job seekers to determine which job postings are relevant to them. Each individual and each job ad represents a node in a bipartite network in which applications link workers w to contemporaneous ads a, i.e. the ad a must be available during the search window of w. In this way, the network connects two workers through a common link if they *have applied to the same job posting*. For each job seeker w, we define the set of relevant job postings $\mathcal{A}_w^1$ as the union of all job postings applied to by the set of all job seekers linked to w. Since we only consider their immediate links for each individual (1 degree of separation), we define this as a network of degree 1. Our approach has some similarities to the literature of community detection social networks (Karrer and Newman, 2011) and its application to local labor markets in Nimczik (2023). As these authors do, we use agents' decisions to back up a labor market structure that is consistent with those choices instead of relying on predefined characteristics defining labor markets.

Following this logic, the network of degree 0 is the original set of job ads applied to by individual w, denoted analogously as $\mathcal{A}_w^0$. On the other hand, a network of degree 2 is defined as the network that considers both job seekers linked directly to w in addition to those who are linked to the connections of w (job seekers have 2 degrees of separation), giving rise to the set $\mathcal{A}_w^2$. We can continue with this logic iteratively until we form the set $\mathcal{A}_w^\infty$, which is the cross-product of each job seeker w and all job postings a as long as they are connected somehow through the network.[6]

Figure 1 shows an example of the network algorithm and the resulting datasets. In the figure, there are three

[6]The set $\mathcal{A}_w^\infty$ and the exploded dataset differ if there are isolated pairs or groups of individuals who are not connected to the rest of the applicants through any ad.

workers, $\{w_1, w_2, w_3\}$ and six job postings, $\{a_1, a_2, a_3, a_4, a_5, a_6\}$. Consider worker w_1. She applies to three jobs, thus $\mathcal{A}^0_{w_1} = \{a_1, a_2, a_3\}$ and is linked to w_2 through applications to $\{a_2, a_3\}$. Since w_2 also applies to job position a_4, if we consider networks of degree 1, a_4 would be included in the set of relevant ads for the w_1.

Again, considering the first worker, we have $\mathcal{A}^0_{w_1} = \{a_1, a_2, a_3\}$, and as discussed above, $\mathcal{A}^1_{w_1} = \{a_1, a_2, a_3, a_4\}$. Given that w_1 and w_2 are linked and that w_2 is linked with w_3, the relevant job ads for w_1, given a network of degree 2, is $\mathcal{A}^2_{w_1} = \{a_1, a_2, a_3, a_4, a_5, a_6\}$. In our simple example, the network of degree 2 is already the "exploded" network (the cross-product of all ads and all workers).

The formal definition of a one-degree-of-separation ad set for a worker w is

$$\mathcal{A}^1_w = \bigcup_{v:\mathcal{A}^0_w \cap \mathcal{A}^0_v \neq \emptyset} (\mathcal{A}^0_w \cup \mathcal{A}^0_v)$$

which can be generalized for other degrees of separation.[7] While we could construct consideration sets using an arbitrarily number of separation degrees, s, it becomes computationally unfeasible soon. In what follows, we will concentrate on networks of degree 1 only.

Table 2: Number of relevant ads (a) per worker (w)

	Potential ads for a worker		
	All	U	E
percentile 10	2	2	2
percentile 50	16	16	19
percentile 90	96	104	87
mean	38.5	40.7	36.8
standard deviation	68.1	73.8	57.1
mean applications (%)	22.3	23.2	20.9

Notes: The table shows the number of relevant job postings per job seeker given a network of degree 1 (see main text). Statistics separated by labor force status of job seeker (U = unemployed, E = employed).

In table 3.2, we present information on the resulting number of relevant job postings per worker and workers per job posting, given a network of degree 1. The median number of relevant job postings (a) is 16 per job seeker, with employed seekers being related to more posts (19) than those unemployed (16). The number of potential ads exhibits quite a bit of variation, going from 2 (tenth percentile of distribution) to 104 and 89 for the unemployed and employed, respectively (ninetieth percentile). Given the sets of related job ads,

[7] The generalization follows a recursive definition

$$\mathcal{A}^s_w = \bigcup_{v:\mathcal{A}^{s-1}_w \cap \mathcal{A}^0_v \neq \emptyset} (\mathcal{A}^{s-1}_w \cup \mathcal{A}^0_v)$$

which depends on $\mathcal{A}^0_w$ and the definition of $\mathcal{A}^1_w$.

mean application rates,[8] are 22.3% for the entire sample, with unemployed seekers applying to 23.2%, while employed ones do so for 20.9% of their relevant ads.

3.3 Similarity / Proximity metrics

The relevance of ads within a consideration set should vary. We posit that greater application overlap between workers implies higher similarity. Consequently, for a worker w and a non-applied ad a, proximity should increase with the similarity between w and another worker v who applied to a and is linked to w. For instance, in figure 1, w_1 and w_2's shared applications suggest similar preferences/qualifications, increasing the likelihood of w_2 considering w_1's applied ads. Conversely, w_2 is less likely to consider $\{a_5, a_6\}$, given the fewer shared applications with w_3.

We formalize this ideas using the Jaccard (1901) metric to quantify set similarity:

$$b(w, v) \equiv \frac{\left|\mathcal{A}_w^0 \cap \mathcal{A}_v^0\right|}{\left|\mathcal{A}_w^0 \cup \mathcal{A}_v^0\right|} \tag{1}$$

where $|S|$ denotes the cardinality of set S.

The similarity metric b allows us to define $q(w, a)$, a proximity metric between worker w and ad a, loosely representing the probability of w considering a. If w and a are linked solely through worker v, then $q(w, a) = b(w, v)$ is a straightforward choice. With multiple linking workers, we define $q(w, a)$ to be the *max-proximity* as the maximum similarity b across all paths connecting w and a:

$$q(w, a) = \max_{v: a \in \mathcal{A}_v^0} \{b(w, v)\} \tag{2}$$

This metric assigns a weight to ad a for worker w, determined by the highest similarity between the set of ads chosen by w and the set of applications done by any other worker applying to a. The resulting max-proximity satisfies properties akin to a probability weight.

1. $q(w, a) = 1$ if and only if $a \in \mathcal{A}_w^0$ (w applies to a);
2. $q(w, a) \in [0, 1)$ if and only if $a \notin \mathcal{A}_w^0$ (w does not apply to a);
3. $q(w, a) = 0$ if and only if $a \notin \mathcal{A}_w^1$ (a is not in w's choice set).

Several criteria justify the max-proximity:

[8] Defined as the number of effective applications to total ads for worker w: $\frac{|\mathcal{A}_w^0|}{|\mathcal{A}_w^1|}$

1. *Reduced gap across separation degrees:* As shown in Proposition 1 (appendix A.2), similarity $b^s(w,v)$ weakly increases with separation degree s. Max-proximity mitigates discrepancies from varying s.

2. *Robustness to network variations:* Maximizing similarity between workers sharing ad a makes our measure robust to minor network changes unless a larger proximity emerges.

3. *Shortest path interpretation:* With random ad choices, $b(w,v)$ represents the probability of workers w and v choosing a common ad.[9] Hence, $q(w,a)$ selects the most likely path linking a and w.

While the max-proximity is not the only way to construct a metric, the average-proximity $\frac{\sum_{v:a\in\mathcal{A}_v^0} b(w,v)}{\sum_{v:a\in\mathcal{A}_v^0} 1}$ yields similar empirical results (see results in section 5, table 4), suggesting that the exclusion of potential applications via consideration sets is more critical than the precise weighting by q. In the appendix A.3 we develop a simple example and compute maximum and average similarity metrics as an example.

4 Estimation of the application equation

For the constructed dataset, we estimate a linear regression of the form

$$y_{wa} = X_{wa}\beta + \sum_{k}\sum_{p=1}^{\bar{P}}\{\gamma_{kp}(z_{k,wa})^p\} + \gamma_j z_{j,wa} + \epsilon_{wa} \tag{3}$$

where y_{wa} is a dummy variable that takes the value of one if a job seeker w applies to posting a and zero otherwise. In X_{wa}, we include a linear trend and monthly dummies to control for secular trends and seasonal patterns in website usage.[10] We also control for observed job and worker characteristics. The list of variables for the job includes firm size, dummies for firm industry, specific job requirements (computer knowledge or some other form of specific knowledge), and controls for specific job characteristics: type of contract (full or part time), number of vacancies needed to be filled, and controls for job title relevant words following Marinescu and Wolthoff (2020) and Banfi and Villena-Roldán (2019). For individuals, we control binary variables male, marriage, and an interaction between them. We also include quintic polynomials for the age of the job seeker, the amount of time (measured in weeks) in either the current job (for those employed), or unemployment (for unemployed seekers) and finally, the number of total related jobs to the worker within a network of degree of separation 1. We handle some missing values following guidelines discussed in the appendix A.4.

[9] The Jaccard metric b equals the density of the local network defined between two workers connected via coincident applications, i.e. the share of actual links out of total potential links. When computing the $q(w,a)$ metric as the maximum of Jaccard metrics b, we simply choose the most likely path linking the ad a and w under the previous assumption.

[10] We do not explicitly include time subindices since application time is exactly determined by a pair (w,a). This occurs because worker w can only apply to ad a just once. As an example, a June month dummy J_{wa} takes a value of 1 if w applies to a in June.

For both seekers and ads, we include a variable indicating whether the wage expectation (for seekers) or the wage expected to be paid (for jobs) is made explicit or not. To control for business cycle conditions, we consider the unemployment rates of the applicant's region during the month in which the application took place[11]. We also include a quadratic term of the regional unemployment rate to capture potential non-linear effects, which follow Hazell and Taska (2024)[12]. The effects of these characteristics impact the level of the probability of application and therefore are related to an *average component* of the application process that has been more profusely studied in the literature (DeLoach and Kurt, 2013; Gomme and Lkhagvasuren, 2015; Baker and Fradkin, 2017; Ahn and Shao, 2017; Leyva, 2018; Mukoyama, Patterson, and Şahin, 2018; Faberman and Kudlyak, 2019; Bransch, 2021).

As mentioned in the introduction, our main focus is measuring the *selective component* of job search: to that end, we include a set of controls for the *misalignment* or gap (which we denote by z) between characteristics required by job positions versus the characteristics of the job seeker. For continuous variables, which we index by k, we define z_k as the simple difference between the value of the characteristic required by the position and the value of the characteristic possessed by the job seeker. We do this for years of education, years of experience, and log wages. Notice that this definition allows for *negative* values, which is the case when the value in the job ad is *lower* than the value of the characteristic for the worker. For regional distance, we compute misalignment as kilometers between regional capital cities.[13] For occupations, the variable z_j is defined as a dummy that takes the value of 1 when the category in the job posting is different from the occupation of the current/last job of the worker (when the individual is searching on the job/from unemployment) and 0 when they are the same.

4.1 Worker-ad characteristic gaps within consideration sets

In table 3 we show statistics regarding misalignment for different sets of job ads: (i) those that the worker applied to; (ii) ads attached to the worker using our network algorithm; and (iii) a set of randomly selected ads assigned to each worker. For the latter, we select 330 random job ads in the entire dataset, number which represents the 99-th percentile in terms of the distribution of network assigned job ads.[14] From the table we observe that the level and dispersion of misalignment is generally higher for randomly assigned job ads

[11] More details the appendix A.4

[12] However, it's true that the notion of cyclical behavior using regional variation does not exactly fit the theoretical counterpart, as shown by Kuhn, Manovskii, and Qiu (2021).

[13] See in the appendix A.4 our strategy to handle observations with missing region.

[14] We pick this high number to minimize the potential effect of dissimilar network sizes for different workers: e.g., certain types of workers may be predisposed to apply to more jobs, making their networks larger by construction.

Table 3: Difference between ads and workers

	Mean	SD
log wages		
Applied ads	-0.0524	0.4821
Relevant ads (Network)	-0.0461	0.4819
Relevant ads (Random)	-0.0593	0.6894
Education		
Applied ads	-0.1623	0.8010
Relevant ads (Network)	-0.1615	0.7445
Relevant ads (Random)	-0.2235	0.8655
Experience		
Applied ads	-5.7141	6.5949
Relevant ads (Network)	-5.7051	6.6680
Relevant ads (Random)	-5.6981	6.7677
Regional distance		
Applied ads	159.15	387.80
Relevant ads (Network)	166.58	366.52
Relevant ads (Random)	216.58	404.47
Different Occupation		
Applied ads	0.3349	0.4393
Relevant ads (Network)	0.4120	0.3165
Relevant ads (Random)	0.5247	0.1362

Notes: The table shows misalignment measures for different sets of job ads (see main text).

than those arising from the network formation algorithm discussed above: the difference in wages, education level, and regional distance are all higher when we look at the set of random ads. In terms of occupation, the likelihood that they represent a different occupation is also significantly higher (52 vs 41 per cent). The only exception is years of experience, which exhibit similarities between the random and network ads. We relate these statistics in table 3 to our discussion of consideration sets in section 3.1 because sizable discrepancies in observable characteristics between randomly picked ads and consideration set ads is exactly the setup in which a biased estimation occurs, according to Tenn and Yun (2008) model. This reinforces the need of some approach defining local labor markets, instead of assuming that job seekers could apply to any job.

In equation (3), for each of the continuous dimensions k we include in the regression a polynomial of order $P = 5$ to assess whether non-linearities exist in the effect of these *misalignments* on application decisions. In this way, we capture if *over*-qualified ($z_k < 0$) jobseekers behave differently from *under*-qualified ($z_k > 0$) ones. We estimate equation (3), separating our sample between the employed and unemployed[15] to assess whether on-the-job search behavior differs from unemployed search behavior. We also consider interaction effects between different *misalignment* levels, the penultimate term in (3).

[15] We construct consideration sets before splitting our sample, allowing us to take both employed and unemployed applications to define local labor markets.

4.2 Weighting observations

We argue in section 3.3 that weighting observations using the proximity between workers and ads, $q(w, a)$, is crucial for obtaining unbiased estimates, as outlined by Tenn and Yun (2008) and detailed in appendix A.1.

However, this is insufficient. We must also account for changes in the market composition of applicants and job ads. Since we exclude applications before the last CV update, our sample over-represents individuals from later periods. Notably, about a quarter of our data (approximately 2 million observations) are applications from 2016 Q3. Balancing the composition is essential to control for cyclical search behavior and compositional effects, especially given the website's increasing penetration in the Chilean labor market.

To address this issue, we use the reweighting technique of DiNardo, Fortin, and Lemieux (1996). We model the probability of an application occurring in 2016 Q3 as a function of applicant and job ad observables using a probit model. For categorical dummies, we drop those with an average below 0.2 or above 0.8 in 2016 Q3 or other quarters to avoid extreme predicted probabilities. We then compute predicted probabilities $\widehat{p}(w, a)$. Our final weight for worker-ad observations is:

$$\varphi(w, a) = q(w, a)\frac{\widehat{p}(w, a)}{1 - \widehat{p}(w, a)}$$

for applications within the common support of observations from 2016 Q3 and other quarters.

For consistent estimation, we require zero covariance between the network weight $q(w, a)$ and the error term ϵ in equation (3), conditional on the observables of the pair (w, a), i.e., $cov(\sqrt{q(w, a)}, \epsilon | X, \{z_{k,wa}\}_k, z_j) = 0$.[16] As shown in section 3.3, $q(w, a)$ reflects choices by anonymous co-applicants of w, generally unaware of worker w's existence and unobserved characteristics. This aligns with online job board operation: applicants typically ignore who else is applying. However, the residual ϵ may correlate with unobserved ad features. Given our extensive controls for ad characteristics, this is likely only if applicant w and others have private information about the job not in the ad.

5 Results

Table 4 shows coefficients multiplied by 100 from the estimating equation (3) using ordinary least squares. We report estimates by employment status and whether we perform weights through maximum or average proximity between a worker and the ads in her network.[17]

[16] A formal description of this condition is in appendix A.5.

[17] For further results with alternative weights, see the Appendix.

5.1 Results on the effects of traits of Applicants and ads

Table 4: Average component coefficients by labor status

VARIABLES	(1) Employed weights^A	(2) Employed weights^B	(3) Unemployed weights^A	(4) Unemployed weights^B
Married	-2.769***	-2.774***	-0.150	-0.113
	(0.931)	(0.936)	(0.498)	(0.507)
Male	1.609***	1.633***	2.324***	2.425***
	(0.066)	(0.067)	(0.050)	(0.051)
Explicit wage (w)	0.261***	0.266***	0.808***	0.833***
	(0.057)	(0.058)	(0.047)	(0.048)
Explicit wage (a)	-2.724***	-2.977***	-0.955***	-1.045***
	(0.088)	(0.090)	(0.067)	(0.068)
No. of Vacancies (a)	-0.007	-0.003	0.045***	0.059***
	(0.011)	(0.011)	(0.006)	(0.006)
Ad duration (weeks)	-0.001	0.000	-0.087***	-0.088***
	(0.002)	(0.002)	(0.001)	(0.001)
Observations	2,124,244	2,124,244	3,184,675	3,184,675
R-squared	0.12	0.12	0.11	0.11
Mean app prob	22.91	24.18	26.87	24.82

Notes: Regression coefficients from a linear regression on application decisions. Dependent variable is y_{wa}, a dummy for the existence of a job application. Each regression controls also for polynomials and interactions in *misalignment* as well as age of the worker, firm size, contract type, dummies for different types of requirements of the job and characteristics of the firm. Column *weights*^{A} denotes results under weights constructed using the max proximity between workers and ads; *weights*^{B} is an alternative average proximity. See main text for further details. Standard errors in parentheses. One, two, and three asterisks indicate significance at 10%,5%, and 1%, respectively.

The first point to notice is that unemployed job seekers apply more frequently than employed ones to the ads in their consideration sets. Among the employed, married individuals apply less than their non-married counterparts, while there is a non significant gap for the unemployed. Male job seekers, especially unemployed ones, apply more often to ads, keeping other applicant and ad characteristics constant.

Job seekers who explicitly state their wage expectations apply more frequently than those who do not, particularly among the unemployed. Conversely, job ads that include an explicit wage tend to receive fewer applications on average, although this effect is less pronounced for unemployed individuals. This observation aligns with findings in Banfi and Villena-Roldán (2019), which suggest that ads with hidden wages attract more applicants due to the perceived possibility of wage flexibility or negotiation, as proposed by the Michelacci and Suarez (2006) model. Furthermore, unemployed individuals significantly increase their application probability

by 0.045 percentage points for each additional vacancy, while employed individuals show no significant change in application behavior. The limited positive response to a slightly higher likelihood of receiving an offer points to a substantial role for employer-side selection, potentially through non-sequential employer search (van Ours and Ridder, 1992; van Ommeren and Russo, 2013) or signaling of less favorable job conditions.

The effect of the perceived "age" of the job ad has a negative effect for the unemployed, who dislike job ads that are older (in weeks). The negative effect for the unemployed can be related to stock-flow matching behavior[18]: new job seekers in the website (the flow) apply to the stock of job ads. When time passes, the inflow of job seekers becomes part of the stock of individuals, who then try to match with the new flow of job positions, as suggested by evidence in Gregg and Petrongolo (2005) and Coles and Petrongolo (2008). Our results for the unemployed are also consistent with applicants reacting to "phantom" ads, which may be filled positions by the time of the potential application, as in Albrecht, Decreuse, and Vroman (2023) and Chéron and Decreuse (2016). The evidence reported by Davis and Samaniego de la Parra (2024) is also qualitatively consistent with our findings. The effects are not significant for the employed, which suggests a different pattern of search for this group in this website.

5.2 Selective Component: Misalignment and applications.

Next we present the effect of *misalignment* in continuous dimensions (education, experience, log wages, and distance), which we claim represents how *selective* workers are in terms of complying with quantifiable job ad requirements and how sensitive they are to not fill them.

In figure 2 we present graphically results of the effect of misalignment in years of education, years of experience, log wages, and regional distance (in hundred of kms) on application decisions. The figure shows predicted application probabilities ($\hat{y}_{wa}$ from the estimates of equation 3), when a particular continuous dimension misalignment (z_k) varies, keeping all other observables at their sample mean, including the misalignment in other dimensions. Given that each misalignment dimension enters the equation as a fifth-order polynomial and that there are interactions between them, the computed effect is potentially highly non-linear and depends on which value the other control variables take. The considered range for z_k is bounded by its 1st and 99th percentiles of the variables and all figures display 95% confidence bands.[19]

As seen in the figure, job seekers in both labor market states tend to align themselves with the advertised requirements of job postings. This is represented by an inverted U-shaped relationship between *misalignment*

[18] References are Taylor (1995); Coles and Muthoo (1998); Coles and Smith (1998); Ebrahimy and Shimer (2010)

[19] Figure A3 in the appendix shows the same exercise but presenting relative application probabilities rather than levels.

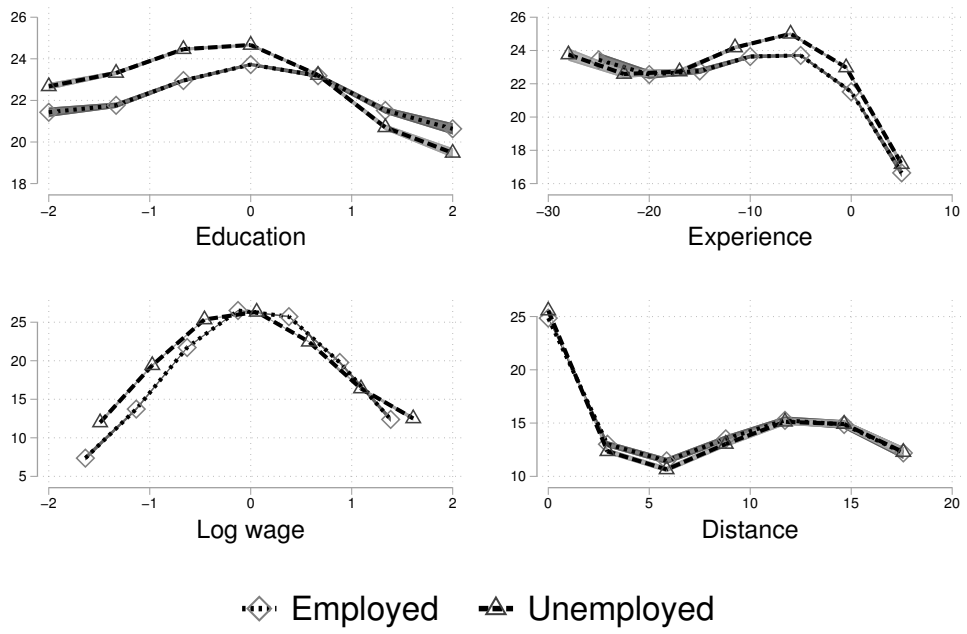

Figure 2: Predicted application probabilities, given results from eq. (3) and different levels of misalignment in the selected variable x (see main text for details). The rest of regressors are at their sample means.

and application probability (all else constant) for education, experience, log wages, and by a mostly decreasing line in the case regional distance. The figures also show that all estimates are sharp, given the narrow confidence intervals.

Education misalignment : In the upper-left panel of figure 2, the application probability for both employed and unemployed peaks at zero, e.g., an exact alignment between required and realized years of education. Nevertheless, the shapes of employed and unemployed are asymmetric. The application probability is larger for the unemployed when the applicant is overqualified, whereas the pattern reverses when the applicant is underqualified. Therefore, employed seekers seem less reluctant to apply to jobs for which they are underqualified in terms of education. In other words, employed workers are more ambitious or daring to take the next rung of the job ladder, assuming that jobs requiring more education are better.

Experience misalignment : For the unemployed, the experience dimension curves (northeast panel of the same figure) peak around -7 and show a steeper decline for higher values of misalignment in that dimension. This means that job seekers tend to have more than seven years of experience than the minimum required by positions and do not refrain from applying if they are even more overqualified in experience. The main reason for the average misalignment in this dimension is that most of our sample consists of individuals with

a significant number of years of experience while experience requirements in job ads often represents a lower bound. The application probability curve for the employed peaks a bit to the left of the one for the unemployed, and the gap between the two groups becomes wider for experience gaps between -15 and 0, suggesting that the unemployed are slightly more prone to apply to jobs for which they are overqualified in terms of experience.

Offered-Expected Wage Gap: The plot in the lower-left panel reveals that differences in log wages greatly affect application probabilities: the application probabilities for the unemployed fluctuate between 12% and 25%, while for employed seekers, the range is wider, from around 7% to 25%. Given that our estimates control for all other observables across job positions and job seekers and that the regression controls for interactions, we can interpret the misalignment in $\log$-wages as a gap in job and worker unobserved productivities. Controlling for all observables, higher-paying jobs and job seekers with higher earnings expectations must have higher skill levels on average, and vice versa. These interpretations align with our findings on high positive assortative matching at the application stage (Banfi, Choi, and Villena-Roldán, 2022). Overall, the unemployed application curve lies above its employed counterpart for negative wage gaps. Unemployed individuals are more likely to apply to jobs for which they are overqualified in terms of productivity. Conversely, for jobs where applicants are underqualified (to the right of the peaks), employed applicants apply more often. These patterns suggest that on-the-job searchers are more daring than their unemployed counterparts, probably due to the former's better outside options. In contrast, the unemployed more often apply to jobs paying below their wage expectations due to the urgency of finding employment. Workers applying less to jobs paying well above their expectations, despite the utility gain from higher wages, suggests strategic behavior. A natural explanation is that workers weigh the desirability of the high wage against a perceived lower probability of being hired.

Distance: The lower-right panel depicts the predicted probability as a function of the distance between the regional capital of the applicant and the regional capital of the job in hundreds of kilometers. For ads located relatively close to the applicants, the likelihood of an application decreases quite quickly: from 25% at zero-distance to nearly 10% for a 600 kilometers distance. For higher distances there is some increase, which may be related to the geography of the country: because approximately 77% of the population lives less than 600 kilometers away from Santiago, which in itself represents 40%, individuals from north and south extremes of the country may have internalized moving to the central part of the country for better labor outcomes. The employed-unemployed gap is only noticeable in mid-distance applications. Marinescu and Rathelot (2018) and Manning and Petrongolo (2017) estimates imply a much larger drop in the likelihood of applying to jobs as distance increments, although our estimates are not directly comparable because we measure an intention

rather than an effective reallocation and control for a substantially richer set of variables.

5.3 Other results

Our analysis also reveals how application probabilities vary with age, search duration, and business cycle conditions. We find that unemployed individuals' application probabilities exhibit a non-monotonic relationship with age, and a decreasing relationship with search duration, consistent with findings in Mukoyama, Patterson, and Şahin (2018), Faberman and Kudlyak (2019) and DellaVigna, Heining, Schmieder, and Trenkle (2021). Employed jobseekers show a flatter relationship with search duration, potentially due to offsetting factors like match-specific human capital and increased outside options, as suggested by Jovanovic (1979), Li and Weng (2017) Pissarides and Wadsworth (1994); Fujita (2012) and Menzio, Telyukova, and Visschers (2016). Application probabilities also display a decreasing pattern with the regional unemployment rate, aligning with DeLoach and Kurt (2013) and Gomme and Lkhagvasuren (2015), though some studies show different patterns. While these relationships have been explored in previous literature, accounting for them in our baseline estimation is crucial for obtaining consistent estimates of the misalignment effects, which are the main focus of our study. Appendix A.7 explain the results in more detail.

5.4 Varying weights

We further assess our methodological assumptions in appendix A.8. Re-estimating our model using only DiNardo, Fortin, and Lemieux (1996) (DFL) weights within the degree-1 network consideration set (Table A7, columns 2 & 5) yields estimates with generally lower absolute magnitudes compared to our baseline (which uses both DFL and proximity $q(w, a)$ weights). This suggests the proximity metric $q(w, a)$ captures relevant application likelihood information beyond aggregate composition. As average- and max-proximity metrics yield similar baseline results (Table 4), we infer that continuous proximity weighting is important; just inclusion/exclusion in the network set alone seems insufficient for reliable estimates. The lower average application rate in the raw network set compared to the proximity-weighted average further indicates $q(w, a)$ upweights more relevant ads. This aligns with Tenn and Yun (2008)'s argument that bias arises if irrelevant ads are included when factors driving considerations and applications are correlated.

Table A7 also shows results are sensitive to the consideration set definition. Using randomly selected non-applied ads (columns 3 & 6) instead of the network-defined set yields estimates differing considerably from both the baseline and the specifications using consideration sets and DFL. This highlights that all components—the consideration set definition (network vs. random), the proximity weights within the network set,

and the compositional adjustment (DFL)—are important factors influencing the final estimates.

Comparing misalignment curves under alternative specifications (Figures A3 and A4, normalized by mean application probability) to the baseline (Figure 2) reveals loose similarities but significant differences. Notably, the relative application probability declines much more steeply around the peak when using random consideration sets (Figure A4) for most dimensions except distance. This likely reflects the inclusion of irrelevant ads, potentially conflating application probability response with initial pre-screening. Under random sets, the decline around the log-wage peak is also weaker for unemployed seekers, and employed applicants appear less willing to apply for jobs requiring more education than they possess, compared to the baseline. These substantial differences underscore the impact of using behaviorally grounded consideration sets over randomly generated ones.

6 Conclusions

We use data from a Chilean job posting website and a network algorithm to define choice sets for individuals. This approach allowed us to uncover several key insights into the nature of online job search and the differences in behavior between employed and unemployed job seekers.

Our analysis documents how various demographic characteristics of individuals correlate with higher application rates. Notably, we find that males tend to apply more frequently to job positions, while single individuals engage in more on-the-job search. Additionally, we observed that certain job features attract more applications; newer jobs with a higher number of vacancies are particularly appealing, especially to the unemployed.

The highlight of our findings lies in the selective component of job search, that is, how application decisions respond to the potential match fit along several dimensions. To describe these decisions, we focus on the concept of misalignment, defined as the gap between a job ad's requirements and the relevant characteristics of the worker. The richness of our database allowed us to jointly estimate the behavior of job seekers facing misalignment in education, experience, log wages, geographical distance, and occupation. Our analysis reveals that all workers exhibit a negative response to misalignment across these dimensions. We also find that employed job seekers show more ambitious behavior, tending to apply for jobs that require more education than they possess and job ads with wages exceeding their declared expectations. Moreover, our results indicate that log wage misalignment significantly influences application probability.

A core methodological contribution of our work is the construction of consideration sets based on a bipartite network. This approach marks a departure from usual methods defining local labor markets using fixed bins

of occupation, location, or other characteristics. These conventional methods seem inappropriate in our case at least, as job seekers frequently apply to ads across occupational and regional boundaries. Our approach is based on revealed preferences of applicants that potentially integrate multiple job and worker features, observed or not.

Using consideration sets is key to avoid biased estimators because the variables that determine which jobs applicants consider are often the same that influence their application decisions. Since applicants routinely employ search filters and other tools to pre-screen job ads, the naive assumption that all jobs are within every applicant's choice set leads to biased estimations.

In sum, our research offers a nuanced understanding of online job search behavior, emphasizing the importance of appropriately defining choice sets. The empirical results highlight the significance of job seeker characteristics and job ad attributes in the application process, as well as the crucial role of misalignment in shaping job search strategies.

References

ABALUCK, J., AND A. ADAMS-PRASSL (2021): "What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses*," *The Quarterly Journal of Economics*, 136(3), 1611–1663.

AHN, H. J., AND L. SHAO (2017): "Precautionary on-the-job search over the business cycle," *Available at SSRN 2897533*.

ALBRECHT, J., B. DECREUSE, AND S. VROMAN (2023): "Directed search with phantom vacancies," *International Economic Review*, Forthcoming.

AZAR, J., AND I. MARINESCU (2024): "Monopsony Power in the Labor Market: From Theory to Policy," *Annual Review of Economics*, 16(Volume 16, 2024), 491–518.

BAKER, S. R., AND A. FRADKIN (2017): "The impact of unemployment insurance on job search: Evidence from Google search data," *Review of Economics and Statistics*, 99(5), 756–768.

BANFI, S., S. CHOI, AND B. VILLENA-ROLDÁN (2022): "Sorting on-line and on-time," *European Economic Review*, 146, 104128.

BANFI, S., AND B. VILLENA-ROLDÁN (2019): "Do High-Wage Jobs Attract More Applicants? Directed Search Evidence from the Online Labor Market," *Journal of Labor Economics*, 37(3), 715–746.

BELZIL, C. (1996): "Relative efficiencies and comparative advantages in job search," *Journal of Labor Economics*, 14(1), 154–173.

BRANSCH, F. (2021): "Job search intensity of unemployed workers and the business cycle," *Economics Letters*, 205, 109927.

BRENČIČ, V., AND M. PAHOR (2019): "Exporting, demand for skills and skill mismatch: Evidence from employers' hiring practices," *The World Economy*, 42(6), 1740–1773.

CARRILLO-TUDELA, C., AND L. VISSCHERS (2023): "Unemployment and endogenous reallocation over the business cycle," *Econometrica*, 91(3), 1119–1153.

CHÉRON, A., AND B. DECREUSE (2016): "Matching with Phantoms," *Review of Economic Studies*, (0), 1–30.

CHOI, S., A. JANIAK, AND B. VILLENA-ROLDÁN (2015): "Unemployment, Participation and Worker Flows Over the Life-Cycle," *The Economic Journal*, 125(589), 1705–1733.

CLEMENS, J., L. B. KAHN, AND J. MEER (2021): "Dropouts need not apply? The minimum wage and skill upgrading," *Journal of Labor Economics*, 39(S1), S107–S149.

COLES, M., AND B. PETRONGOLO (2008): "A Test between Stock-Flow Matching and the Random Matching Function Approach," *International Economic Review*, 49(4), 1113–1141.

COLES, M. G., AND A. MUTHOO (1998): "Strategic Bargaining and Competitive Bidding in a Dynamic Market Equilibrium," *Review of Economic Studies*, 65(2), 235–260.

COLES, M. G., AND E. SMITH (1998): "Marketplaces and Matching," *International Economic Review*, 39(1), 239–254.

DAVIS, S. J., AND B. SAMANIEGO DE LA PARRA (2024): "Application flows," Discussion paper, National Bureau of Economic Research.

DELLAVIGNA, S., J. HEINING, J. F. SCHMIEDER, AND S. TRENKLE (2021): "Evidence on Job Search Models from a Survey of Unemployed Workers in Germany," *The Quarterly Journal of Economics*, 137(2), 1181–1232.

DELOACH, S. B., AND M. KURT (2013): "Discouraging workers: estimating the impacts of macroeconomic shocks on the search intensity of the unemployed," *Journal of Labor research*, 34, 433–454.

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.

Ebrahimy, E., and R. Shimer (2010): "Stock-flow matching," *Journal of Economic Theory*, 145(4), 1325–1353.

Elsby, M. W., B. Hobijn, and A. Şahin (2013): "Unemployment Dynamics in the OECD," *Review of Economics and Statistics*, 95(2), 530–548.

Fabel, O., and R. Pascalau (2013): "Recruitment of seemingly overeducated personnel: insider–outsider effects on fair employee selection practices," *International Journal of the Economics of Business*, 20(1), 57–82.

Faberman, R. J., and M. Kudlyak (2019): "The intensity of job search and search duration," *American Economic Journal: Macroeconomics*, 11(3), 327–57.

Faberman, R. J., A. I. Mueller, A. Şahin, and G. Topa (2022): "Job search behavior among the employed and non-employed," *Econometrica*, 90(4), 1743–1779.

Fluchtmann, J., A. M. Glenny, N. Harmon, and J. Maibom (2024): "Unemployed job search across people and over time: Evidence from applied-for jobs," *Journal of Labor Economics*, 42(4), 1175–1217.

Fredriksson, P., L. Hensvik, and O. N. Skans (2018): "Mismatch of talent: Evidence on match quality, entry wages, and job mobility," *American Economic Review*, 108(11), 3303–3338.

Fujita, S. (2012): "An empirical analysis of on-the-job search and job-to-job transitions," Working paper, Federal Reserve Bank of Philadelphia.

Gomme, P., and D. Lkhagvasuren (2015): "Worker search effort as an amplification mechanism," *Journal of Monetary Economics*, 75, 106–122.

Gregg, P., and B. Petrongolo (2005): "Stock-flow matching and the performance of the labor market," *European Economic Review*, 49(8), 1987–2011.

Hazell, J., and B. Taska (2024): "Downward Rigidity in the Wage of New Hires," *American Economic Review*, forthcoming.

Holzer, H. J. (1987): "Job search by employed and unemployed youth," *ILR Review*, 40(4), 601–611.

HORNSTEIN, A., P. KRUSELL, AND G. L. VIOLANTE (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *The American Economic Review*, 101(7).

JACCARD, P. (1901): "Étude de la distribution florale dans une portion des Alpes et du Jura," *Bulletin de la Societe Vaudoise des Sciences Naturelles*, 37, 547–579.

JACKSON, M. O. (2008): *Social and Economic Networks*. Princeton University Press.

JAROSCH, G., J. S. NIMCZIK, AND I. SORKIN (2024): "Granular search, market structure, and wages," *Review of Economic Studies*, 91(6), 3569–3607.

JOVANOVIC, B. (1979): "Job matching and the theory of turnover," *Journal of political economy*, 87(5, Part 1), 972–990.

KAMBOUROV, G., AND I. MANOVSKII (2009): "Occupational Mobility and Wage Inequality," *Review of Economic Studies*, 76(2), 731–759.

KARRER, B., AND M. E. NEWMAN (2011): "Stochastic blockmodels and community structure in networks," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 83(1), 016107.

KUDLYAK, M., D. LKHAGVASUREN, AND R. SYSUYEV (2013): "Systematic Job Search: New Evidence from Individual Job Application Data," mimeo, Federal Reserve Bank of Richmond.

KUHN, M., I. MANOVSKII, AND X. QIU (2021): "The Geography of Job Creation and Job Destruction," Working Paper 29399, National Bureau of Economic Research.

LAMADON, T., M. MOGSTAD, AND B. SETZLER (2022): "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 112(1), 169–212.

LEYVA, G. (2018): "Against All Odds: Job Search During the Great Recession," Unpublished Manuscript. Banco de Mexico.

LI, F., AND X. WENG (2017): "Efficient Learning and Job Turnover in the Labor Market," *International Economic Review*, 58(3), 727–750.

LONGHI, S., AND M. TAYLOR (2011): "Explaining differences in job search outcomes between employed and unemployed job seekers," Discussion paper, IZA Discussion Papers.

——— (2013): "Occupational change and mobility among employed and unemployed job seekers," *Scottish Journal of Political Economy*, 60(1), 71–100.

MANNING, A., AND B. PETRONGOLO (2017): "How local are labor markets? Evidence from a spatial job search model," *American Economic Review*, 107(10), 2877–2907.

MARINESCU, I., AND R. RATHELOT (2018): "Mismatch unemployment and the geography of job search," *American Economic Journal: Macroeconomics*, 10(3), 42–70.

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.

MATSUDA, N., T. AHMED, AND S. NOMURA (2019): "Labor market analysis using big data: The case of a Pakistani online job portal," *World Bank Policy Research Working Paper*, (9063).

MCFADDEN, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. by P. Zarembka. Academic Press, New York.

MENZIO, G., I. A. TELYUKOVA, AND L. VISSCHERS (2016): "Directed search over the life cycle," *Review of Economic Dynamics*, 19, 38 – 62, Special Issue in Honor of Dale Mortensen.

MICHELACCI, C., AND J. SUAREZ (2006): "Incomplete Wage Posting," *Journal of Political Economy*, 114(6), 1098–1123.

MUKOYAMA, T., C. PATTERSON, AND A. ŞAHIN (2018): "Job Search Behavior over the Business Cycle," *American Economic Journal: Macroeconomics*, 10(1), 190–215.

NAUDON, A., AND A. PÉREZ (2018): "Unemployment dynamics in Chile: 1960-2015," *Journal Economía Chilena (The Chilean Economy)*, 21(1), 4–33.

NG'OMBE, J. N., AND B. W. BRORSEN (2022): "The effect of including irrelevant alternatives in discrete choice models of recreation demand," *Computational Economics*, 60(1), 71–97.

NIMCZIK, J. S. (2023): "Job mobility networks and endogenous labor markets," working paper.

PISSARIDES, C. A., AND J. WADSWORTH (1994): "On-the-job search: Some empirical evidence from Britain," *European Economic Review*, 38(2), 385–401.

ŞAHIN, A., J. SONG, G. TOPA, AND G. L. VIOLANTE (2014): "Mismatch unemployment," *American Economic Review*, 104(11), 3529–64.

SHIMER, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95(1), 25–49.

TAYLOR, C. R. (1995): "The Long Side of the Market and the Short End of the Stick: Bargaining Power and Price Formation in Buyers', Sellers', and Balanced Markets," *The Quarterly Journal of Economics*, 110(3), 837–855.

TENN, S., AND J. M. YUN (2008): "Biases in demand analysis due to variation in retail distribution," *International Journal of Industrial Organization*, 26(4), 984–997.

TRAIN, K. (2009): *Discrete Choice Methods with Simulation*. Cambridge Univ Press, Second edition.

VAN NIEROP, E., B. BRONNENBERG, R. PAAP, M. WEDEL, AND P. H. FRANSES (2010): "Retrieving unobserved consideration sets from household panel data," *Journal of Marketing Research*, 47(1), 63–74.

VAN OMMEREN, J., AND G. RUSSO (2013): "Firm Recruitment Behaviour: Sequential or Non-sequential Search?," *Oxford Bulletin of Economics and Statistics*, pp. 1–24.

VAN OURS, J., AND G. RIDDER (1992): "Vacancies and the Recruitment of New Employees," *Journal of Labor Economics*, 10(2), 138–155.

A Online Appendix

A.1 Estimation bias due to ignoring consideration sets

As pointed out before, Tenn and Yun (2008) and Ng'ombe and Brorsen (2022) show that including irrelevant alternatives into the consideration set significantly bias results because there is an intuitive correlation between characteristics driving applications and consideration sets. We further elaborate this argument in here:

We build on Tenn and Yun (2008) who mainly concern with the availability of products in retail stores for demand estimation. Consider a discrete choice model in which a worker w values attributes of a job ad a according to a linear utility function

$$U_{wa} = X_{wa}\beta + \epsilon_{wa}$$

where ϵ_{wa} follows a Gumbel distribution.

As widely known in the demand estimation literature, the probability that the worker w applies to the job a takes a multinominal logit form (McFadden, 1973; Train, 2009).

$$\tilde{\pi}_{wa} = \frac{\exp(X_{wa}\beta)}{\sum_b \exp(X_{wb}\beta)}$$

Tenn and Yun (2008) introduce the notion of consideration set by defining an indicator variable that takes value 1 if a product is available in a store and 0 otherwise. They refer to this as "heterogenous store logit model". In our context, we consider weights $q(w,a)$ between a worker w and an ad a to have a role that is similar to the one denoting store availability of a product. Assuming that weights $q(w,a)$ are an increasing function of the true probability that the job a is the consideration set of worker w. This originates a logit model with heterogeneous consideration sets where the probability of observing an application of w to and ad a is

$$\pi_{wa} = \frac{q(w,a)\exp(X_{wa}\beta)}{\sum_{b\in\mathcal{A}_w} q(w,b)\exp(X_{wb}\beta)}$$

We assume an outside option $a = 0$ (jobs in another platform, current job, unemployment, etc) whose value is normalized to zero. Therefore, we obtain

$$\frac{\pi_{wa}}{\pi_{w0}} = \frac{q(w,a)\exp(X_{wa}\beta)}{q(w,0)\exp(X_{w0}\beta)} = \frac{q(w,a)}{q(w,0)}\exp(X_{wa}\beta)$$

Taking logarithms, we obtain

$$\log\left(\frac{\pi_{wa}}{\pi_{w0}}\right) = X_{wa}\beta + \log\left(\frac{q(w,a)}{q(w,0)}\right)$$

In case that our algorithm generates $q(w,a) = 0$ because the worker w does not consider ad a, we take an approximation of $q(w,a) = \varepsilon > 0$, for an arbitrary low value of ε.

In the case of an unweighted estimation where all available ads are included into the consideration set, so $q(w,a) = 1$ for all ads $a = 0, 1, 2, ...$, the previous equation becomes

$$\log\left(\frac{\pi_{wa}}{\pi_{w0}}\right) = X_{wa}\beta$$

Therefore, the naive unweighted model actually has an omitted variable $Z_{wa} = \log\left(\frac{q(w,a)}{q(w,0)}\right)$. If we estimated this model without recognizing this variable, our estimates have an omitted variable in that

$$B(\hat{\beta}_j) = \frac{cov(X_{-jwa}, Z_{wa})}{var(X_{-jwa})}$$

where X_{-jwa} is the observation for the pair (w, a) of the attribute X_j when all other covariates have been partialed-out, i.e. the OLS residual obtained from a regression with X_j as dependent variable and all the other X as independent variables. Variables that increase the likelihood of application (conditional on other covariates) are likely to increase the chance of being into the consideration set, leading to a positive bias. The same argument applies with opposite effects with attributes that, conditionally on other covariates, deter applications.

Moreover the bias can be re-written as

$$B(\hat{\beta}_j) = \frac{cov(X_{-jwa}, Z_{wa}|X_{-jwa} \geq 0)Prob(X_{-jwa} \geq 0) + cov(X_{-jwa}, Z_{wa}|X_{-jwa} < 0)Prob(X_{-jwa} < 0)}{var(X_{-jwa})}$$

The last characterization help us think in cases of misalignment. If $X_{-jwa} \geq 0$, there is a negative correlation between this positively misaligned attribute and the likelihood that the job ad a shows up into the consideration set of the worker w, $\mathcal{A}_w$. In this case, the bias becomes negative, leading to underestimation of β_j. If $X_{-jwa} < 0$, there is a positive correlation between this negatively misaligned attribute and the likelihood that the job ad a shows up into the consideration set of the worker w, $\mathcal{A}_w$. In this case the bias becomes positive, leading to overestimation of β_j. This suggests that in most cases there is an attenuation bias in the estimates of misalignment characteristics.

A.2 Network metrics analysis

Using basic results of set cardinality, the Jaccard similarity metric for application sets of w and v can be expressed as follows

$$b(w, v) \equiv \frac{\left|\mathcal{A}_w^0 \cap \mathcal{A}_v^0\right|}{\left|\mathcal{A}_w^0 \cup \mathcal{A}_v^0\right|} \equiv \frac{\left|\mathcal{A}_w^0\right| + \left|\mathcal{A}_v^0\right|}{\left|\mathcal{A}_w^0 \cup \mathcal{A}_v^0\right|} - 1$$

Extending the definition of similarity for sets of degree 1, we obtain[1]

$$b^1(w, v) \equiv \frac{\left|\mathcal{A}_w^1 \cap \mathcal{A}_v^1\right|}{\left|\mathcal{A}_w^1 \cup \mathcal{A}_v^1\right|}$$

Consequentially, the max similarity measure between and ad a and a worker w, is defined using sets of degree 1, i.e.[2]

$$q^1(w, a) = \max_{u:\, a \in \mathcal{A}_u^1} \left\{b^1(w, u)\right\}$$

For every worker w, the set of degree 1 equals

$$\mathcal{A}_w^1 = \bigcup_{u: \mathcal{A}_w^0 \cap \mathcal{A}_u^0 \neq \emptyset} \left(\mathcal{A}_w^0 \cup \mathcal{A}_u^0\right) = \mathcal{A}_w^0 \cup \mathcal{A}_v^0 \cup \mathcal{A}_{-w,-v}^0$$

where

$$\mathcal{A}_{-w,-v}^0 = \bigcup_{u: \mathcal{A}_w^0 \cap \mathcal{A}_u^0 \neq \emptyset, u \neq w, u \neq v} \mathcal{A}_u^0$$

Likewise, we have

$$\mathcal{A}_v^1 = \mathcal{A}_v^0 \cup \mathcal{A}_w^0 \cup \mathcal{A}_{-v,-w}^0$$

[1] We choose to avoid overloading notation, but strictly speaking we should denote the similarity as $b(w, v) \equiv b^0(w, v)$. We omit the superindex 0 as our baseline case. If it is not zero, we are explicit about this by denoting $b^s(w, v)$.

[2] Similarly, $q(w, a) \equiv q^0(w, a)$.

where

$$\mathcal{A}^0_{-v,-w} = \bigcup_{u:\mathcal{A}^0_v \cap \mathcal{A}^0_u \neq \emptyset, u \neq v, u \neq w} \mathcal{A}^0_u$$

Notice than, in general, $\mathcal{A}^0_{-w,-v} \neq \mathcal{A}^0_{-v,-w}$

With these elements, we can establish the following

Lemma 1: expansion of unions in network degree: The union of two sets of degree 1 for workers w and v has at least the same cardinality as the union of the two sets of degree 0 for both workers.

$$\mathcal{A}^1_w \cup \mathcal{A}^1_v = \left(\mathcal{A}^0_w \cup \mathcal{A}^0_v\right) \cup \left(\mathcal{A}^0_{-w,-v} \cup \mathcal{A}^0_{-v,-w}\right)$$

and therefore it follows that

$$\left|\mathcal{A}^1_w \cup \mathcal{A}^1_v\right| \geq \left|\mathcal{A}^0_w \cup \mathcal{A}^0_v\right|$$

The result can be extended through an induction argument for any network of degree s.

$$\left|\mathcal{A}^{s+1}_w \cup \mathcal{A}^{s+1}_v\right| \geq |\mathcal{A}^s_w \cup \mathcal{A}^s_v|$$

Moreover we can also establish that

Lemma 2: expansion of intersections in network degree: The intersection of two sets of degree 1 for workers w and v has at least the same cardinality as the union of the two sets of degree 0 for both workers.

$$\mathcal{A}^1_w \cap \mathcal{A}^1_v = \left(\mathcal{A}^0_w \cup \mathcal{A}^0_v\right) \cup \left(\mathcal{A}^0_{-w,-v} \cap \mathcal{A}^0_{-v,-w}\right).$$

Therefore, the following result is apparent

$$\left|\mathcal{A}^1_w \cap \mathcal{A}^1_v\right| \geq \left|\mathcal{A}^0_w \cup \mathcal{A}^0_v\right|$$

Through an induction argument for any network of degree s, the result is extended.

$$\left|\mathcal{A}^{s+1}_w \cap \mathcal{A}^{s+1}_v\right| \geq |\mathcal{A}^s_w \cup \mathcal{A}^s_v|$$

Proposition 1: similarity between workers increases in network degree: In addition to Lemmas 1 and 2, we also know that

$$\left|\mathcal{A}^{s+1}_w \cup \mathcal{A}^{s+1}_v\right| \geq \left|\mathcal{A}^{s+1}_w \cap \mathcal{A}^{s+1}_v\right|$$

Therefore, as the network degree s increases, the similarity measure between workers w and v increases, but is never greater than 1, i.e

$$b^{s+1}(w,v) \equiv \frac{\left|\mathcal{A}^{s+1}_w \cap \mathcal{A}^{s+1}_v\right|}{\left|\mathcal{A}^{s+1}_w \cup \mathcal{A}^{s+1}_v\right|} \geq b^s(w,v) \equiv \frac{|\mathcal{A}^s_w \cap \mathcal{A}^s_v|}{|\mathcal{A}^s_w \cup \mathcal{A}^s_v|}$$

Based on Proposition 1, we know that using networks of higher and higher degree to define consideration sets would lead to weights near 1 for all ads that are not isolated, i.e. for every ad a there is some w and v such that $a \in \mathcal{A}^0_w \cap \mathcal{A}^0_v$.

Computing the proximity $q(w,a)$ by taking the maximum out of all similarities $b(w,v)$ between workers v who have a into their consideration sets is justified under these criteria:

1. Minimization of discrepancies between networks defined by different degrees: since we establish in Proposition 1 that $b(w,v)^s$ with $s = 0, 1, ...$ increases in the degree of the network, by focusing on the maximum of proximities, we effectively reduce as much as possible the discrepancy between different networks of different degree.

2. Robustness to minor variations in the network: since we take the maximum proximity between pairs of

workers sharing a reference ad a in their consideration sets $\mathcal{A}^s$, our preferred measure does not change unless a variation in the network generates a larger proximity.

3. Shortest path: the Jaccard metric b equals the density of the local network defined between two workers connected via coincident applications, i.e. the share of actual links out of total potential links. If ads are randomly chosen, the metric $b(w, v)$ can be interpreted as the probability that workers w and v choose a common job ad. When computing the $q(w, a)$ metric as the maximum of Jaccard metrics b, we simply choose the most likely path linking the ad a and w under the previous assumption.

By no means, the maximum criteria is a "correct" choice of measuring the proximity between ads and workers. This claim applies more generally to size or distance metrics in networks (Jackson, 2008).

A.3 Network simple example

To illustrate the computation of consideration sets using different methods, we propose the following network portrayed in Figure A.3

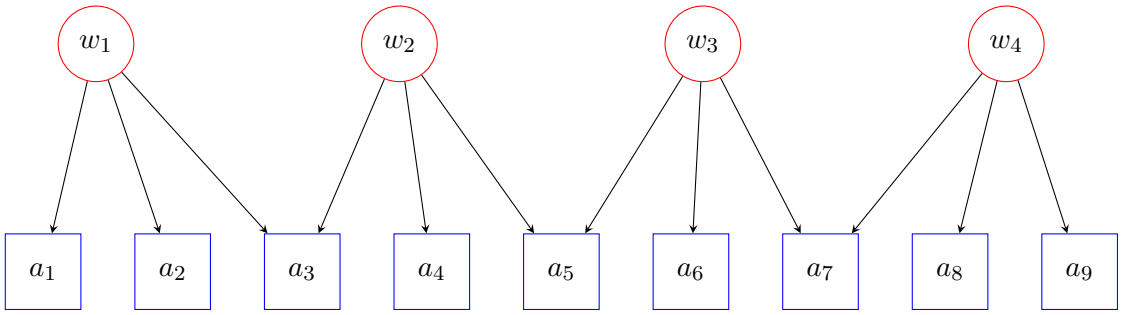

Figure A1: Example bipartite worker-ad network

In this example,

- $\mathcal{A}_1^0 = \{a_1, a_2, a_3\}$
- $\mathcal{A}_2^0 = \{a_3, a_4, a_5\}$
- $\mathcal{A}_3^0 = \{a_5, a_6, a_7\}$
- $\mathcal{A}_4^0 = \{a_7, a_8, a_9\}$

Therefore, using our definitions, we obtain that

- $\mathcal{A}_1^1 = \{a_1, a_2, a_3, a_4, a_5\}$
- $\mathcal{A}_2^1 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$
- $\mathcal{A}_3^1 = \{a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$
- $\mathcal{A}_4^1 = \{a_5, a_6, a_7, a_8, a_9\}$

The proximity between two ads, $b(w, v)$ is

$$b^s(w, v) = \frac{|\mathcal{A}_w^s \cap \mathcal{A}_v^s|}{|\mathcal{A}_w^s \cup \mathcal{A}_v^s|}.$$

Table A.3 shows the proximity between networks of degree zero for different workers.
Table A.3 shows the proximities between workers in a networks of degree one is substantially larger.

	w_1	w_2	w_3	w_4
w_1	1	1/5	0	0
w_2	1/5	1	1/5	0
w_3	0	1/5	1	1/5
w_4	0	0	1/5	1

Table A1: Worker proximity matrix $b(w, v)$ with zero-degree networks, $\mathcal{A}_w^0$

	w_1	w_2	w_3	w_4
w_1	1	5/7	1/3	1/9
w_2	5/7	1	5/9	1/3
w_3	1/3	5/9	1	5/7
w_4	1/9	1/3	5/7	1

Table A2: Worker proximity matrix with one-degree networks, $\mathcal{A}_w^1$

The proximity measures are computed according to the definition

$$q^s(w, a) = \max_{u:\, a \in \mathcal{A}_u^s} \{b^s(w, u)\}$$

so that, for non-trivial cases in which a belongs to the set of applications $\mathcal{A}_w^0$, we have that

$$\begin{aligned}
q^0(a_4, w_1) &= \max_{u:\, a_4 \in \mathcal{A}_u^0} \left\{b^0(w_1, u)\right\} \\
&= \max\left\{b^0(w_1, w_2)\right\} = 1/5 \\
q^0(a_5, w_1) &= \max_{u:\, a_5 \in \mathcal{A}_u^0} \left\{b^0(w_1, u)\right\} \\
&= \max\left\{b^0(w_1, w_2)\right\} = 1/5 \\
q^0(a_6, w_1) &= \max_{u:\, a_6 \in \mathcal{A}_u^0} \left\{b^0(w_1, u)\right\} \\
&= \max\left\{b^0(w_1, w_2), b^0(w_1, w_3)\right\} = \max\{1/5, 0\} = 1/5 \\
q^0(a_7, w_1) &= \max_{u:\, a_7 \in \mathcal{A}_u^0} \left\{b^0(w_1, u)\right\} \\
&= \max\left\{b^0(w_1, w_3)\right\} = 0 = q^0(a_8, w_1) = q^0(a_9, w_1)
\end{aligned}$$

Table A.3 shows the proximities between ads and workers in a networks of degree zero.

For the case of one-degree networks, the ad-worker proximity $q^1(w, a)$ is computed in the following cases as an illustration

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
w_1	1	1	1	1/5	1/5	1/5	0	0	0
w_2	1/5	1/5	1	1	1	1/5	1/5	0	0
w_3	0	0	1/5	1/5	1	1	1	1/5	1/5
w_4	0	0	0	1/5	1/5	1/5	1	1	1

Table A3: Ad-worker proximity matrix with zero-degree networks, $q^0(w, a)$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
w_1	1	1	1	1	1	5/7	5/7	1/3	1/3
w_2	1	1	1	1	1	1	1	5/9	5/9
w_3	5/9	5/9	1	1	1	1	1	1	1
w_4	1/3	1/3	5/7	5/7	1	1	1	1	1

Table A4: Ad-worker proximity matrix with one-degree networks, $q^1(w, a)$

$$
\begin{aligned}
q^1(a_6, w_1) &= \max_{u:\, a_6 \in \mathcal{A}_u^1} \left\{ b^1(w_1, u) \right\} \\
&= \max \left\{ b^1(w_1, w_2), b^1(w_1, w_3), b^1(w_1, w_4) \right\} = \max \{5/7, 1/3, 1/9\} = 5/7 \\
q^1(a_7, w_1) &= \max_{u:\, a_7 \in \mathcal{A}_u^1} \left\{ b^1(w_1, u) \right\} \\
&= \max \left\{ b^1(w_1, w_2), b^1(w_1, w_3), b^1(w_1, w_4) \right\} = \max \{5/7, 1/3, 1/9\} = 5/7 \\
q^1(a_8, w_1) &= \max_{u:\, a_8 \in \mathcal{A}_u^1} \left\{ b^1(w_1, u) \right\} \\
&= \max \left\{ b^1(w_1, w_3), b^1(w_1, w_4) \right\} = \max \{1/3, 1/9\} = 1/3 \\
q^1(a_9, w_1) &= \max_{u:\, a_9 \in \mathcal{A}_u^1} \left\{ b^1(w_1, u) \right\} \\
&= \max \left\{ b^1(w_1, w_3), b^1(w_1, w_4) \right\} = \max \{1/3, 1/9\} = 1/3
\end{aligned}
$$

Table A.3 finally shows the proximities between ads and workers in a networks of degree one, which are much greater values.

A.4 Details of data handling

In this section, we report some details about data management.

Elapsed duration of labor status: Among the employed, around 40% of the sample have no measured tenure since the starting date of the job is unreported. To keep these observations in our sample, we define a dummy variable for missing tenures and impute a value of zero to all unobserved tenures. In this way, the estimated tenure profile should be interpreted as conditional on declaring a starting date for the current job. The coefficient of the missing tenure binary variable, in turn, is interpreted as the differential effect in application probability of an undeclared starting date with respect to an observed zero tenure. The same strategy is used for unemployed job seekers, but in this case around 7% of starting dates are missing.

Consideration dates for not-applied jobs: For worker-ad pairs that are matched given our network algorithm, the date of an actual application does not exist. In those cases, we impute the date of application by the mode date of applications of the linked workers to that particular job ad.

Missing region data: To avoid losing observations due to missing region, we use dummy variables for those cases and impute a value of zero for unobserved values. Hence, we do not consider ads offering jobs with unknown, multiple, or international locations. The estimated application response to distance should be interpreted as the effect of that variable conditional on observing both the location of the job ad and the applicant (though this last case in very rare). The coefficient of the missing region binary variable should be interpreted as the differential effect in application probability of missing job ad region with respect to an observed zero distance (intraregional application).

A.5 Consistency condition

A more detailed explanation of the required condition result is as follows. Consider that all the right-hand side variables comprising worker w and ad a characteristics $X_{wa}, \{z_{k,aw}\}_k, z_{j,aw}$ are stacked into the vector S_{wa} and all the corresponding parameters are stacked into the vector θ so that the linear probability model would be $y_{wa} = S_{wa}\theta + \epsilon_{wa}$. Then, using weights $q(w,a)$, we run regressions of the form $\sqrt{\varphi(w,a)}y_{wa} = \sqrt{\varphi(w,a)}S_{wa}\theta + \sqrt{\varphi(w,a)}\epsilon_{wa}$ to estimate coefficients weighted by q, as in a standard version of Generalized Least Squares. To ensure that we obtain consistent estimators, it must be true that $E[\sqrt{\varphi(w,a)}\epsilon_{wa}|S_{wa}] = \sqrt{\frac{\widehat{p}(w,a)}{1-\widehat{p}(w,a)}}cov(\sqrt{q(w,a)}, \epsilon_{wa}|S_{wa}) = 0$, which occurs only occurs if the last covariance equals zero. A stronger condition is to require independence between weights $q(w,a)$ and the application error term ϵ.

A.6 Additional descriptive stats

Table A5: Percentage of matching characteristics of ads and applicants

match variables	ads applied (%)	ads not applied (%)	ads not applied (%) (weighted)
absolute wage gap $\leq 10\%$	17.3	12.5	12.9
absolute wage gap $\leq 25\%$	42.8	32.4	33.3
match education level	54.5	44.7	46.4
match experience	8.5	7.5	8.1
absolute experience ≤ 1 year	43.7	42.4	44.1
match region	50.8	58.3	56.8
distance ≤ 100 km	52.1	59.4	58.1
distance ≤ 200 km	56.1	63.0	62.2
match occupation	51.7	39.8	41.5

Table A6: Average consideration set sizes by characteristics and labor status of workers

	employed	unemployed	total
female	37.4	37.9	37.7
male	39.7	38.7	39.2
age 18-24	37.6	38.5	38.3
age 25-29	38.2	37.9	38.0
age 30-34	37.5	37.9	37.7
age 35-39	39.2	38.0	38.6
age 40-44	40.6	39.4	39.9
age 45-54	42.5	39.3	40.5
age 55+	42.1	39.8	40.5
high school	37.1	35.8	36.1
tech tertiary	41.2	41.6	41.4
college	38.3	38.4	38.4
graduate	36.6	39.7	37.9
total	38.8	38.3	38.5

A.7 Life-cycle, duration and business cycle effects

We report the predicted application probability varying age, duration of employment status, and unemployment between the 1st and 99th percentiles of their sample values, while keeping the other covariates at their mean values in figure A2.

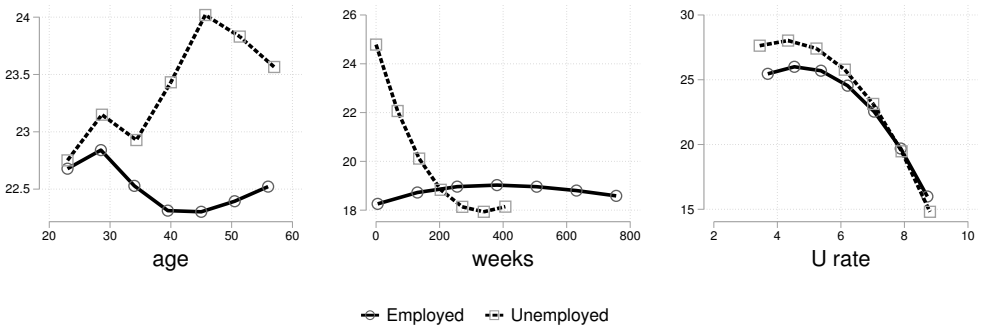

Figure A2: Predicted application probabilities for different ages, number of weeks in the current labor force status, and national unemployment rate at the time of the application decision, given results from equation (3), no compositional adjustment. The figure is computed using the coefficients associated to a polynomial of order 5 on each variable and leaving the rest of regressors at their sample mean.

In the left panel, we observe that the unemployed apply more often to ads in their consideration sets at all ages, and their probability of application increases with age, with an overall peak at age 45 to decrease until mid-fifties. For the unemployed, the application probability is higher for individuals under 30, and then decreases until the mid-forties, and then slightly increases. While this evidence might seem only partially consistent with job finding rates and employment-to-employment transitions over the life-cycle as reported by Choi, Janiak, and Villena-Roldán (2015) and Menzio, Telyukova, and Visschers (2016), and Naudon and Pérez (2018) for Chile, we point out two reasons why this is not the case. First, in these papers, job finding rates refer to the larger frequency of realized transitions. In contrast, our evidence here is about search or application effort. Indeed, Mukoyama, Patterson, and Şahin (2018) show a slightly increasing profile of effort on the intensive

margin of time devoted to job search until age 50. Second, the self-selected sample of older workers using the online job board may be somewhat different from the average worker in the labor force of that age.

The middle panel in the figure shows a decreasing application probability as the search duration increases, measured as the time elapsed between the finishing date of the previous job and the application date. The extended range of durations suggests that equalizing traditional unemployment duration with our measure of search duration is far-fetched. Thus, an appropriate interpretation is that individuals who have lost jobs and are website users make most of their applications soon after the separation. For employed jobseekers, the application likelihood seems to be flat for the most part even though there is a slightly increasing trend up to 400 weeks, or nearly eight years. Two offsetting factors may be at play: a growing match-specific human capital deterring on-the-job search and a market-learning process of the worker that increases the outside value of the applicant.

A decreasing probability of application as the unemployed search duration increases, is an important issue for the design of unemployment insurance policies, as stated by Faberman and Kudlyak (2019) and DellaVigna, Heining, Schmieder, and Trenkle (2021), among others. Results for employed seekers are consistent with theory (Jovanovic, 1979; Li and Weng, 2017) and previous evidence (Pissarides and Wadsworth, 1994; Fujita, 2012). This also qualitatively consistent with the evidence of realized job-to-job flows in Menzio, Telyukova, and Visschers (2016). This finding is relevant to discipline models explaining job-to-job transitions and frictional wage dispersion, as in Hornstein, Krusell, and Violante (2011).

In terms of business cycle conditions, the right panel of figure A2 shows a decreasing relationship between the unemployment rate, our cyclical variable, and application decisions. The application probability remains flat between 25-30% and slightly higher for the unemployed when the regional labor market exhibits low unemployment, i.e. below 5.5%. When moving to regional labor markets showing unemployment rates between 6% and 8.3%, the average application probability declines from 25% to 15% and is very similar for both employed and unemployed applicants.

Since it is well-known that the job finding probability is procyclical (Shimer, 2005; Elsby, Hobijn, and Şahin, 2013; Naudon and Pérez, 2018) the larger effort exerted in slack labor (high-unemployment) markets is simply not sufficiently high to generate a countercyclical job finding probability. Hence, the general decreasing pattern of application probability in unemployment rate suggests that job seekers find that their search effort cannot compensate for the scarcity of available jobs when unemployment is high, unlike Faberman and Kudlyak (2019), Mukoyama, Patterson, and Şahin (2018) and Bransch (2021). In contrast, the finding aligns with DeLoach and Kurt (2013) and Gomme and Lkhagvasuren (2015). Yet Leyva (2018) finds roughly acyclical search effort, as in the lower-end of the unemployment rate in our sample. Our finding of non-monotonicity of the effect helps reconciling these heterogeneous pieces of evidence in the literature.

A.8 Results under alternative consideration sets

Table A7: Average component coefficients by labor status, alternative estimations

VARIABLES	(1) Employed weights^A	(2) Employed DFL	(3) Employed Random	(4) Unemployed weights^A	(5) Unemployed DFL	(6) Unemployed Random
Married	-2.769***	-0.713	-37.047	-0.15	-0.897***	-13.662
	(0.931)	(0.455)	(490.373)	(0.498)	(0.237)	(257.673)
Male	1.609***	0.320***	0.114***	2.324***	0.443***	0.089***
	(0.066)	(0.032)	(0.005)	(0.050)	(0.025)	(0.005)
Explicit wage (w)	0.261***	-0.02	-0.051***	0.808***	0.050**	-0.018***
	(0.057)	(0.028)	(0.005)	(0.047)	(0.023)	(0.005)
Explicit wage (a)	-2.724***	-0.662***	-0.145***	-0.955***	-0.240***	0.067***
	(0.088)	(0.042)	(0.007)	(0.067)	(0.032)	(0.007)
No. of Vacancies (a)	-0.007	0.008	0.002***	0.045***	0.018***	0.003***
	(0.011)	(0.006)	(0.000)	(0.006)	(0.003)	(0.000)
Ad duration (weeks)	-0.001	0.001	0.065***	-0.087***	-0.017***	0.094***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	2,124,244	2,124,244	15,210,765	3,184,675	3,184,675	18,093,735
R-squared	0.12	0.04	0.02	0.11	0.04	0.03
Mean app. prob.	20.91	3.98	0.82	26.87	4.15	0.98

Notes: Regression coefficients from a linear regression on application decisions. Dependent variable is y_{wa}, a dummy for the existence of a job application. Each regression controls also for polynomials and interactions in *misalignment* as well as age of the worker, firm size, contract type, dummies for different types of requirements of the job and characteristics of the firm (see details in the main text). Standard errors in parentheses. One, two, and three asterisks indicate significance at 10%,5%, and 1%, respectively.

Figure A3: Predicted application probabilities, relative to mean application probability, given results from eq. (3) and different levels of misalignment in the selected variable x (see main text for details). The rest of regressors are at their sample means. no compositional adjustments. Network sample.

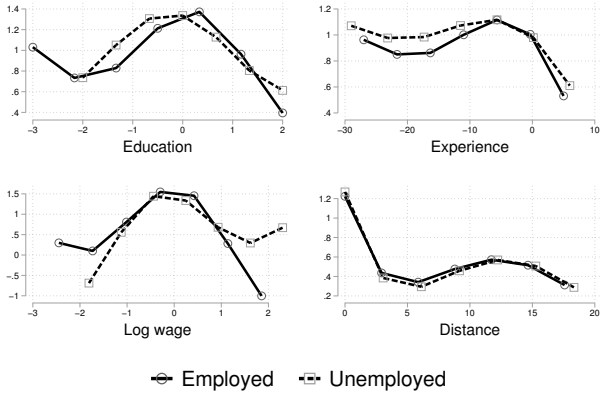

Figure A4: Predicted application probabilities, relative to mean application probability, given results from eq. (3) and different levels of misalignment in the selected variable x (see main text for details). The rest of regressors are at their sample means. no compositional adjustments. Random sample.