Job-to-job Mobility, Outside Options, and Bargaining: A Machine Learning Approach*

Pietro Geuna

Paris School of Economics, Institut des Politiques Publiques, and EHESS

pietro.geuna@psemail.eu

October 30, 2026

Abstract

Outside options play a central role in models of wage determination, yet the employment opportunities available to individual workers are not directly observable. I develop an individual-level measure of the value of alternative employment, rooted in a discrete choice model of labor supply estimated from observed job-to-job transitions using neural networks to account flexibly for differences in workers' preferences and their compatibility with potential employers. I validate the measure in linked employer–employee data from the Veneto region of Italy, showing that workers with higher measured outside options recover more quickly from mass layoffs. Improvements in outside options lead to greater job-to-job mobility and higher wages within ongoing employment spells: a one standard deviation increase raises wages by approximately 1.4%. An instrumental variable design based on large-firm openings confirms the causal reading of these results. The findings replicate in Portuguese linked employer–employee data, where increases in employers' value added also pass through more strongly to the wages of workers with better outside options. Taken together, these findings support wage bargaining as an empirically relevant mechanism of wage determination.

JEL Codes: J31, J42, J62, C35, C45.

Keywords: Outside options; wage bargaining; employer mobility; monopsony; discrete choice; neural networks.

*I thank Luc Behaghel, Richard Blundell, Leonard Bocquet, Thomas Breda, Alexander Busch, Sydnee Caldwell, David Card, Áureo de Paula, Xavier D'Haultfœuille, Monica Essig Aberg, Theodor Friederiszick, Christophe Gaillac, Giuseppe Grasso, Patrick Kline, Attila Lindner, Alan Manning, David Margolis, Gina Markov, Eric Maurin, Maria Montoya-Aguirre, Martin Mugnier, Fabien Postel-Vinay, Steven Raphael, Jesse Rothstein, Nina Roussille, Emmanuel Saez, Benjamin Schoefer, Lorenzo Sinno, Elena Spadini, and Nelson Sunuwar for helpful comments and suggestions. I also thank participants at the UC Berkeley IRLE seminar, the UC Berkeley Labor Lunch seminar, the PSE Winter Retreat, the UCL Ph.D. seminar, the 2nd Berlin Ph.D. Conference in Applied Microeconomics, EEA-ESEM 2026, and EALE 2026. All remaining errors are my own.

1 Introduction

In standard models of wage determination, a worker's outside option plays a central role. Whether wages are set through bargaining, where outside options serve as the threat point in negotiations, or through competitive forces that push compensation toward the value of the next-best alternative, what a worker could earn elsewhere shapes what they earn today. Workers whose skills are demanded by a broad range of employers have greater scope to seek favorable employment terms than those whose skills are valued by only a few firms. Geography also matters: workers in areas with several relevant employers have more alternatives to their current job than those whose employment opportunities are concentrated within a few local employers. In a broad class of models, outside options determine bargaining power, constrain employer wage-setting, and ultimately affect the distribution of surplus between workers and firms.

Yet, despite their theoretical importance, outside options are not directly observable. Employment records reveal the jobs workers hold and the transitions they make, but not the full set of employers they could have joined, or how they value those alternatives. Broad measures based on geographical and industry categories can miss differences in the opportunities available to individual workers. A nearby employer may have no jobs suited to a worker's skills, or may offer pay and working conditions that the worker finds unattractive. A fine-grained measure of outside options therefore requires reconstructing the value of alternative employment for each worker, accounting for the accessibility and relevance of potential employers, as well as differences in preferences.

This paper addresses this issue by developing an individual-level measure of outside options and uses it to test the empirical relevance of wage bargaining. Drawing on observed employer-to-employer transitions, I estimate the attractiveness of potential alternative destinations to each worker, allowing valuations to depend flexibly on worker and employer characteristics, so as to incorporate rich heterogeneity. I then aggregate these destination utilities into a measure of the expected value of alternative employment, grounded in discrete choice theory. I use the derived measure to test firms' wage setting conduct, showing that improvements in labor market outside options raise wages even within an ongoing employment spell, providing evidence that workers use opportunities elsewhere to negotiate higher pay from their current employer.

To construct this measure, I develop and flexibly estimate a simple nested-logit model of labor supply, drawing on the tradition of modelling workers' choices among differentiated employers (Manning, 2003; Card, Cardoso, et al., 2018). In the upper nest, workers choose between remaining with their current employer and leaving, and, conditional

on leaving, choose a destination employer in the lower nest. Under standard distributional assumptions on utility shocks, this structure yields a closed-form inclusive value summarising the expected maximum utility available from employers other than the incumbent (Train, 2009). The resulting measure, which I use as a proxy for outside options, captures the value of alternative employment by considering each potential destination employer individually, accounting for its geographical proximity, compatibility with the worker, and the value the worker attaches to its characteristics.

I apply the model to linked employer–employee data from the Veneto region of northeastern Italy over 1995–2000. I estimate destination choices using a neural multinomial logit model (Hruschka, Fettes, and Probst, 2004), replacing the standard linear utility specification with a neural network, while retaining the standard logit choice probabilities. This preserves the structural advantages of discrete choice models, including a closed-form expression for the expected maximum utility available from an individual's choice set. Further, to capture compatibility with potential destinations, I assign workers to eight types using a clustering approach based on their demographic and employment characteristics. This follows a growing literature in labor economics that reduces the dimensionality of worker and firm heterogeneity by representing it through a finite number of types (Bonhomme, Lamadon, and Manresa, 2019; Lentz, Piyapromdee, and Robin, 2023). For each worker and potential destination employer, I include the number and share of the destination's employees who belong to the same cluster as the worker. These measures help address an important difficulty in discrete choice models of labor supply: an employer may offer attractive jobs and be geographically close without being willing to hire a particular worker. The presence of employees with similar characteristics at a potential destination provides a proxy for the employer's demand for the worker's skill profile. Incorporating this information helps distinguish relevant employment opportunities from nearby firms whose workforce composition suggests a poor match, allowing the inclusive value to reflect both workers' preferences and employers' demand, and significantly improves predictive accuracy. I show that, in held-out data, the neural specification outperforms both a parametric multinomial logit with type-specific preference heterogeneity and a gradient-boosted tree classifier in ranking workers' destination employers.

To validate the measure, I then turn to examining workers' recovery from mass layoffs. Having been forced to leave their employer, displaced workers must draw on the alternative employment opportunities available to them, making their subsequent outcomes informative about the quality of their outside options. Workers with higher pre-displacement inclusive values return to employment more quickly and experience

smaller losses in earnings and employment. Their subsequent jobs are also closer to their former employer and more often in the same industry. These findings support interpreting the inclusive value as a measure of access to relevant alternative employment opportunities.

I then use my measure of outside options to analyse workers' reactions to changes in their external labor market prospects, as well as firms' wage-setting conduct. Better outside options predict greater employer mobility and higher wages, even within ongoing employment relationships. A one standard deviation improvement in outside options corresponds to a 1.8 percentage-point increase in the probability of changing employers in the following year, approximately 12.5% relative to the baseline rate. This finding is consistent with the nested logit structure of the model. Turning to wages, in my preferred specification, a one standard deviation improvement in outside options corresponds to approximately 1.4% higher wages. This specification includes worker–employer match fixed effects, which absorb permanent characteristics of the employment relationship, and employer–year fixed effects, which absorb common changes in the employer's wage policy and other shocks common to its workforce. The estimate therefore exploits differential changes in coworkers' outside opportunities while their respective employment relationships continue. The positive wage response provides evidence of wage bargaining: workers obtain higher pay as their alternatives improve, even after accounting for changes in their employer's wage policy. Estimated wage and mobility responses are larger for men than for women, complementing the evidence in Caldwell and Danieli (2024) on the role of outside opportunities in gender pay differences. Wage responses are also stronger among workers with shorter tenure and at growing firms, whereas mobility responses are larger among longer-tenure workers and at shrinking firms, suggesting that the scope for renegotiation depends on both the stage of the employment relationship and the incumbent employer's capacity to share rents.

To strengthen the causal interpretation of these findings, I then develop an instrumental variable strategy based on large-firm openings. I construct a shift-share instrument combining employment at newly opening firms with worker-type-specific transition probabilities between industry–province cells, computed on the five years preceding the beginning of my analysis sample. Openings thus receive greater weight when they occur in destination cells that historically attracted more workers of a given type from the same origin industry and province. The design thus exploits differences in how relevant the same openings are to different workers, with exposure weights fixed before the analysis period. With worker, employer and year fixed effects, the 2SLS estimates imply approximately 2.2–2.3% higher wages per standard deviation improvement in outside op-

tions, reinforcing the evidence of a causal wage increase in response to an improvement in a worker's external labor market prospects. The corresponding mobility estimates are also positive, although somewhat less precisely estimated, and close to the OLS coefficients in the preferred specification.

I then replicate the analysis using Portuguese linked employer–employee data for 2013–2019. The wage results extend to this different institutional setting and period: in the specification with worker–employer match and employer–year fixed effects, a one standard deviation improvement in outside options corresponds to approximately 1.4% higher wages, almost identical to the Veneto estimate. Linking these records to firm accounts also allows me to examine rent sharing more directly. Workers with better outside options obtain larger wage gains when their employer's value added per worker increases, providing further evidence that alternative employment opportunities strengthen workers' ability to share in the rents stemming from improvements in employer performance. This provides further evidence that outside options shape workers' bargaining power, and that bargaining plays a role in wage determination.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature and highlights this paper's contributions, and Section 3 describes the data and sample construction. Sections 4 and 5 present the conceptual framework and estimation procedure. Section 6 describes the estimated outside option measure, and Section 7 validates it using mass layoffs. Section 8 presents the wage and mobility results, Section 9 develops the instrumental-variable strategy, and Section 10 presents the Portuguese replication and extension. Section 11 concludes.

2 Related Literature

This paper contributes to several strands of the literature. I organize my discussion around three main themes: the measurement of outside options and their pass-through to wages, the role of bargaining and wage posting in determining wages, and methodological approaches combining discrete choice models with network-based analysis of labor markets.

2.1 Measuring Outside Options

A small but growing literature has sought to quantify workers' employment opportunities outside their current job. The most closely related paper is Caldwell and Danieli (2024). They develop an "outside options index" (OOI) that measures workers' alter-

natives based on the cross-sectional concentration of similar workers across job types, grounding their measure in a two-sided matching model. Their approach uses the distribution of workers with comparable characteristics to infer the breadth of each worker's option set, finding that differences in outside options explain approximately 20% of the gender earnings gap in Germany. My approach complements theirs by recovering outside opportunities from observed employer-to-employer transitions rather than cross-sectional employment patterns. I bring the analysis to the level of individual destination firms, estimating each worker's probability of choosing a particular employer conditional on a job-to-job move. Flexible neural estimation of destination utilities accommodates rich individual heterogeneity in labor market transition choices, allowing worker characteristics to interact with firm attributes and mobility frictions. The resulting inclusive value summarizes the attractiveness of alternative employers at the worker-year level.

Schubert, Stansbury, and Taska (2024) measure outside opportunities across occupations using observed occupational transitions. Their outside-occupation option index combines transition patterns with wages in alternative occupations, capturing opportunities that conventional occupation-specific labor market definitions exclude. Dodini et al. (2024) take a complementary task-based approach, measuring employer concentration across jobs with similar skill requirements rather than within conventional industry or occupation boundaries.

More recently, Barrela (2026) identifies labor markets from observed job mobility and measures the pay quality of outside opportunities through an employment-weighted average of alternative employers' wage premia, excluding the current employer. My approach instead estimates the attractiveness of individual destination firms to each worker, allowing worker characteristics to interact flexibly with firm attributes and mobility frictions. Thus, a given firm's attractiveness as a potential destination need not depend solely on its wage policy, but can reflect any dimension of the worker–firm match captured by the observed characteristics and their interactions.

2.2 Pass-Through of Outside Options to Wages

A key question in labor economics is whether improvements in workers' outside options translate into higher wages at their current employer. The theoretical prediction from bargaining models is clear: better outside options should strengthen workers' bargaining position and raise wages. However, the empirical relevance of bargaining for wage setting has been called into question.

Jäger, Schoefer, et al. (2020) provide a direct test using quasi-experimental variation from unemployment insurance reforms in Austria. Surprisingly, they find that wages are largely insensitive to changes in the value of nonemployment: a $1 increase in UI benefits translates to less than $0.01 in wage gains for employed workers. This finding presents a puzzle for standard Nash bargaining models, which would predict much larger elasticities. In my context, I define outside options based on transition probabilities to other employers, so that the value of unemployment does not enter the choice set of a worker.

Earlier evidence from Beaudry, Green, and Sand (2012) shows that changes in local industrial composition affect wages across sectors through the employment opportunities available to workers.

Caldwell and Harmon (2019) demonstrate that information about outside options flows through coworker networks. When former coworkers experience increased labor demand at their current firms, incumbent workers see higher wages and greater job-to-job mobility, suggesting that workers who have better information about their outside options are able to capture some of their value.

Lachowska, Mas, Saggio, et al. (2022) use dual jobholders to show that improvements in outside options raise separations among lower-paid workers but incumbent wages among workers in the highest wage quartile. Caldwell, Haegele, and Heining (2025) provide linked survey and administrative evidence that individual wage bargaining is widespread and that differences in bargaining behavior contribute to within-firm wage inequality. Ek et al. (2025) show that better outside opportunities increase the wage return to match quality among incumbent workers, consistent with workers using outside offers to extract a larger share of match-specific surplus.

Jäger, Roth, et al. (2024) offer a complementary perspective: using linked survey-administrative data from Germany, they show that workers systematically underestimate their outside options, particularly at low-wage firms. This suggests that incomplete pass-through may partly reflect workers' failure to leverage options they objectively possess, and grants employers monopsony power.

A related literature examines how labor market concentration affects wages - a very concentrated labor market is one in which workers have fewer outside options, all things equal. Azar, Marinescu, and Steinbaum (2022) document substantial employer concentration in U.S. labor markets using online vacancy data, finding that the average market has an HHI of approximately 4,378 - well above thresholds typically considered concentrated. They show that higher concentration is associated with lower wages. Jarosch, Nimczik, and Sorkin (2024) develop a theoretical framework where large employers ex-

ercise market power by effectively removing their own vacancies from workers' outside options. In their "granular search" model, firm size directly translates into bargaining power, and they estimate that this mechanism depresses wages by approximately 10% in Austria. My measure incorporates a related mechanism by excluding the current employer from the set of alternative destinations. Holding the worker's characteristics and the rest of the opportunity set fixed, employment at a larger firm removes a larger potential source of outside employment from that set; the effect on the inclusive value depends on how attractive that employer would otherwise be to the worker.

My contribution to this literature is to show that employment-to-employment outside options do causally raise wages. Using mass layoffs as a validation exercise and large firm openings as an instrumental variable, I find that workers with better job-to-job mobility prospects earn significantly higher wages at their current employer, thus implying that bargaining does matter for wage setting, although the effect - one standard deviation higher outside option leads to a 1.4% higher wage in my preferred specification - is economically relatively small.

2.3 Bargaining, Wage Posting, and the Role of Firms

The seminal work of Abowd, Kramarz, and Margolis (1999) demonstrated the importance of firms for wage dispersion through a decomposition of log wages into additive worker and firm components. Outside options, in my framework, depend jointly on individual characteristics and the current employer, thus varying at the worker–firm (ij) level in a way that need not decompose into separate worker and firm effects. The positive impact of this measure on wages therefore calls into question log additive separability as an accurate description of wage determination. This relates to the broader discussion of the AKM restrictions reviewed by Kline (2024), and to Sorkin and Warwar (2026), whose paired-mover tests provide evidence against additive separability. My analysis identifies worker-specific outside opportunities as an economic source of such nonseparable wage variation. Nevertheless, the modest economic magnitude of the estimated wage response suggests that log additive separability may remain a useful first approximation to wage determination.

Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay, and Robin (2006) develop equilibrium models of on-the-job search in which competition over outside offers shapes wages and workers' progression across employers. These models provide a theoretical foundation for the role of alternative employment in wage determination. My analysis shares their emphasis on employment opportunities beyond the current match, but fo-

cuses on measuring their value from observed destination choices rather than estimating a full dynamic model of search and wage bargaining.

A recent contribution by Di Addario et al. (2023) extends the AKM framework to include "origin effects"—fixed effects for the firm from which a worker was hired. Using Italian administrative data, they find that origin effects explain only 0.7% of the variance in hiring wages, while destination effects explain over 23%. This suggests that where a worker comes from is relatively inconsequential compared to where they currently work. In a sense, my paper asks the opposite question: I look at how wages are impacted by where one *can* go.

Bonhomme, Lamadon, and Manresa (2019) develop a framework for matched employer–employee data that allows for two-sided unobserved heterogeneity and complementarities in earnings, combining a first-stage classification of firms with latent worker types. My approach also uses a discrete classification of workers, but for a different purpose: worker types help characterize potential destinations through the presence of workers with similar characteristics. Their employment at a destination provides a proxy for that employer's demand for workers like the individual considering a move, or for the compatibility between the worker's characteristics and the jobs available there. Types therefore enter the construction of destination attributes without restricting outside options to vary only at the type level.

In a related paper, Abowd, McKinney, and Ian M Schmutte (2019) address endogenous mobility directly: using U.S. LEHD data and a Bayesian latent-type framework, they show that the exogenous mobility assumption underlying AKM is rejected and that correcting for it increases the share of earnings variance attributed to worker heterogeneity while reducing the role of firm effects.

2.4 Discrete Choice Models of Labor Markets

Discrete choice models provide a framework for understanding labor supply when workers view employers as imperfect substitutes. Manning's (2003) *Monopsony in Motion* establishes the broader role of labor market frictions and employer differentiation in giving firms wage-setting power, a perspective developed further in Manning (2011). Card, Cardoso, et al. (2018) present a discrete choice framework in which heterogeneous preferences over employers generate upward-sloping labor supply to individual firms, linking firm productivity and wage policies to wage inequality. Drawing on the differentiated-product demand estimation tradition, Azar, Berry, and Marinescu (2022) estimate work-

ers' choices among job vacancies using application data, recovering own-wage application elasticities and their implications for labor supply to individual employers.

A related literature incorporates employer choice into equilibrium models with heterogeneous workers and firms. Berger, Herkenhoff, and Mongey (2022) study how imperfect substitution across employers and strategic interaction between granular firms shape wages and labor market power. Berger, Herkenhoff, Jeong, et al. (2026) explicitly incorporate discrete employer choice alongside heterogeneity in worker characteristics and labor supply elasticities, showing how workforce composition and sorting influence firms' wage setting. Lamadon, Mogstad, and Setzler (2022) estimate an equilibrium model with two-sided heterogeneity in which workers' preferences over nonwage job characteristics make employers imperfect substitutes, allowing them to quantify compensating differentials, rent sharing, and labor market rents.

Closely related work by Friederiszick, Malmcrona, and Sunuwar (2026a) and Friederiszick, Malmcrona, and Sunuwar (2026b) models workers' labor supply to individual jobs, defined as occupations within firms, and uses the resulting substitution patterns to characterize workers' alternative employment opportunities. Their framework allows them to study how mergers and mass layoffs affect both directly affected workers and other workers whose relevant employment alternatives change.

Observed mobility provides information about how workers value these differentiated employers. Sorkin (2018) uses employer-to-employer transitions to recover a revealed-preference ranking of firms, showing that wage differences alone do not account for their attractiveness to workers. This approach establishes a link between the structure of worker flows and the valuation of employment opportunities. Roussille and Scuderi (2025) similarly recover firms' nonwage attractiveness from workers' choices over observed job menus, using the resulting amenity rankings to distinguish between models of wage-setting conduct.

Drawing on this discrete choice tradition, I shift the focus from labor supply to a given employer to the value of employment at alternative firms. Under the standard assumptions of the model, the discrete choice structure implies a welfare measure of the available alternatives—the inclusive value. I use this implication to recover a worker-specific measure of the value of outside employment and study its role in wage determination and mobility.

2.5 Mobility Networks and Labor Market Boundaries

A final strand of literature uses observed mobility to recover the structure and boundaries of labor markets. Ian M. Schmutte (2014) applies community detection methods to job transitions to identify labor market segments, while Nimczik (2020) uses stochastic block models to group firms according to their mobility patterns. These approaches allow relevant labor markets to emerge from observed connections between jobs rather than imposing boundaries based on conventional geographic, industry, or occupational classifications.

A complementary approach emphasizes that workers' opportunities can overlap without belonging to a single, mutually exclusive market. Manning and Petrongolo (2017) develop a spatial job-search model in which the attractiveness of employment opportunities declines with distance, generating overlapping local labor markets. Datta (2024) develops and estimates a spatial job-search model in which commuting costs generate probabilistic, overlapping labor markets, highlighting substantial variation in relevant employment alternatives within conventional geographic boundaries.

Network connections also shape access to subsequent employment opportunities. Bocquet (2024) models skill-related mobility frictions through an occupational network constructed from expert assessments of skill transferability, showing how bridge occupations connecting distinct skill clusters shape workers' opportunities and the speed of labor reallocation. Del Prato (2024) studies how firms' connectivity in the employer-to-employer mobility network affects external offer arrival, generating a "springboard" value that workers trade off against current pay.

My approach shares this literature's emphasis on employment opportunities that depend on workers' characteristics and their position in the labor market. I use observed employer-to-employer choices to estimate the attractiveness of alternative firms to each worker, allowing workers employed at the same firm to face different valuations of outside employment rather than assigning them a common market-level measure.

3 Data and Sample

The main analysis uses the Veneto Worker Histories (VWH), a matched employer–employee dataset constructed from the administrative records of the Italian Social Security Institute (INPS).[1] The data cover private-sector employees outside agriculture

[1]Dataset developed by the Economics Department in Università Ca' Foscari Venezia under the supervision of Giuseppe Tattara.

who worked at establishments located in Veneto, northeastern Italy, during 1975–2001, regardless of their place of residence. VWH has been used extensively to study wage determination, worker–firm sorting, and labor mobility, including by Card, Devicienti, and Maida (2014), Bartolucci, Devicienti, and Monzón (2018), Serafinelli (2019), Kline, Saggio, and Sølvsten (2020), and Kline (2024). The records link workers to employers and report employment spells, earnings, paid weeks, worker demographics, and employer characteristics, including industry and location. I use records from 1990 to 2001 to construct employment histories and estimation variables. The main analysis covers 1995–1999, with employment outcomes observed through 2000. Transitions during 1990–1994 provide pre-period measures of mobility between industries, while transitions in 2000 are held out to assess the model's predictive performance.

Manufacturing accounts for approximately 59% of employment among eligible employers in the Veneto sample during 1995–1999. Appendix Table A1 reports the employment distribution across the ten largest industries, alongside employer-year counts and average employer size. Small employers are prevalent: the median employer has 10 employees, approximately 76% of employer-year observations have fewer than 20, and only 6.9% have at least 50. Nevertheless, employers with more than 100 employees account for 33% of employment, despite representing only 2.7% of employer-year observations (Appendix Figure A1). Employment opportunities also extend beyond workers' municipalities of residence. Among Veneto-resident worker-year observations, the share working outside their municipality ranges from approximately 59% in the province of Rovigo to 72% in the province of Padova. The share with a residence-to-workplace distance exceeding 30 km ranges from 5.5% in the province of Verona to 15.7% in the province of Rovigo (Appendix Figure A2). These patterns motivate accounting for both employer size and geographic proximity when characterizing alternative employment opportunities.

The raw data comprise linked worker, employer, and annual employment records. Worker records report gender, date and place of birth, nationality, and municipality and province of residence. Employment records identify the employer and contribution year and report earnings, paid months, weeks and days, contract type, contractual grade, employment qualification, and the cessation date where applicable. Employment qualification identifies broad categories such as blue-collar workers, white-collar workers, managers, and apprentices, rather than educational attainment. Importantly, the data do not contain detailed occupational codes identifying the specific occupation performed within a firm. Reported earnings are gross of taxes, include cash benefits, exclude benefits in kind, and are supplied at constant 2003 prices. Employer records contain industry clas-

sifications, address and municipality, and information on establishment, suspension, and cessation of activity.

I reconstruct employment histories at monthly frequency using the months of activity reported in the employment records. When a worker holds multiple jobs in a month, I designate the job with the highest weekly wage as the primary job. I measure wages as average real weekly earnings, calculated by dividing the supplied annual job earnings at 2003 prices by reported paid weeks, and winsorize this measure at the 0.5th and 99.5th percentiles. For the wage and mobility analysis, I construct an annual panel with one observation per worker and year, assigning each worker to the employer paying the highest weekly wage; ties are resolved by tenure and then employer identifier. This annual primary-job panel identifies the worker's current employer and wage, while the monthly histories distinguish employer transitions with and without intervening months of observed nonemployment.

To estimate workers' choices among alternative employers, I use employer transitions between adjacent calendar months, without intervening months of observed nonemployment. I restrict the sample to workers residing in Veneto for whom the worker and origin-employer characteristics used in estimation are available. Potential destination employers must have at least five employees, and the realized destination must be eligible in both the origin and arrival years. These restrictions yield 449,335 transitions: 358,412 with arrival years between 1995 and 1999, used for model estimation, and 90,923 with arrival year 2000, reserved for predictive validation. This sample captures destination choices conditional on an observed employer transition; the annual panel used to study wages and mobility also includes workers who remain with their employer. Appendix Table A2 details the sample restrictions.

From the linked employment records, I construct measures of workers' employment histories to capture differences in labor market attachment and accumulated experience that may shape their destination choices. I measure tenure as the number of years since the worker's first observed employment at the current firm. Prior employment history is summarized by an indicator for any month without recorded employment in the preceding two years, the number of such months in the preceding five years, and cumulative months of recorded employment, all measured before January of the worker's first observed year at the current employer. Since absence from the records does not distinguish unemployment from inactivity or employment outside the data's coverage, I restrict the destination-choice sample to employer-to-employer transitions identified by observing a worker at two different employers in consecutive calendar months. At the employer level, I construct employment, employment growth, and workforce composition from

the linked worker records. I also use geographic information to measure distances between workers' residences and potential employers, and transitions during 1990–1994 to characterize historical mobility between industries.

Drawing upon a growing literature that represents worker and firm heterogeneity through a finite number of types (Bonhomme, Lamadon, and Manresa, 2019; Bonhomme, Lamadon, and Manresa, 2022; Lentz, Piyapromdee, and Robin, 2023), I group workers into eight types to characterize employers' workforce composition. I use the k-prototypes algorithm of Huang (1998), which accommodates both numerical and categorical characteristics. The clustering uses age, gender, nationality, contract category, employment qualification, tenure, log wages, and wage growth, with each worker represented by their first usable observation during 1995–2001. I select eight types using the elbow criterion, balancing a reduction in within-group dissimilarity against the complexity of a finer partition. I then construct the number and share of workers of each type employed at each firm in each year. These measures describe how closely a potential destination's workforce resembles the worker considering that destination, providing a proxy for employers' demand for different worker profiles. The types supplement the individual characteristics entering the destination-choice model; they do not restrict workers of the same type to have identical destination preferences. Table 1 summarizes the characteristics of the eight types, while Table 2 describes their observed mobility patterns.

The resulting types distinguish workers at different stages of their careers and with different employment qualifications. Table 1 identifies two young groups, types 4 and 8, whose modal employment qualifications are apprentice and blue-collar, respectively, alongside older workers with longer tenure, particularly type 5. Gender also differentiates the groups: types 6 and 7 are almost exclusively female and male, respectively. These profiles are accompanied by differences in observed mobility. In Table 2, average distances between origin and destination employers range from approximately 12 to 17 km, while the share of transitions within the same two-digit industry ranges from approximately 32% for types 3 and 8 to 61% for type 5. These differences motivate allowing destination choices to vary flexibly with individual characteristics and their interaction with potential employers.

These patterns also illustrate, in this context, the limitations of defining workers' opportunities through coarse labor market boundaries. Studies linking labor market characteristics, such as employer concentration, to workers' outcomes often define markets by combining occupational and geographic categories, as in Marinescu, Ouss, and Pape (2021). In the absence of detailed occupational codes, an analogous definition here would combine industry and province. Yet, in the destination-model training sample, 57.6% of

employer transitions cross two-digit industry boundaries, even before imposing a geographic restriction. An industry–province definition would therefore place more than half of realized transitions outside the worker's origin market. Treating that market as the worker's choice set would exclude many employers that workers demonstrably move to, misrepresenting the set of relevant alternative employers.

The Portuguese analysis uses Quadros de Pessoal (QP), an annual matched employer–employee dataset reporting employment and remuneration in October. The data link workers to firms and establishments and provide information on workers' demographics, education, occupation, contractual status, hours, and pay, alongside employer characteristics such as industry and employment. Unlike VWH, QP contains detailed occupational codes. Geographic information is considerably coarser, however: the available extract identifies establishment and employer locations at the NUTS II regional level, rather than providing the municipality-level information available for Veneto. I use records from 2010 onward to construct employment histories and pre-period mobility patterns, with the main analysis covering 2013–2019. When a worker has multiple jobs in the October snapshot, I retain the job with the highest total monthly earnings, breaking ties by normal working hours and then employer identifier. Monthly earnings comprise base pay, regular bonuses and allowances, and overtime remuneration. The wage comparison with Veneto uses log monthly earnings in nominal euros, winsorized at the pooled 0.5th and 99.5th percentiles. Employer transitions are identified by changes in the primary employer between consecutive annual observations; these may include intervening periods without employment. For the rent-sharing analysis, I link QP to firm accounts from the Sistema de Contas Integradas das Empresas (SCIE), which provide information on value added, employment, and capital.

Table 1: Summary Statistics by Worker Type

Type	Workers	Mean tenure (years)	Mean log weekly earnings	Mean change in log earnings	Mean age (years)	Male share	Modal employment qualification	Transitions per employment record
1	166,695	4.3	6.42	0.019	37	0.15	2	0.11
2	238,505	4.3	6.60	0.016	43	0.83	1	0.13
3	116,132	2.8	6.58	0.036	28	0.69	2	0.17
4	99,916	1.8	6.15	0.050	22	0.07	5	0.26
5	154,663	4.5	6.56	0.019	54	0.71	1	0.09
6	209,916	3.2	6.21	0.005	28	0.00	1	0.18
7	316,525	3.2	6.49	0.025	29	1.00	1	0.20
8	93,382	1.8	6.21	0.061	20	0.98	1	0.30

Notes: Statistics use reference employment records in 1995–1999 with an assigned worker type. The Workers column counts distinct individuals; means and shares weight employment records, which may include multiple employers within a worker-year. Tenure is the observation year minus the first observed year at the employer. Weekly earnings are reported job-year earnings at constant 2003 prices divided by paid weeks, winsorized at the 0.5th and 99.5th percentiles; the table reports the mean natural logarithm. The change in log earnings is the difference from the worker's preceding record after sorting by worker and year. It is not restricted to adjacent years and can compare records within the same year; it should not be interpreted as annual wage growth. Male share is a proportion. Modal employment qualification is the most frequent broad qualification category: 1 = blue-collar, 2 = white-collar, 5 = apprentice; these are not detailed occupational codes. The final column divides observed employer transitions, including those with intervening months without covered employment, by reference employment records for the type. It does not identify voluntary quits or the probability of quitting. Worker types are constructed using the characteristics and clustering procedure described in the text.

Table 2: Mobility by Worker Type: Destination-Model Training Sample

Type	Transitions	Mean employer-to-employer distance (km)	Same 2-digit industry share	Same 1-digit industry share	Firms in mobility network	Directed network density
1	28,460	13.84	0.497	0.669	17,036	0.000064
2	57,530	13.70	0.508	0.660	26,082	0.000058
3	29,409	16.78	0.322	0.573	16,611	0.000084
4	24,647	12.78	0.364	0.555	19,052	0.000059
5	19,349	14.52	0.607	0.736	12,701	0.000074
6	53,094	12.89	0.467	0.659	22,668	0.000075
7	116,258	14.00	0.380	0.541	36,005	0.000071
8	29,665	11.51	0.321	0.496	20,613	0.000063

Notes: The sample is the 358,412 employer-to-employer transitions used to estimate destination choices, with arrival years 1995–1999. It includes Veneto residents observed at different primary employers in adjacent calendar months, with the worker and employer information required by the model. Origin employers must be eligible in the origin year; realized destinations must be eligible in both the origin and arrival years. Eligible employers have at least five employees. Worker types are assigned using the clustering described in the text. Mean distance is the great-circle distance between origin employer coordinates in the origin year and destination employer coordinates in the arrival year; 347,839 transitions have usable coordinates. Industry shares are fractions of transitions within the same two-digit or one-digit ATECO 1981 industry, using employer attributes at these respective dates. Firms in the mobility network counts distinct origin or destination employers for each type over the sample period. Directed network density is $m/[n(n-1)]$, where n is this firm count and m is the number of distinct directed origin–destination pairs, excluding self-links. Repeated transitions along a pair count once for density and separately for transition means and shares. Firms can appear in multiple type-specific networks; their union contains 55,106 firms. Shares and densities are proportions, not percentages.

4 Conceptual Framework

Consider a worker i employed at firm ℓ in period t. Let $\mathcal{J}_t$ denote the set of all firms in the economy, and $\mathcal{J}_{it} = \mathcal{J}_t \setminus \{\ell\}$ the set of alternative employers. I represent the worker's employment choice using a nested-logit structure. At the upper level, the worker chooses between remaining with the current employer (*Stay*) and moving to another firm (*Leave*). At the lower level, conditional on leaving, the worker chooses a destination $j \in \mathcal{J}_{it}$. This sequential representation does not require the decisions to occur sequentially in time. Figure 1 illustrates this structure.

While encompassing standard multinomial logit as a special case, this framework can capture unobserved switching costs through shocks shared across alternative employers, distinguishing staying with the incumbent from joining an otherwise observationally identical firm.

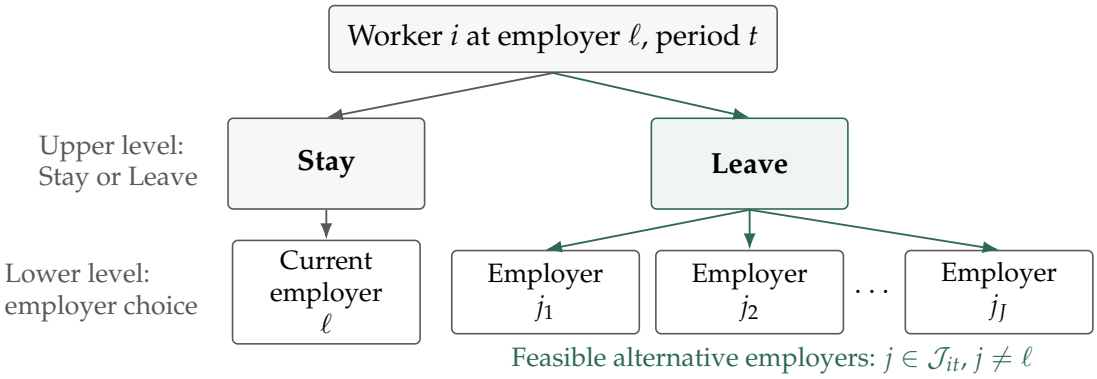


Figure 1: Nested Structure of Employer Choice

Within the Leave nest, utility associated with destination j is

$$U_{i\ell jt} = u_{i\ell jt} + \varepsilon_{i\ell jt}, \tag{1}$$

where $u_{i\ell jt}$ is systematic utility and $\varepsilon_{i\ell jt}$ is an idiosyncratic random utility shock. I assume that utility shocks follow the joint generalized extreme-value distribution underlying nested logit, with type-I extreme-value marginals (Train, 2009). The dissimilarity parameter σ governs dependence within the Leave nest: smaller values imply stronger correlation across alternative employers' utility shocks, while $\sigma = 1$ corresponds to independence. These shocks remain independent of the shock associated with staying at the current firm. With the upper-level scale normalized to one, the standard random-utility specification requires $0 < \sigma \leq 1$. I estimate the inclusive-value coefficient without

imposing this restriction. In Section 8, I show that the estimates lie within this interval, consistent with the nested-logit structure of the decision problem.

Following the discrete choice approach to employer differentiation discussed by Card, Cardoso, et al. (2018), systematic utility can reflect remuneration, amenities, mobility costs, and other determinants of employment choices. At this stage, I impose no restriction on how $u_{i\ell jt}$ depends on the worker i, destination employer j, current employer ℓ, or period t. Section 5 describes how I reconstruct these utilities from observed employer transitions.

Conditional on Leave, destination probabilities take the multinomial-logit form

$$P_{i\ell jt|\text{Leave}} = \frac{\exp(v_{i\ell jt})}{\sum_{\substack{k\in\mathcal{J}_{it}\\ k\neq\ell}} \exp(v_{i\ell kt})}, \qquad v_{i\ell jt} \equiv \frac{u_{i\ell jt}}{\sigma}. \tag{2}$$

The discrete choice structure also provides a welfare measure of the value of alternative employment. Under this generalized extreme-value assumption, expected maximum utility over alternative employers is, as shown by McFadden (1981),

$$\begin{aligned} V^{\text{out}}_{i\ell t} &\equiv \mathbb{E}\left[\max_{\substack{j\in\mathcal{J}_{it}\\ j\neq\ell}} U_{i\ell jt}\right] \\ &= \sigma \log\left(\sum_{\substack{j\in\mathcal{J}_{it}\\ j\neq\ell}} \exp(v_{i\ell jt})\right) + \gamma_{\text{EM}}, \end{aligned} \tag{3}$$

where $\gamma_{\text{EM}} \simeq 0.5772$ is the Euler–Mascheroni constant. Define the inclusive value in within-nest utility units as

$$I_{i\ell t} = \log\left(\sum_{\substack{j\in\mathcal{J}_{it}\\ j\neq\ell}} \exp(v_{i\ell jt})\right). \tag{4}$$

The expected maximum provides a model-based measure of the worker's employment outside option. The inclusive value $I_{i\ell t}$ summarizes this object in within-nest utility units, with $V^{\text{out}}_{i\ell t} = \sigma I_{i\ell t} + \gamma_{\text{EM}}$. The Euler–Mascheroni constant remains part of the expected-maximum expression. However, because the inclusive value is standardized in the subsequent empirical analysis, this common additive constant will effectively drop out. The

measure reflects both the availability and the attractiveness of alternative employers and depends on individual characteristics and the current employer, whose identity affects destination utilities and determines which employer is excluded from the Leave nest.

I interpret observed direct employer-to-employer transitions as choices among alternative employment opportunities, maintaining that these transitions reflect voluntary mobility. Accordingly, I define outside options over alternative employment opportunities. Unemployment and inactivity are excluded from the destination set, and I do not estimate their utility or their contribution to the value of leaving the current employer. The resulting measure therefore characterizes the employment component of workers' outside options.

With the upper-level error scale normalized to one, the probability of Leave is

$$\Pr(\text{Leave} \mid i, \ell, t) = \Lambda(A_{i\ell t} + \sigma I_{i\ell t}), \tag{5}$$

where $\Lambda(\cdot)$ is the logistic distribution function and $A_{i\ell t}$ collects the systematic components common to Leave relative to Stay, potentially including common switching costs.

This structure admits sequential estimation: conditional destination choices are estimated first, and the resulting inclusive value enters the upper-level Stay/Leave equation. Under the maintained model assumptions, this procedure is consistent, although less efficient than joint maximum likelihood (Train, 2009). The objective of this paper, however, is not the full joint structural estimation of the mobility decision. Rather, I use the destination-choice structure to measure the value of alternative employment and examine how changes in this value influence workers' subsequent labor market trajectories. Section 5 describes the construction, normalization, and calibration of the estimated inclusive value.

The framework yields an unambiguous prediction for employer mobility, holding the value of the current employment relationship and switching costs fixed: an increase in the value of alternative employment raises the probability of leaving. Writing $p_{i\ell t} = \Lambda(A_{i\ell t} + \sigma I_{i\ell t})$,

$$\left.\frac{\partial p_{i\ell t}}{\partial I_{i\ell t}}\right|_{A_{i\ell t}} = \sigma p_{i\ell t}(1 - p_{i\ell t}) > 0. \tag{6}$$

The corresponding prediction for wages depends on the model of wage determination. To illustrate this distinction, consider two simple wage-setting benchmarks.

Under wage posting, the employer specifies a wage schedule that does not respond to individual workers' outside opportunities. Conditional on this schedule and the worker's pay-relevant characteristics, an improvement in the worker's outside option therefore

leaves their wage unchanged. Adjustment occurs through the worker's decision to remain with or leave the employer.

Under wage bargaining, outside opportunities can instead affect the wage within the current employment relationship. Suppressing indices, suppose that the worker's utility from agreement is $\alpha w + a$, where $\alpha > 0$ and a captures nonwage value, while disagreement delivers the expected value of alternative employment, $V^{\text{out}} = \sigma I + \gamma_{\text{EM}}$. Let match revenue be y, normalize the employer's disagreement payoff to zero, and denote the worker's bargaining weight by $\beta \in (0, 1)$. The Nash bargain solves

$$\max_{w} \left(\alpha w + a - V^{\text{out}}\right)^{\beta} (y - w)^{1-\beta}. \tag{7}$$

For a match with positive surplus, this yields

$$w^N = \beta y + (1 - \beta)\frac{V^{\text{out}} - a}{\alpha}, \qquad \frac{\partial w^N}{\partial I} = \frac{(1-\beta)\sigma}{\alpha} > 0. \tag{8}$$

Holding productivity, nonwage value, and bargaining power fixed, better outside options therefore raise negotiated wages. I use the estimated inclusive value to test these contrasting predictions about wage-setting conduct.

5 Estimating Outside Options

To construct an empirical counterpart to the value of alternative employment defined in Section 4, I first predict workers' destination choices conditional on leaving their current employer. This corresponds to estimating the lower nest of the decision problem. The prediction task is to assign a probability to each potential destination firm, allowing these probabilities to depend flexibly on the worker's characteristics, their current employer, and the attributes of alternative employers. I estimate a neural multinomial-logit model using observed employer-to-employer transitions, with predicted destination probabilities

$$\widehat{P}_{i\ell jt|\text{Leave}} = \frac{\exp(\widehat{v}_{i\ell jt})}{\sum_{\substack{k \in \mathcal{J}_{it} \\ k \neq \ell}} \exp(\widehat{v}_{i\ell kt})}, \tag{9}$$

where $\widehat{v}_{i\ell jt}$ denotes the model's fitted destination utility in within-nest units, corresponding to $v_{i\ell jt} = u_{i\ell jt}/\sigma$ in Section 4.

I then aggregate these predicted utilities into an inclusive value,

$$\widehat{I}_{i\ell t} = \log\left(\sum_{\substack{j\in\mathcal{J}_{it}\\ j\neq\ell}} \exp(\widehat{v}_{i\ell jt})\right). \tag{10}$$

The sum is taken over all employers except the current one, so as to reconstruct the ex ante expected maximum utility from leaving one's current employer, as characterized in Section 4. The resulting inclusive value provides an individual measure of the value of alternative employment for each worker-year.

To model destination utilities, I replace the usual linear index of the multinomial logit with a neural network. This follows a growing literature combining discrete choice models with neural utility specifications, including Aouad and Désir (2025) and S. Wang, Q. Wang, and Zhao (2020). The network learns a flexible mapping from observed characteristics to destination utilities, while retaining the multinomial-logit formula for conditional choice probabilities. This approach preserves the familiar discrete choice structure, including the welfare interpretation of the inclusive value under the maintained assumptions, while allowing substantially greater flexibility in the functional form of utility.

Figure 2 provides a schematic overview of the neural destination-choice model. Consider first the vector x_{jt} of nineteen destination-employer characteristics, including firm size, employment growth, industry indicators, and the AKM firm wage effect. Appendix Table A3 lists all inputs and details their construction. These characteristics are common to all workers considering the same employer in a given year.

Next, consider the vector $m_{i\ell jt}$ of five characteristics linking the worker and their current employer to a potential destination. These include geographic distance, an indicator for whether the origin and destination employers are in the same province, historical flows between their industries, and the number and share of workers of the individual's type employed at the destination. Historical industry flows are constructed from transitions observed in 1990–1994, before the estimation period. Unlike destination characteristics alone, these inputs describe how a given destination relates to the worker considering it.

Type compatibility helps address a general concern in discrete choice models of labor supply: observed employment choices reflect both workers' preferences and employers' willingness to hire them. An attractive employer need not be an attainable destination, because workers may face hiring constraints that are unobserved in employment records (see the discussion in Hirsch et al., 2022). The number and share of same-type workers

at each destination provide proxies for this compatibility: they indicate whether the firm employs workers with similar observed characteristics. Including these measures allows the model to assign lower utility scores to destinations with weaker compatibility, reducing their contribution to the inclusive value. Worker types thus help capture differences in employment opportunities arising from labor demand, alongside differences in workers' preferences.

Stacking x_{jt} and $m_{i\ell jt}$ produces a vector of twenty-four observed destination and pair characteristics. Workers may value these characteristics differently depending on their own attributes and their current employment relationship. The vector $z_{i\ell t}$, containing forty-four worker and current-employer characteristics, allows the model to capture this heterogeneity. It includes demographic characteristics, employment qualification, earnings, reconstructed employment histories, and attributes of the current employer.

Allowing $z_{i\ell t}$ to enter an unrestricted utility network directly could introduce a component common to every destination considered by the same worker. To see the problem, suppose that

$$v_{i\ell jt} = \mu(z_{i\ell t}) + g_{i\ell jt}. \tag{11}$$

The common component cancels from destination probabilities:

$$P_{i\ell jt|\text{Leave}} = \frac{\exp\big(\mu(z_{i\ell t}) + g_{i\ell jt}\big)}{\sum_{k\in\mathcal{J}_{it}} \exp(\mu(z_{i\ell t}) + g_{i\ell kt})} = \frac{\exp(g_{i\ell jt})}{\sum_{k\in\mathcal{J}_{it}} \exp(g_{i\ell kt})}. \tag{12}$$

It nevertheless enters the inclusive value:

$$I_{i\ell t} = \mu(z_{i\ell t}) + \log\left(\sum_{j\in\mathcal{J}_{it}} \exp(g_{i\ell jt})\right). \tag{13}$$

Consequently, the model could assign different values of alternative employment to workers without changing any predicted destination probability. Destination-choice fit alone therefore cannot discipline this common component, although it affects the measure subsequently related to wages and mobility.

If I allowed worker and current-employer characteristics to enter the neural network freely, it could therefore learn worker-specific utility components that leave destination probabilities unchanged but shift the inclusive value. To tackle this issue, I channel these characteristics through a bilinear matching layer, drawing on the low-rank bilinear pooling approach of Kim et al. (2017). The model maps the worker/current-employer vector and the stacked destination/pair vector into two twelve-dimensional representations, $h^z_{i\ell t}$

and $h^b_{i\ell jt}$, each formed from learned weighted combinations of its inputs. Corresponding components are then multiplied:

$$\iota_{i\ell jt,r} = h^z_{i\ell t,r}\, h^b_{i\ell jt,r}, \qquad r = 1, \ldots, 12. \tag{14}$$

This architecture captures heterogeneity in the valuation of destination attributes through a compact set of learned interactions, rather than estimating a separate coefficient for every worker–destination characteristic pair. Worker and current-employer characteristics thus enter through their interactions with destination attributes, without entering the final utility network directly. For example, women may place greater weight on geographic proximity, consistent with the evidence in Le Barbanchon, Rathelot, and Roulet (2021), while workers' industry of origin may influence the attractiveness of destinations with related skill requirements. The network learns these patterns from observed transitions without prespecifying how the valuation of each destination attribute varies across workers.

Finally, the nineteen destination characteristics, five pair characteristics, and twelve matching interactions form a thirty-six-dimensional input to f_θ. This function is a multilayer perceptron: a neural network that passes its inputs through successive layers of learned weighted combinations and nonlinear transformations. Here, two hidden layers of twenty units, using exponential linear unit activations, map the thirty-six inputs into a single predicted destination utility:

$$v_{i\ell jt} = f_\theta\left(x_{jt},\, m_{i\ell jt},\, \iota_{i\ell jt}\right). \tag{15}$$

I estimate the network parameters by minimizing a regularized cross-entropy loss for observed destination choices. For computational tractability, each training choice set contains the realized destination and 199 sampled alternatives, with a correction for their unequal sampling probabilities, following the importance-sampling approach to neural probability models of Bengio and Sénécal (2003). The loss penalizes the model when it assigns a low probability to the observed destination. The worker and destination projections, matching interactions, and final utility network are learned jointly from this prediction objective. Appendix E.1 details the sampling procedure, training objective, optimization, and model selection.

My objective is to reconstruct destination utilities and the value of alternative employment they imply, rather than to identify a separate effect of every observed characteristic. In fact, it is easy to prove (see Appendix E.2) that, for a given destination set, $|\widehat{I}_{i\ell t} - I_{i\ell t}| \leq \max_{j \in \mathcal{J}_{it}} |\widehat{v}_{i\ell jt} - v_{i\ell jt}|$. The difference between the estimated and true

inclusive values cannot exceed the largest destination-utility prediction error. Accurate reconstruction of the utility function would therefore deliver an accurate approximation to the value of the lower nest, without requiring a unique or causal interpretation of every network parameter.

The emphasis on prediction motivates the flexible neural specification, which allows nonlinear relationships and heterogeneous valuations of destination attributes to be learned from observed transitions. I assess its predictive performance using held-out employer transitions in year 2000, excluded from the data used to train the model. Figure 3 compares the neural model with a type-specific multinomial logit and an XGBoost classifier. The type-specific multinomial logit accommodates preference heterogeneity by allowing each worker type to have its own coefficients on destination and pair characteristics. XGBoost provides a flexible prediction benchmark based on boosted decision trees, without imposing the bilinear matching architecture of the neural specification. Performance is measured by Recall@K: the fraction of transitions for which the realized destination appears among the K highest-ranked employers. Appendix E.3 details how the two benchmark models are trained and evaluated.

Across the 4,932 validation transitions, the neural model places the realized destination among its ten highest-ranked employers in 23.0% of cases, compared with 18.3% for the type-specific multinomial logit and 17.9% for XGBoost. At fifty employers, these shares rise to 39.6%, 32.6%, and 32.1%, respectively. The neural model achieves higher recall throughout the range of rank cutoffs shown in Figure 3, indicating that its utility specification captures patterns in destination choices that improve the ranking of realized employers relative to both benchmarks.

Appendix Figure A3 reports predictive accuracy separately by worker type. Destination choices are easier to predict for more experienced workers with longer tenure (types 1 and 5), and harder for early-career blue-collar workers (type 8). Recall at ten employers (the share of observed transitions for which the realized destination is among the ten firms assigned the highest predicted probabilities) ranges from approximately 19% to 37% across worker types.

The inclusive value is a generated regressor, constructed from an estimated destination utility function. Following the cross-fitting approach of Chernozhukov et al. (2018), I divide workers into five groups, keeping all observations of a given worker together. Let $f(i) \in \{1, \ldots, 5\}$ denote worker i's group, and let $\widehat{\theta}^{(-f)}$ denote parameters estimated after excluding group f from the training sample. The utilities assigned to worker i are then $\widehat{v}_{i\ell jt} = v_{i\ell jt}(\widehat{\theta}^{(-f(i))})$. This separates the observations used to train the utility func-

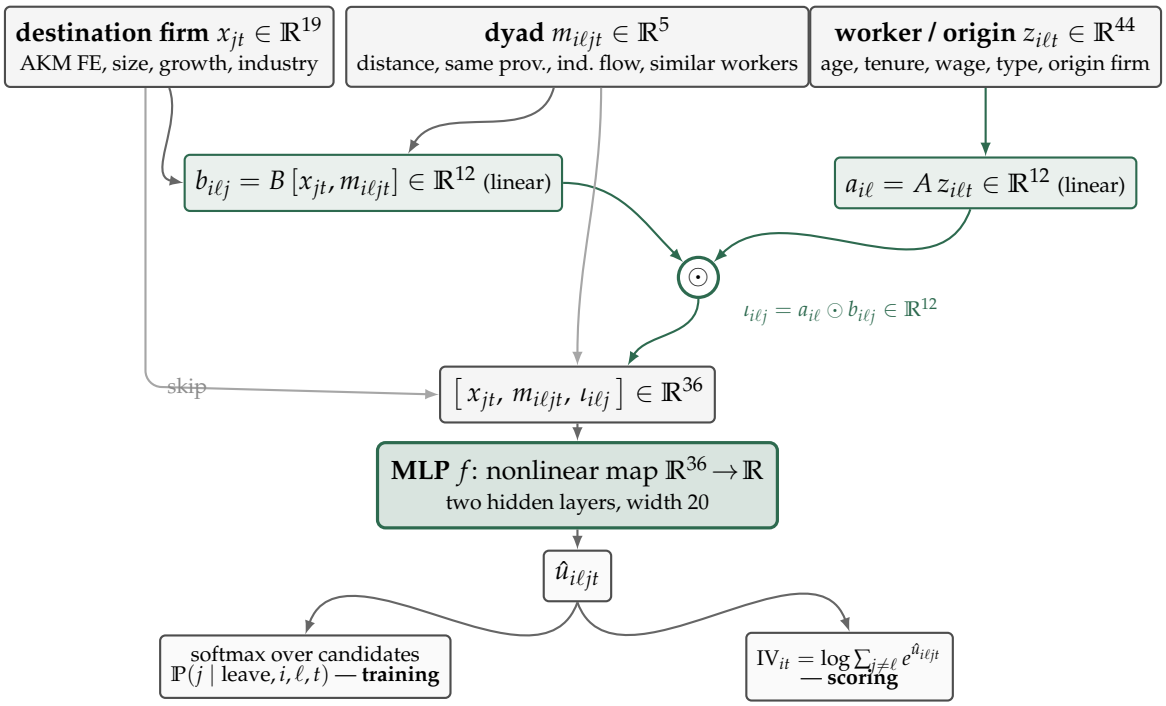
tion from those for which it generates predictions, preventing a worker's own transitions from entering the parameter updates used to construct their outside-option measure.

Because the five networks are estimated separately, their utility scores can have different numerical levels and spreads. Before combining their predictions, I shift and rescale each network's scores to match a common reference model estimated on the full training sample. I calculate these adjustments on a shared sample of worker-year observations, matching the average utility and its dispersion across destinations. This places the five sets of predictions on a comparable numerical scale before they are aggregated into inclusive values. Appendix E.4 details the calculation.

I construct the yearly, worker-specific inclusive value I_{it} for both workers who change employer and those who remain with their incumbent, writing I_{it} for the value associated with worker i and their current employer in year t. Applying a model estimated on transitions to stayers requires an additional assumption: conditional on observed worker, current-employer, and destination characteristics, the same systematic destination-utility function applies to both groups. I assume that the flexible neural specification approximates this common function sufficiently well to reconstruct the expected value of alternative employment for stayers using the preferences revealed by observationally equivalent leavers. For each worker-year, I then score the eligible destination employers, as defined in Appendix E.1, exclude the incumbent, and aggregate the adjusted utilities using the log-sum expression defined above.

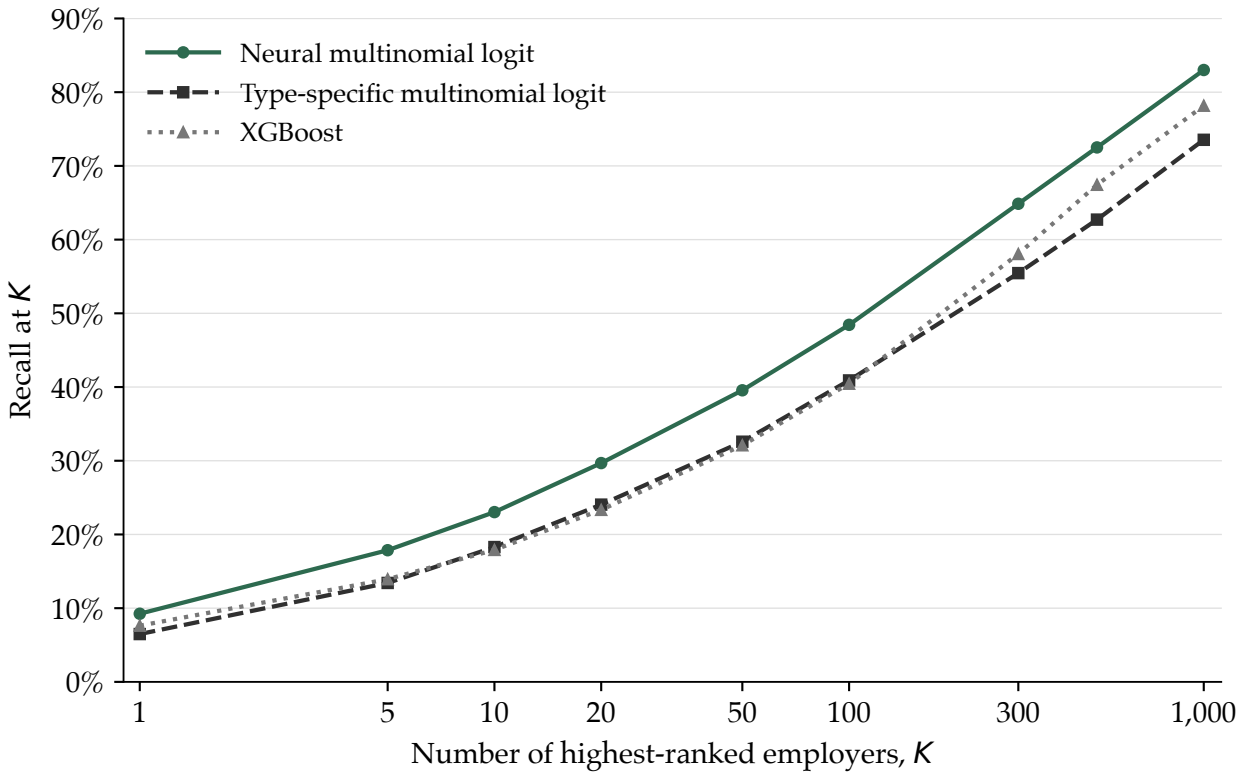
I attach the resulting inclusive value to each worker's annual primary employer. This produces an analysis panel containing 4,143,521 worker-year observations for 1,134,186 workers over 1995–1999, before outcome-specific restrictions. For the annual wage and mobility regressions, I standardize the index using its mean and standard deviation across this pooled panel, so coefficients are expressed per standard deviation of outside options.

Figure 2: How We Estimate: the DeepMNL Architecture



Notes: Worker/origin features enter only through interactions. Parameter counts: A: 540; B: 300; f: 1,181; total: 2,021.

Figure 3: Destination-Choice Validation



Notes: Recall at K is the share of observed transitions for which the realized destination is among the K highest-ranked employers. All three models are evaluated on the same 4,932 held-out transitions in 2000, ranking destinations among eligible employers and excluding the current employer. The horizontal axis uses a logarithmic scale. Appendix E.3 describes the benchmark models.

6 Inside the Outside Option

This section sheds light on what outside options, as measured by the inclusive value, look like in the data. To build intuition about the variation captured by this measure, consider a simpler multinomial logit, without any of the neural network architecture described in Section 5, with common coefficients and no interactions. Destination utility is then written as $v_{i\ell jt} = \alpha' x_{jt} + \beta' m_{i\ell jt}$, where x_{jt} is the vector of destination-employer characteristics and $m_{i\ell jt}$ is the vector of dyadic features capturing geographical proximity and compatibility between the worker, their current employer, and a potential destination. Worker and current-employer characteristics do not enter the model as separate covariates.

Define $Z_t = \sum_{k \in \mathcal{J}_t} \exp(\alpha' x_{kt})$ and $q_{jt} = \exp(\alpha' x_{jt}) / Z_t$. Factoring Z_t out of the sum over alternative employers gives

$$I_{i\ell t} = \log Z_t + \log \left[\sum_{\substack{j \in \mathcal{J}_t \\ j \neq \ell}} q_{jt} \exp(\beta' m_{i\ell jt}) \right]. \tag{16}$$

The first term is common to all workers in a given year. Variation across workers therefore stems from heterogeneous access and compatibility with employers, as measured by the dyadic features—geography, same-type employment shares, and historical industry flows—weighted by those employers' attractiveness, as well as the exclusion of the current employer. Having many attractive employers in the economy offers little benefit if they are distant or poorly suited to the worker.

Allowing valuations to vary with individual characteristics in a flexible way, through the neural network-augmented MNL, extends this intuition: access along each dimension matters more for workers who place greater weight on it, while the attractiveness of employer characteristics can also differ across workers.

In our data, these differences have a clear geographical dimension. Figure 4 shows the average standardized inclusive value among workers living in each municipality of Veneto. Outside options are generally weaker in the mountainous north and stronger across the central plain, where manufacturing activity is concentrated in industrial districts and around the region's main cities.

These municipality averages reflect differences in the composition of workers and their current employers, as well as the opportunities available nearby. Nevertheless, the map illustrates how unevenly economic opportunities are spread across the territory: mobility is highly local, with average distances between origin and destination employers of

approximately 12–17 km across worker types (Table 2), and access to attractive employers depends on where workers live.

Figure 5 describes which employers contribute to workers' outside options. For each worker-year and a group of alternative employers $G \subseteq \mathcal{J}_{it}$, I calculate the probability share $s_{it}(G) = \sum_{j \in G} \widehat{P}_{i\ell jt|\text{Leave}}$, using the destination probabilities defined in Section 5. The figure reports averages of these shares across sampled worker-years, grouping employers by distance, wage premium, province, and industry. Approximately two-thirds of this probability lies within 10 km of residence, and employers with higher AKM wage premiums generally account for larger shares. Opportunities are also concentrated in the current employer's province and industry, but these boundaries leave substantial alternatives outside: approximately 41% of predicted destination-choice probability is assigned to employers outside the current employer's two-digit industry.

These patterns suggest that coarse market definitions, such as province-by-industry cells, can exclude a substantial share of relevant employment alternatives. Concentration measures constructed within such boundaries may therefore misrepresent the employers competing for a worker's labor. My approach is closer to the notion of continuous, overlapping labor markets developed by Manning and Petrongolo (2017), Datta (2024), and Nimczik (2020): employers' relevance varies with their proximity and compatibility with each worker, rather than being determined by membership in a common geographical and industry category.

To describe how concentrated these opportunities are, I calculate the effective number of alternative employers, $N_{it}^{\text{eff}} = \left[\sum_{j \in \mathcal{J}_{it}} \widehat{P}_{i\ell jt|\text{Leave}}^2\right]^{-1}$. This equals N when destination-choice probability is evenly distributed across N employers. The effective number is the inverse of a worker-specific Herfindahl–Hirschman index (HHI), with employer shares given by predicted destination probabilities conditional on leaving.

This allows me to measure the concentration of alternative employment opportunities at the worker level, rather than assigning a common concentration measure to all workers within a province-by-industry market. Employers outside these boundaries can contribute to the measure, while workers within the same cell can face different levels of concentration because their proximity and compatibility with potential employers differ. Appendix Figure A4 shows that, in the diagnostic sample, the median effective number is approximately six, and 63% of worker-years have an effective number no greater than ten.

Figure 6 examines how destination utilities respond to changes in the model's inputs. Let $\widehat{v}_{i\ell jt}^{(k)}$ denote utility after increasing continuous input k by one standard deviation, or reversing it if binary, while holding the estimated network and all other inputs fixed. For

each input, I calculate $S_k = N^{-1} \sum_{it} \sum_{j \in \mathcal{J}_{it}} \widehat{P}_{i\ell jt|\text{Leave}} |\widehat{v}^{(k)}_{i\ell jt} - \widehat{v}_{i\ell jt}|$, averaging absolute utility changes across destinations using their original predicted choice probabilities, and then across the N sampled worker-years. These coefficients are weighted averages of absolute utility changes: they are strongly informative about the model's sensitivity to each input, but they do not indicate whether that input raises or lowers utility.

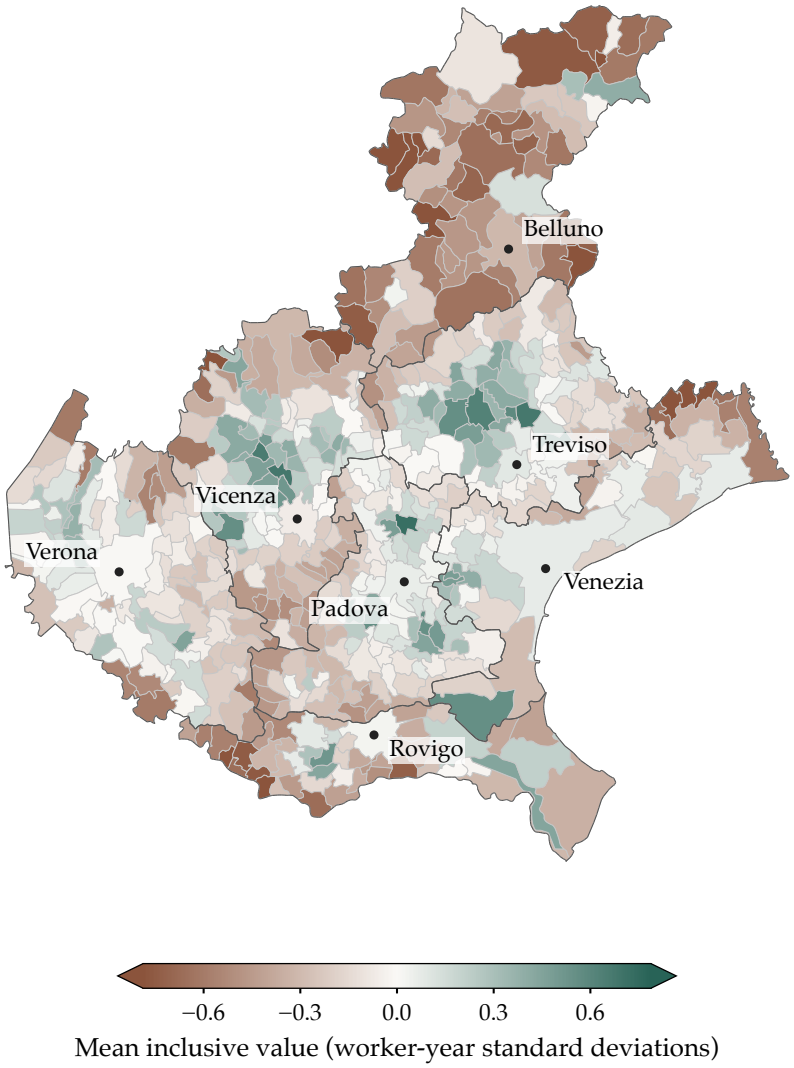
Geography and the number of same-type workers at the destination produce the largest responses. The importance of same-type employment supports the use of worker types as a proxy for compatibility and labor demand: whether a firm already employs workers with similar characteristics is informative about its relevance as a potential destination. This complements the geographical dimension of access. A nearby employer may offer few relevant opportunities if its workforce is poorly matched to the worker, while an employer with many similar workers may represent a more plausible alternative. The model picks up this dimension of compatibility and uses it to predict workers' destination choices.

Validating the intuition from the previous results, Appendix Figure A5 shows that outside options are higher where more employers are located within 10 or 25 km of residence. The relationship is stronger when counting employers that already employ workers of the same type, consistent with the importance of compatibility alongside proximity. Appendix Figure A6 complements this evidence by considering worker, current-employer, and local-market characteristics separately and then jointly. Older workers and workers with longer tenure also tend to have higher measured outside options. Strikingly, women have higher outside options as well. Although explaining this difference is beyond the scope of the analysis, one hypothesis is that geographical selection into employment plays a role: differences in female labor-force participation between urban and rural areas could affect which employment opportunities are available to the women observed working.

Appendix Figure A7 relates the inclusive value to realized outcomes among workers who change employers. When changing jobs, workers in the top decile by outside options move directly between employers in approximately 69% of cases, while workers in the lowest decile do so in approximately 55% of cases. Average nonemployment before the next job falls from about 2.5 months in the lowest decile to 1.5 months in the highest. This difference reflects both the greater frequency of direct transitions and shorter spells among workers who experience nonemployment: within the latter group, average duration falls from approximately 5.4 to 4.7 months. Workers with higher inclusive values also find their next employer closer to home. The median residence-to-destination distance declines from approximately 12.5 km in the lowest decile to 7.2 km in the highest, with similar patterns for direct transitions and moves involving nonemployment. These

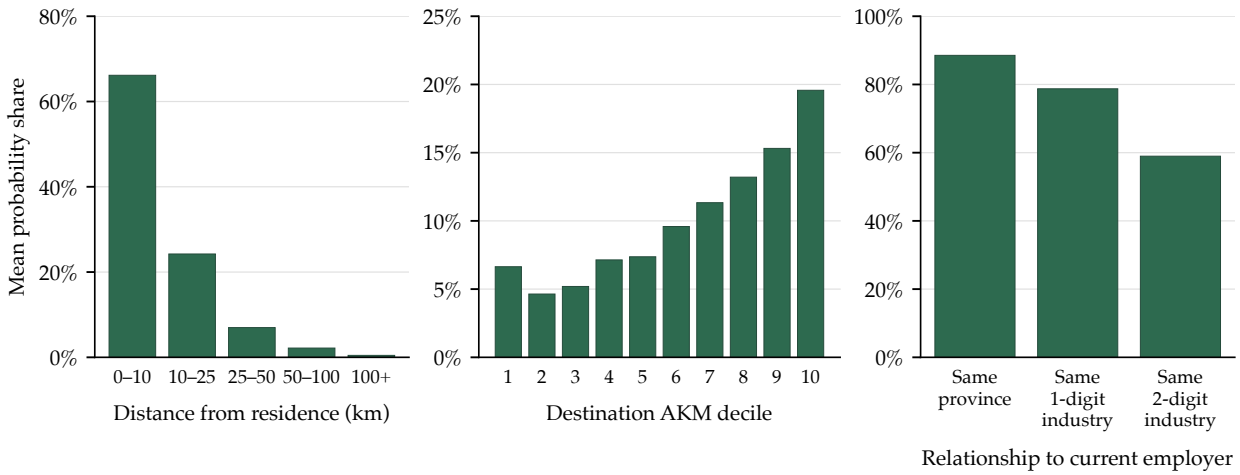
patterns are consistent with the measure capturing access to employment opportunities that workers can draw upon when changing jobs.

Figure 4: Outside Options Across Veneto



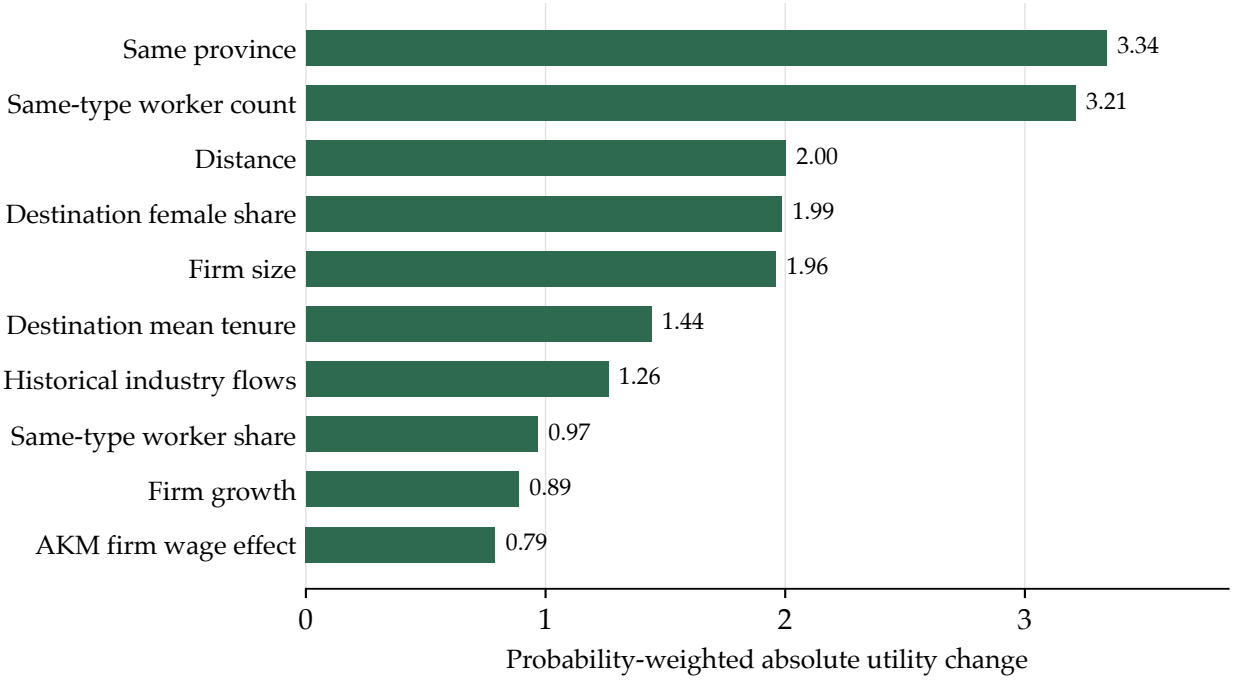
Notes: Municipality means of the cross-fitted inclusive value, standardized using the mean and standard deviation across mapped worker-year observations in 1995–1999. The map covers 563 municipalities and 4,625,612 worker-years. Historical residence codes are matched to current municipal boundaries; workplace location is used when residence coordinates are unavailable. Thin lines mark municipal boundaries, thicker lines mark province boundaries, and dots identify provincial capitals. The colour scale is centred at zero, with symmetric endpoints determined by the second and ninety-eighth percentiles of municipality means; values beyond these endpoints receive the endpoint colours.

Figure 5: What's Inside the Inclusive Value?



Notes: Each bar reports the average share of predicted destination-choice probability assigned to a group of employers, using the cross-fitted model and a sample of 1,000 worker-year observations. For each observation, the share is the sum of exponentiated destination utilities within the group divided by the corresponding sum over all eligible alternative employers, excluding the current employer. The left panel groups employers by distance from the worker's residence; the middle panel groups them by AKM firm wage-effect decile, from lowest to highest; the right panel reports shares in the current employer's province and one- and two-digit industries. The three groups in the right panel overlap. Vertical scales differ across panels.

Figure 6: Local Sensitivity of Destination Utilities



Notes: Bars show the average absolute change in predicted destination utility following a change in one input, weighted across alternative employers by the worker's original predicted destination-choice probabilities. Results use the full-training-sample network and 800 worker-year observations. Continuous inputs are increased by one standard deviation of their values in the evaluated destination arrays; binary inputs are flipped between zero and one. The network parameters, worker characteristics, and other inputs remain fixed. Utilities are recalculated through the full network, including its matching interactions. The current employer is excluded.

7 Validation Using Mass Layoffs

Next, I turn to a test of whether the inclusive value does indeed capture the quality of workers' outside options. To validate my measure, I look at displaced workers who lose their job in a mass layoff, and track whether their recovery from job loss varies by the level of their pre-displacement outside option. The reasoning is that, having been forced to leave their employer, these workers must draw on the alternative employment opportunities available to them, making their subsequent labor market trajectories informative about the quality of those opportunities. A good measure of outside options should therefore predict better recovery from displacement for workers with higher measured outside options.

Appendix Figure A8 describes the timing of re-employment following displacement.

Appendix Figure A9 provides a first descriptive comparison using Kaplan–Meier survival curves for the broader displacement sample. Workers with higher pre-displacement inclusive values return to employment more quickly: after twelve months, approximately 36% remain without observed re-employment, compared with 40% among workers below the median. This pattern provides initial support for the measure. To further explore whether workers with better outside options recover relatively better from job loss, I turn to the matched event-study design.

I investigate this question using event studies, following a long tradition of research on the consequences of job displacement (Jacobson, LaLonde, and Sullivan, 1993; Lachowska, Mas, and Woodbury, 2020; Brandily, Hémet, and Malgouyres, 2022). I identify mass layoffs as employment contractions of at least 50% at Veneto employers with at least ten employees, excluding events in which workers move together to a common employer. I match each displaced worker to a worker who remains with their employer, matching exactly on calendar year, worker type, gender and qualification, and selecting the nearest neighbour in pre-displacement age, tenure, wages and the employer's AKM wage effect. The resulting sample contains 11,595 matched pairs. I follow their outcomes from three years before to three years after the event, comparing recovery separately for displaced workers with above- and below-median pre-displacement inclusive values. Appendix F.1 provides the full sample-construction and matching procedure.

I then estimate, within the matched sample, the following regression separately for workers with above- and below-median pre-displacement inclusive values:

$$Y_{ipk} = \alpha_p + \delta_k + \sum_{\substack{h=-3 \\ h\neq -1}}^{3} \beta_h D_i \, \mathbf{1}\{k = h\} + \varepsilon_{ipk}, \tag{17}$$

where Y_{ipk} is the outcome of worker i, belonging to matched pair p, at event time k, and D_i indicates the displaced member of the pair. Matched-pair fixed effects α_p absorb differences in outcome levels across pairs, while event-time fixed effects δ_k capture the common evolution of outcomes around the event. The year preceding the event, $k = -1$, is the omitted reference period.

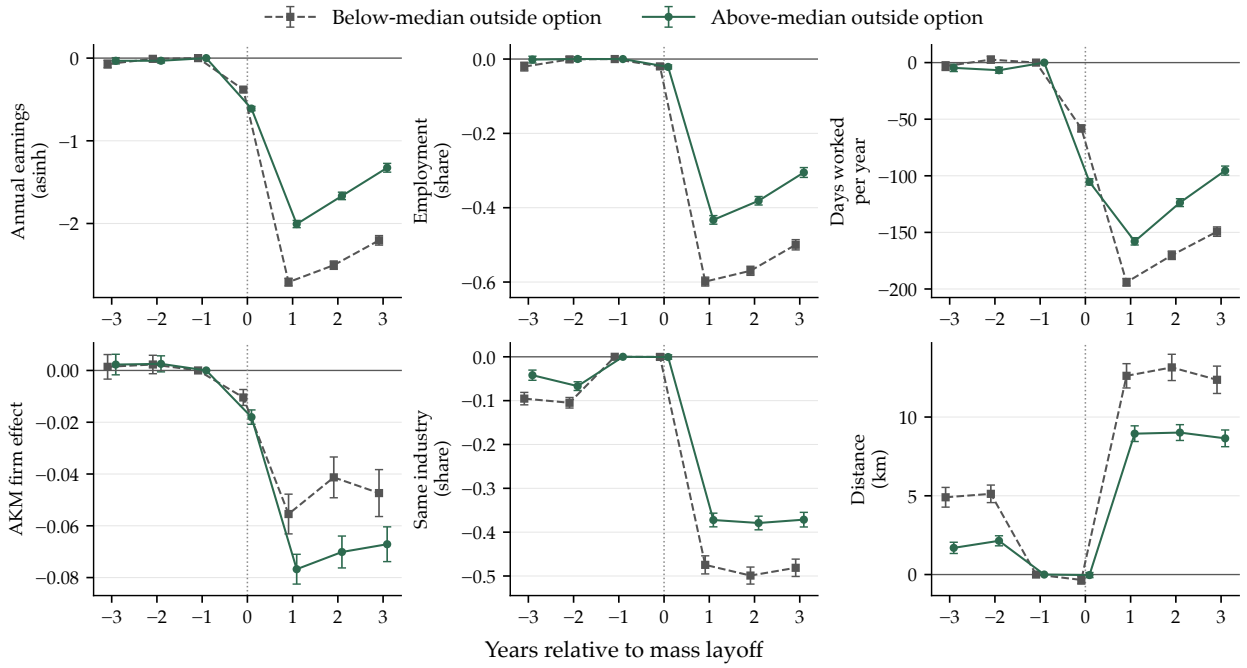
Figure 7 presents the event-study estimates, with 95% confidence intervals based on heteroskedasticity-robust standard errors. Following displacement, workers in both groups experience substantial earnings losses relative to their matched controls. However, workers with higher pre-displacement outside options recover more strongly, and their earnings shortfall is smaller throughout the subsequent three years. Employment follows a similar pattern: in the first year after the event, the employment shortfall is approximately 43 percentage points for workers above the median inclusive value, compared with 60 percentage points for those below it. By the third year, these gaps narrow to approximately 31 and 50 percentage points, respectively. Workers with higher outside options also experience smaller reductions in days worked: the third-year shortfall is approximately 95 days, compared with 149 days for workers below the median.

The characteristics of subsequent employment help interpret these differences. Among those observed in employment, workers with higher pre-displacement inclusive values work closer to their original employer and are more likely to remain in its two-digit industry. These patterns are consistent with interpreting the measure as capturing access to compatible employers: workers with better outside options appear able to find alternative employment without moving as far geographically or changing industry as frequently.

Interestingly, this stronger recovery does not appear to reflect access to employers with higher wage policies, as measured by their AKM fixed effects.

Taken together, these results support interpreting the inclusive value as a measure of access to alternative employment: workers with higher pre-displacement outside options recover more strongly from job loss, finding employment closer to their previous employer and more often within the same industry.

Figure 7: Labor-Market Outcomes Around Mass Layoffs



Notes: Event-study coefficients for displaced workers relative to matched controls, estimated separately for workers with above- and below-median pre-layoff inclusive values. Specifications include matched-pair and event-time fixed effects; the year before displacement ($t = -1$) is the reference period. Bars show 95% confidence intervals based on heteroskedasticity-robust standard errors. The vertical dotted line marks the layoff year. Annual earnings are transformed as $\text{asinh}(\text{earnings}/1{,}000)$, including zeros. The remaining panels report employment, annual days worked, the employer's AKM wage effect, employment in the pre-layoff two-digit industry, and distance to the employer. Vertical scales differ across outcomes.

8 Outside Options, Mobility, and Wages

The preceding sections have examined what the inclusive value captures about workers' alternative employment opportunities. I have described how these opportunities vary across workers and employers, and shown that workers with higher pre-displacement inclusive values recover more strongly from mass layoffs. This evidence supports interpreting the measure as a proxy for workers' outside options. I now use it to investigate how outside options relate to mobility and wages during the course of workers' employment relationships. In particular, do workers with better outside options leave their employer more frequently, and do they receive higher wages while remaining in the same job?

The conceptual framework in Section 4 guides this analysis. Its prediction for mobility is unambiguous: holding the value of the current employment relationship fixed, better outside options make leaving more attractive and increase the probability of changing employers. The wage response, however, depends on the wage-setting protocol. Under wage posting, improvements in an individual worker's outside options need not change the wage offered by their current employer. Under bargaining, those improvements strengthen the worker's position in the negotiation and can translate into higher pay within the existing relationship. I first examine the relationship between outside options and employer mobility, before turning to wages and the evidence they provide on wage-setting conduct.

I first estimate a logit for changing employers in the following year, conditional on employment in both years. Individual fixed effects mean that I exploit changes in outside options over time for the same worker, while employer fixed effects account for persistent firm-specific factors, such as remoteness. Year fixed effects absorb common aggregate changes. Thus, the remaining variation exploits changes in workers' outside options over time, including differential changes among workers employed at the same firm, as the surrounding labor market evolves in ways that appeal differently to different workers. Finally, controlling for current wages allows me to examine the relationship between outside options and mobility holding pay fixed, since outside options may themselves influence the wage paid by the incumbent.

Column (1) of Table 3 reports a positive coefficient of 0.1409 on the standardized inclusive value, corresponding to an average marginal effect of approximately 1.78 percentage points for a one standard deviation increase in outside options. Relative to the employer-change rate of 14.2% in the eligible annual sample, this amounts to an increase of approximately 12.5%. To compare the estimated coefficient with the nested-logit restriction

discussed in Section 4, it must be expressed in the units of the unstandardized index. Writing $I_{it}^{\text{std}} = (I_{it} - \bar{I})/s_I$, the corresponding coefficient is $\widehat{\beta}/s_I$. Since $s_I = 2.809$, the estimate implies an unstandardized coefficient of approximately $0.1409/2.809 \simeq 0.050$. With the upper-level error scale normalized to one, this value lies within the interval $0 < \sigma \leq 1$, satisfying the necessary sign and scale restriction of the nested-logit representation. The coefficient remains positive and statistically significant when leaving also includes workers no longer observed in employment in the following year, at 0.0993 in column (2).

Appendix Table A4 reports linear probability models with the same controls and fixed effects. A one standard deviation increase in outside options is associated with a 0.98 percentage-point increase in employer changes and a 0.82 percentage-point increase in broad leaving. Both estimates are positive and statistically significant.

I now examine whether better outside options are associated with higher wages. The distinction matters for the wage-setting mechanisms discussed in Section 4: outside opportunities may influence workers' decisions to leave without affecting their incumbent wage, whereas bargaining allows them to influence pay within an existing employment relationship. Table 4 examines this relationship using progressively richer controls for worker, employer, and match heterogeneity. The worker and employer fixed-effects specification takes the form

$$\log w_{it} = \beta I_{it}^{\text{std}} + X_{it}'\delta + \alpha_i + \psi_\ell + \lambda_t + \varepsilon_{it}, \tag{18}$$

where w_{it} denotes real weekly wages, I_{it}^{std} is the standardized inclusive value, and X_{it} contains the observed worker covariates. The fixed effects capture persistent worker characteristics, persistent employer differences, and common year effects, respectively. Throughout the table, the inclusive value uses the same standardization over the annual panel. Standard errors are clustered by employer.

Column (1) includes employer and year fixed effects, alongside observed worker characteristics. Employer fixed effects account for persistent differences in firms' wage policies and other characteristics that might influence both pay and workers' outside opportunities. Year effects absorb aggregate wage movements. The coefficient implies that a one standard deviation increase in outside options is associated with approximately 2.5% higher wages. Adding worker-type fixed effects in column (2), which account for average wage differences across the groups constructed in Section 3, yields a slightly larger estimate of approximately 2.8%.

Columns (3) and (4) introduce individual fixed effects. These absorb persistent differences between workers, including unobserved characteristics associated with both their wages and their access to alternative employers. Once I include individual fixed effects, the variation I exploit comes from changes in outside options over time for the same worker. Appendix Figure A10 sheds light on how much of this variation is left. Most variation in outside options is, of course, across workers, but a substantial share—around 17% in the balanced panel—comes from changes within workers over time. The correlation between inclusive values four years apart is 0.729: this is high, but still leaves a fair degree of variation. This reflects, in part, how the outside labor market evolves in ways that appeal differently to different workers.

The coefficient is 0.0175 without additional worker covariates and 0.0133 when these are included. The positive relationship thus persists after accounting for individual heterogeneity, with the latter estimate corresponding to approximately 1.3% higher wages per standard deviation of outside options.

Columns (3) and (4) correspond to augmented AKM specifications. Appendix Figure A11 provides a graphical illustration of column (3): after residualizing both log weekly wages and the inclusive value on worker, employer and year fixed effects, the remaining variation in outside options is positively associated with the remaining variation in wages. Outside options therefore account for wage variation beyond the additive worker and employer components of the AKM model.

Because the inclusive value depends jointly on worker characteristics and the current employer, and need not decompose into additive worker and employer components, its positive coefficient can be interpreted as evidence of match effects in wage determination. Their modest economic magnitude nevertheless suggests that log-additive separability may remain a useful first approximation to wages.

Column (5) introduces worker–employer fixed effects, giving each pair its own intercept. These absorb persistent match characteristics and let me ask whether wages change as the same worker's outside options change while employed at the same firm, in a specific match. The coefficient remains positive and statistically significant at 0.0103, corresponding to approximately 1.0% higher wages per standard deviation of outside options.

Column (6) adds employer-year fixed effects, a very demanding fixed-effects specification: the only variation left is the change in outside options within a match, netting out the average changes for workers at that particular firm. Employer-year effects absorb, for example, changes in the employer's wage policy, or local shocks that affect all workers at a firm equally. The result survives this specification, with a positive and statistically

significant coefficient of 0.0139, corresponding to approximately 1.4% higher wages per standard deviation of outside options.

Taken together, these results provide evidence of wage bargaining within ongoing employment relationships: improvements in a worker's outside options go along with wage growth. My preferred specification, column (6), shows that this relationship survives after accounting for changes common to workers at the same firm: workers whose labor market prospects improve differentially more also see larger wage gains, within the same match, consistent with wage bargaining.

To shed light on where bargaining plays a greater role, I next turn to heterogeneity. Table 5 summarizes the wage and mobility responses across worker characteristics, collective agreements and employer growth, with detailed results in the appendix. The wage regressions follow the specification in column (4) of Table 4, including worker, employer and year fixed effects alongside worker covariates.

First, men appear to respond more strongly to better outside opportunities, both through higher wages and through a greater probability of leaving their current employer. The gender difference is particularly pronounced for wages, suggesting that men are better able to translate outside opportunities into higher pay. This is consistent with Caldwell and Harmon (2019), who also find smaller earnings responses to improvements in outside opportunities among women.

For age and tenure, a different pattern emerges across the two margins. Younger workers and workers with shorter tenure show larger wage responses, whereas older workers and those with longer tenure show larger mobility responses. One possible explanation is that workers earlier in their careers or employment relationships remain further from the wage they could obtain within their current match. Better outside opportunities may therefore give them leverage to negotiate higher pay without changing employers. Conversely, workers closer to their earning ceiling within a match may need to leave to obtain further gains. This interpretation echoes the mechanism in Postel-Vinay and Robin (2002): outside offers can induce wage increases when the incumbent can match them, while offers exceeding what the incumbent can offer lead to mobility, though a formal test of this mechanism is beyond the scope of this analysis.

A similar pattern appears across employment grades. Among apprentices, blue-collar workers and white-collar workers, the wage response is strongest for apprentices, followed by blue-collar and then white-collar workers. The ranking of mobility responses is reversed. These results are consistent with apprentices having greater scope to use outside opportunities to raise their pay within their current employment relationship, while white-collar workers more often respond by changing employers.

Finally, employer growth reveals an interesting difference between wage and mobility responses. At firms classified as growing on the basis of headcount changes, the wage response to outside options is roughly twice that observed at shrinking firms, while the mobility response is small and statistically insignificant. At shrinking firms, by contrast, better outside opportunities are associated with a larger increase in departures. This pattern is consistent with bargaining over rents. Growing firms may have more rents to share and a stronger incentive to retain workers, allowing employees to use better outside opportunities to renegotiate their pay and take home a larger share of those rents. At shrinking firms, where there may be less room to raise wages, workers instead make greater use of the mobility margin. The distinction suggests that the way outside options translate into workers' outcomes depends on the incumbent employer's rents, and its ability and willingness to respond. A more formal test of this hypothesis will appear in Section 10.

Table 3: Outside Options and Leaving the Current Employer

	(1) Employer change	(2) Broad leaving
Inclusive value (std.)	0.1409***	0.0993***
	(0.0096)	(0.0083)
Average marginal effect	0.0178	0.0123
Odds ratio	1.151	1.104
Current wage and worker covariates	Yes	Yes
Worker FE	Yes	Yes
Employer FE	Yes	Yes
Year FE	Yes	Yes
Eligible observations	3,807,791	4,143,521
Logit sample	1,356,661	1,954,829
Numerical estimation sample	1,308,686	1,911,727
Outcome mean (eligible sample)	0.142	0.212
Outcome mean (logit sample)	0.321	0.337

Notes: Logit estimates. An employer change equals one when annual primary employers differ between t and $t+1$, conditional on observed employment in both years. Broad leaving also includes no recorded employment in $t+1$. Years are 1995–1999. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Both columns control for current log weekly wages, age squared, log tenure and log prior employment experience, with worker, employer and year fixed effects. Linear age is omitted with worker and year effects; worker-invariant characteristics are absorbed by worker effects. Standard errors in parentheses are heteroskedasticity-robust. Fixed-effect groups that perfectly predict the outcome are removed recursively. Extended maximum likelihood additionally treats 47,975 and 43,102 separated observations in columns (1) and (2) as having limiting predicted probabilities of zero or one. These observations remain in the logit sample but do not enter numerical estimation. Average marginal effects are $\widehat{\beta}\,\overline{\widehat{p}(1-\widehat{p})}$, averaged over the logit sample with zero contributions from separated observations, and expressed as probability changes. Odds ratios equal the exponential of the logit coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Outside Options and Current Wages

	(1) Employer FE	(2) + worker type FE	(3) Worker and employer FE	(4) + worker covariates	(5) Match FE	(6) Match and employer-year FE
	Dependent variable: Log weekly wages in t					
Inclusive value (std.)	0.0249***	0.0279***	0.0175***	0.0133***	0.0103***	0.0139***
	(0.0007)	(0.0007)	(0.0009)	(0.0010)	(0.0010)	(0.0006)
Worker covariates	Yes	Yes	No	Yes	Yes	Yes
Worker type FE	No	Yes	No	No	No	No
Employer FE	Yes	Yes	Yes	Yes	No	No
Worker FE	No	No	Yes	Yes	No	No
Worker-employer FE	No	No	No	No	Yes	Yes
Employer-year FE	No	No	No	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	No
Outcome mean	6.500	6.500	6.500	6.500	6.500	6.500
Observations	4,143,153	4,143,153	3,968,053	3,968,053	3,588,982	3,584,307

Notes: OLS estimates; the dependent variable is log real weekly wages in the current primary job. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. Worker-employer fixed effects identify matches; employer-year fixed effects absorb common wage changes within each employer and year. Singleton fixed-effect groups are removed recursively. Standard errors in parentheses are clustered by employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: Heterogeneity in Wage and Mobility Responses to Outside Options

Dimension	Group	Wage response ($100\widehat{\beta}$)	Mobility response (percentage points)	Appendix
Gender	Men	1.92***	1.53***	Table A5
	Women	0.55***	0.94***	
Age	Younger	1.80***	0.95***	Table A6
	Middle	0.95***	1.55***	
	Older	0.96***	1.57***	
Tenure	Less than 2 years	2.45***	0.58	Table A7
	2–4 years	0.94***	1.38***	
	5 or more years	0.75***	1.55***	
Employment grade	Blue-collar	0.89***	1.15***	Table A8
	White-collar	0.67***	1.79***	
	Apprentice	1.06***	1.07*	
	Other	0.57**	2.71***	
Collective agreement	Metalworking	1.01***	0.25	Table A9
	Commerce/Services	1.10***	2.04*	
	Textile/Clothing	0.20	1.14***	
	Wood/Furniture	1.18***	1.34***	
	Construction	2.63***	2.40***	
	Footwear	0.46	2.21	
	Other	1.25***	1.78***	
Employer growth	Shrinking	1.25***	1.58**	Table A10
	Stable	1.37***	1.07*	
	Growing	2.51***	0.16	

Notes: Entries report 100 times the coefficient on the inclusive value from separate subgroup regressions. The index uses the same pooled standardization across groups, so responses correspond to a one standard deviation increase in outside options. The wage outcome is log real weekly wages; entries in the wage column are approximately percentage wage differences. Mobility responses are percentage-point changes in the probability of changing annual primary employer, conditional on employment in both years. Wage specifications follow column (4) of Table 4. All regressions include worker, employer and year fixed effects and worker covariates where not absorbed. Mobility is estimated using linear probability models without current wages as a control. Standard errors are clustered by employer. The linked appendix tables report full coefficients, standard errors, observation counts and group definitions. Cross-group comparisons describe point estimates; no formal tests of coefficient equality are reported. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

9 Identification Using Large-Firm Openings

The demanding fixed-effects specifications presented above provide clear evidence of wage bargaining. I now turn to an instrumental-variable approach to validate the causal interpretation of these results. The strategy exploits large-firm openings, which expand employment opportunities differently across workers depending on their access to the opening employers. I use historical mobility patterns to measure this differential exposure and instrument changes in workers' outside options, allowing me to estimate their causal effect on wages and mobility.

The opening of a large employer close to a worker expands the set of alternative employers they can consider and should therefore improve their outside options. Nevertheless, the opening employer need not be equally relevant to all nearby workers. How valuable this new opportunity is depends on the compatibility between workers' characteristics and the employer's labor demand.

I capture this differential exposure using a shift-share design (Bartik, 1991), in which the shifts come from employment at large opening firms and the shares reflect historical mobility patterns. The strategy is similar in spirit to Schubert, Stansbury, Taska, and Dey (2026).

Let k denote worker type, as defined in Section 3, s the current employer's two-digit industry, and p its province. Let c index destination cells, also defined by two-digit industry and province. Denoting by $N^0_{ksp\to c}$ the number of direct employer-to-employer transitions from origin group (k, s, p) to destination cell c in 1990–1994, I define the historical mobility share as

$$\omega^0_{kspc} = \frac{N^0_{ksp\to c}}{\sum_{c'} N^0_{ksp\to c'}}. \tag{19}$$

These shares measure which destinations have historically attracted workers of a given type from the same origin industry and province, providing weights for the relevance of subsequent openings. In practice, a worker is more exposed to an opening when workers of their type have historically moved more often from their current employer's industry and province to the opening employer's industry and province. Using pre-period flows keeps these shares fixed during the analysis period, so that changes in exposure within an origin group are driven by openings rather than workers' contemporaneous mobility responses.

I classify a firm as an opening in year t if it is active in that year but was inactive in the preceding year. For each threshold $q \in \{50, 100, 200\}$, let O^q_{ct} denote total opening-year employment at firms in destination cell c with at least q employees. Employment is mea-

sured by the number of distinct workers observed at each firm during the year. Appendix Table A11 reports the annual number of openings and total employment at these firms for each size threshold. The annual exposure of worker i to large-firm openings is then

$$Z_{it}^{q} = \sum_{c} \omega_{k_i, s_{it}, p_{it}, c}^{0} O_{ct}^{q}, \tag{20}$$

where k_i is the worker's type and s_{it} and p_{it} identify their current employer's industry and province. This exposure increases when large employers open in destinations historically connected to the worker's origin group. I standardize each exposure measure over the worker-year panel and use it as an excluded instrument for the standardized inclusive value, estimating the model separately for each threshold.

I estimate the effects of outside options on wages and mobility by two-stage least squares. For each opening-size threshold q, the first stage takes the form

$$I_{it}^{\text{std}} = \pi_q \widetilde{Z}_{it}^{q} + X_{it}'\rho + \mathcal{F}_{it}^{I} + \nu_{it}, \tag{21}$$

where $\widetilde{Z}_{it}^{q}$ is standardized opening exposure. The outcome equation is

$$Y_{it} = \beta_q I_{it}^{\text{std}} + X_{it}'\delta + \mathcal{F}_{it}^{Y} + \varepsilon_{it}, \tag{22}$$

with the inclusive value instrumented by $\widetilde{Z}_{it}^{q}$. The outcome Y_{it} is either current log real weekly wages or an indicator for changing annual primary employer between t and $t+1$. The mobility equation is a linear probability model estimated among workers observed in employment in both years.

Both stages use the same sample, controls and fixed-effects structure for each outcome. The controls include age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience, where not absorbed by the fixed effects. For mobility, both stages additionally control for current log weekly wages. I report two specifications: one with employer, worker-type and year fixed effects, and another with worker, employer and year fixed effects. The terms $\mathcal{F}_{it}^{I}$ and $\mathcal{F}_{it}^{Y}$ collect the corresponding fixed effects in each equation. Standard errors are clustered by current employer.

The exclusion restriction is that, conditional on the controls and fixed effects, large-firm openings affect wages and mobility only through the value of outside options. Identification relies on differences in workers' exposure to the same openings, measured using historical mobility patterns. These differences must be unrelated to other changes that independently affect wages or mobility. Using historical flows prevents contemporaneous mobility responses from determining the exposure weights. In the specification

with worker, firm and year fixed effects, persistent differences across workers and firms are absorbed, while year fixed effects account for common macroeconomic shocks. This alleviates concerns that these sources of heterogeneity drive the results: the remaining variation comes from changes in exposure over time, net of persistent worker and firm differences and common yearly shocks.

Table 6 reports positive and significant first stages in all cases, with F-statistics between 11.7 and 25.7, above the conventional benchmark of 10. The implied 2SLS wage effects range from approximately 4–5% per standard deviation of outside options with employer, worker-type and year fixed effects. Including worker fixed effects brings the coefficients to 0.0218–0.0231, corresponding to approximately 2.2–2.3% higher wages and broadly consistent with the OLS coefficient of 0.0139 in column (6) of Table 4. For leaving, the 2SLS coefficients are also positive, although larger and less precisely estimated than the OLS estimate, which nevertheless lies within their 95% confidence intervals. Taken together, these results support the view that better outside options, instrumented by new employment opportunities weighted by their accessibility to each worker, raise wages and encourage mobility, consistent with the bargaining mechanism and the OLS evidence.

Table 6: Instrumental-Variable Estimates Using Large-Firm Openings

	Log weekly wages		Employer change	
	(1)	(2)	(3)	(4)
Panel A: Openings with at least 50 employees				
Inclusive value (2SLS)	0.0409***	0.0218*	0.0165	0.0536
	(0.0144)	(0.0120)	(0.0299)	(0.0379)
Opening exposure (first stage)	0.0506***	0.0518***	0.0498***	0.0511***
	(0.0100)	(0.0113)	(0.0102)	(0.0116)
First-stage F	25.7	20.8	23.8	19.4
Panel B: Openings with at least 100 employees				
Inclusive value (2SLS)	0.0469***	0.0219*	0.0184	0.0572
	(0.0168)	(0.0131)	(0.0324)	(0.0411)
Opening exposure (first stage)	0.0464***	0.0479***	0.0458***	0.0474***
	(0.0102)	(0.0115)	(0.0104)	(0.0118)
First-stage F	20.8	17.3	19.3	16.1
Panel C: Openings with at least 200 employees				
Inclusive value (2SLS)	0.0494**	0.0231	0.0241	0.0685
	(0.0195)	(0.0151)	(0.0369)	(0.0489)
Opening exposure (first stage)	0.0401***	0.0412***	0.0396***	0.0409***
	(0.0103)	(0.0117)	(0.0106)	(0.0119)
First-stage F	15.0	12.5	14.0	11.7
Worker covariates	Yes	Yes	Yes	Yes
Current log weekly wage	No	No	Yes	Yes
Employer FE	Yes	Yes	Yes	Yes
Worker type FE	Yes	No	Yes	No
Worker FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Eligible worker-years	4,112,858	4,112,858	3,780,325	3,780,325
Outcome mean (eligible sample)	6.503	6.503	0.142	0.142

Notes: Each panel reports separate 2SLS regressions. The excluded instrument is standardized annual opening-employment exposure, defined in equation (20). Thresholds count distinct workers at each firm during its opening year. Columns (1)–(2) use log real weekly wages in the current primary job. Columns (3)–(4) use a linear probability model for an annual primary-employer change between t and $t+1$, conditional on employment in both years. Years are 1995–1999. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. Mobility regressions additionally control for current log weekly wages. Controls and fixed effects enter both stages. Standard errors in parentheses are clustered by current employer. First-stage coefficients are on standardized opening exposure, with the standardized inclusive value as the dependent variable. The first-stage F is the squared employer-clustered t-statistic on opening exposure. Eligible counts and outcome means refer to complete-case samples before recursive removal of singleton fixed-effect groups. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

10 Portugal: Replication and Extensions

I now turn to replicating my main results in a different country and time frame: Portugal in 2013–2019. Alongside the linked employer–employee records of Quadros de Pessoal (QP), I have access to firm balance-sheet data from the Sistema de Contas Integradas das Empresas (SCIE). This allows me to examine whether the wage results extend beyond Veneto and, further, to test whether workers with better outside options obtain larger wage gains when their employer's value added increases. In a bargaining framework, the value of alternative employment serves as the worker's threat point. Better outside options may therefore allow workers to bring home a larger share of the firm's rents and benefit more from improvements in their employer's performance.

I construct outside options using the same destination-choice framework, adapting the inputs to the Portuguese data. Precise details on the implementation can be found in Appendix I. Occupational and educational information allows me to measure compatibility through the composition of employment at potential destinations. Geography is considerably coarser, however, with locations identified only at the NUTS II regional level. I therefore supplement regional indicators and historical flows with a latent distance estimated from firm-to-firm transitions, capturing proximity in observed mobility.

The intuition behind this application is straightforward. Latent-space models represent connections in a network through distances in an unobserved space (Hoff, Raftery, and Handcock, 2002). In my application, firms exchanging more workers, conditional on their overall propensity to send and receive workers and their observed characteristics, tend to be placed closer together. The model thus finds a system of coordinates that rationalizes the observed pattern of flows, assigning a position to each firm. These positions form a map of mobility connections: for example, two firms that draw workers from, and send workers to, a similar set of employers may be placed close together even if few workers move between them directly. I then measure the distance between the position of a worker's current employer and that of each potential destination, and include this distance among the dyadic features in the destination-choice model. Appendix I.3 details the latent-space specification and the construction of these distances.

Predictive performance is close to that of the Veneto comparison specification, as reported in Appendix Table A12. This suggests that the richer Portuguese covariates—including occupation, education, contractual status, and workforce composition—together with information from the mobility network can compensate for the lack of detailed geography, despite the central role of geographical proximity in predicting mobility in Veneto.

Table 7 shows that the wage results extend to Portugal. The coefficient on outside options is positive and statistically significant across all four specifications. In the most demanding specification, column (4), worker–employer match effects account for permanent differences within employment relationships, while firm-by-year effects absorb changes common to workers at the same employer, including changes in its wage policy, mimicking column (6) of Table 4. The remaining variation therefore compares workers whose outside options evolve differently within the same firm. A one standard deviation improvement in outside options is associated with approximately 1.4% higher wages, almost identical to the corresponding estimate of 1.4% for Veneto.

I next turn to how outside options affect workers' participation in their employer's rents. The heterogeneity results in Section 8 suggest that workers at growing firms respond to better outside options more through wage gains, whereas those at shrinking firms respond more through mobility. One interpretation is that workers use their outside options to negotiate higher wages when their employer has rents to share. The Portuguese firm accounts allow me to investigate this mechanism more directly: do workers with better outside options obtain larger wage gains when their employer's value added increases? Such a finding would be consistent with outside options strengthening workers' bargaining position and allowing them to bring home a larger share of improvements in their employer's performance.

To investigate this question, I estimate a wage equation allowing the response to employer value added to vary with workers' outside options:

$$\begin{aligned} \log w_{i\ell t} = {} & \beta_1(V_{\ell t} - \bar{V}) + \beta_2 I_{it} + \beta_3(V_{\ell t} - \bar{V})I_{it} \\ & + \kappa \log K_{\ell t} + X'_{i\ell t}\pi + \mathcal{F}_{i\ell t} + e_{i\ell t}, \end{aligned} \tag{23}$$

where $V_{\ell t}$ denotes log value added per worker, $\bar{V}$ is its common-sample mean, and I_{it} is the standardized inclusive value. Capital per worker, $K_{\ell t}$, and the worker and employer characteristics collected in $X_{i\ell t}$ account for other determinants of wages. The fixed effects $\mathcal{F}_{i\ell t}$ vary across the specifications reported in Table 8, with the most demanding specification including worker–employer match and year effects.

The coefficient of interest is β_3. The wage response to log value added per worker is $\beta_1 + \beta_3 I_{it}$: a positive interaction therefore means that workers with better outside options receive larger wage gains when their employer's performance improves. Match effects allow me to examine this relationship within ongoing employment relationships, accounting for permanent differences across matches that could otherwise influence both wages and access to alternative employment.

Following Card, Devicienti, and Maida (2014), I instrument log value added per worker with the log of average revenue per worker at firms in the same four-digit industry but outside the incumbent employer's NUTS II region. I instrument its interaction with outside options using the interaction between this revenue measure and I_{it}. The identifying assumption is that, conditional on the controls and fixed effects, revenue changes at firms in the same industry outside the incumbent's region affect workers' wages through their employer's value added. Appendix I.5 provides further details on the sample, instrument construction, and estimation; Appendix Table A13 reports the first-stage estimates.

Table 8 shows that workers with better outside options obtain larger wage gains when their employer's value added increases. In Panel A, the interaction is positive and statistically significant across all four OLS specifications. This relationship survives worker–employer match effects, with a coefficient of 0.0044 in column (4). The 2SLS estimates in Panel B point in the same direction and, once worker effects are included, imply larger responses. In the match-effects specification, the interaction coefficient is 0.0163. For a 10% increase in value added per worker, the point estimates imply a wage response of approximately 0.15% at $I_{it} = 0$ and 0.31% at $I_{it} = 1$—a response approximately twice as large, or an additional 0.16 percentage points of wage growth.

These findings strengthen the bargaining interpretation of the wage results. Better outside options are associated with higher wages within ongoing employment relationships and with a larger wage response to improvements in employer performance. The Portuguese evidence thus suggests that the value of alternative employment matters for how firms share rents with their workers.

Table 7: Portugal: the Bargaining Gradient Replicates

	(1)	(2)	(3)	(4)
Veneto (log weekly wage)	0.0133***	0.0103***	0.0193***	0.0139***
	(0.0010)	(0.0010)	(0.0006)	(0.0006)
Portugal (log monthly wage)	0.0083***	0.0064***	0.0098***	0.0139***
	(0.0008)	(0.0016)	(0.0006)	(0.0011)
Worker controls	Yes	Yes	Yes	Yes
Worker FE	Yes	No	Yes	No
Worker×Firm (match) FE	No	Yes	No	Yes
Firm FE	Yes	No	No	No
Firm×Year FE	No	No	Yes	Yes
Year FE	Yes	Yes	No	No
Obs. (Veneto)	3.97M	3.59M	3.97M	3.58M
Obs. (Portugal)	17.6M	15.8M	17.0M	15.3M

Notes: Entries report the coefficient on the standardized outside-option index; standard errors are in parentheses. For Portugal, the outcome is log total monthly earnings, winsorized at the pooled 0.5th and 99.5th percentiles. The cross-fitted index is standardized over the pooled annual panel before selecting each regression sample. Controls are age squared, log one plus tenure, fixed-term status, and part-time status, where not absorbed. Match effects are worker-by-employer effects; firm-by-year effects absorb separate firm and year effects. Portuguese standard errors are clustered by the current employer.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 8: Rent Sharing: VA Pass-Through Rises with Outside Options

	(1)	(2)	(3)	(4)
Panel A: OLS				
Log VA/worker (β_1)	0.2845***	0.0272***	0.0187***	0.0183***
	(0.0132)	(0.0029)	(0.0027)	(0.0027)
Outside options (β_2)	0.0465***	0.0101***	0.0090***	0.0089***
	(0.0064)	(0.0018)	(0.0023)	(0.0028)
Log VA/worker $\times I_{it}$ (β_3)	0.0377***	0.0037**	0.0043***	0.0044**
	(0.0101)	(0.0016)	(0.0016)	(0.0019)
Panel B: 2SLS				
Log VA/worker (β_1)	0.4109***	0.1186***	0.0112	0.0151
	(0.0441)	(0.0314)	(0.0868)	(0.0879)
Outside options (β_2)	0.0385***	0.0115***	0.0086***	0.0074**
	(0.0064)	(0.0021)	(0.0026)	(0.0031)
Log VA/worker $\times I_{it}$ (β_3)	0.0352	0.0111*	0.0134*	0.0163*
	(0.0226)	(0.0065)	(0.0069)	(0.0091)
Worker and employer covariates	Yes	Yes	Yes	Yes
Worker FE	No	Yes	Yes	No
Employer FE	No	No	Yes	No
Worker–employer match FE	No	No	No	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	6,007,659	5,643,155	5,641,338	5,464,130

Notes: Panel A reports ordinary least-squares estimates and Panel B two-stage least-squares estimates of equation (23). The dependent variable is log total monthly earnings. Both panels use the same observations in each column. Standard errors, clustered by four-digit industry, are in parentheses. VA/worker is value added at factor cost per employee; its logarithm is centered at the common-sample mean. The outside-option index I_{it} is standardized within year before selecting the rent-sharing sample. In Panel B, log VA/worker and its interaction with outside options are instrumented by the log mean revenue per employee of firms in the same four-digit industry outside the incumbent's NUTS II region, and its interaction with I_{it}. All columns control for log tangible capital per employee; age, tenure, firm age and log employment and their squares; and a broad occupation indicator, where not absorbed. Match effects are worker-by-employer effects. Appendix I.5 gives the sample restrictions, instrument construction, and specifications; Appendix Table A13 reports the first stages. Stars use two-sided cluster-t p-values: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

11 Conclusion

This paper develops a worker-specific measure of outside options and uses it to study how alternative employment opportunities affect wages and mobility. I model employment choice using a nested-logit structure: at the upper level, workers decide whether to stay with their current employer or leave; at the lower level, they choose a destination employer conditional on leaving. I estimate the destination-choice model using a neural network, allowing tastes for destination characteristics to vary flexibly across workers while retaining the structural advantages of discrete choice models. I rely on a clustering approach to reconstruct latent worker types and measure compatibility with each potential employer using the number and share of same-type workers in its workforce. These measures proxy employers' demand for workers with similar skill profiles, helping distinguish relevant employment opportunities from firms that are nearby but poorly suited to the worker. Together with the flexible neural specification, they improve predictive accuracy relative to a parametric logit. The discrete choice structure also provides a closed-form measure of the expected maximum utility available from alternative employment: the inclusive value of the set of employers other than the incumbent. This worker-specific measure reflects which employers are within reach, how compatible they are with the worker, and how the worker values their characteristics.

I estimate the model using linked employer–employee data from the Veneto region of northeastern Italy over 1995–2000. The resulting inclusive value of each worker's choice set, which I use as a proxy for outside options, predicts recovery from job loss. Following mass layoffs, workers with better pre-displacement outside options return to employment more quickly and experience smaller losses in earnings and employment. Their subsequent employers are also closer to their original workplace and more often in the same industry. These findings support interpreting the inclusive value as a measure of access to relevant employment opportunities.

Better outside options also encourage mobility and raise wages within ongoing employment relationships. The wage response survives demanding specifications with worker–employer match and employer-year fixed effects, which absorb permanent match characteristics and changes common to workers at the same firm. The estimates therefore exploit changes in the expected utility of alternative employment during an ongoing match, net of firm-wide and aggregate changes. In my preferred specification, an improvement of one standard deviation in outside options corresponds to approximately 1.4% higher wages. I then develop an instrumental-variable approach using a shift-share instrument that combines large-firm openings with historical transition probabilities into

the industry–province cells where these openings occur. The instrumental-variable estimates provide further support for a causal wage response. These results provide evidence that workers use alternative employment opportunities to bargain over pay, consistent with models of wage bargaining and rejecting the simple wage-posting benchmark.

Workers' responses to changes in the quality of their outside options vary with their own characteristics and those of their incumbent employer. Younger workers and workers with shorter tenure show stronger wage responses, while older workers and those with longer tenure respond more through mobility. At growing firms, improvements in outside options translate more strongly into wage gains; at shrinking firms, departures are the more important margin.

I then replicate the analysis by estimating the destination-choice model using Portuguese linked employer–employee data for 2013–2019. In the most demanding specification, an improvement of one standard deviation in outside options corresponds to approximately 1.4% higher wages. Firm accounts also allow a more direct investigation of rent sharing. Workers with better outside options obtain larger wage gains when their employer's value added increases, consistent with alternative employment strengthening their bargaining position and allowing them to benefit more from improvements in employer performance.

In this paper, I argue that a detailed characterisation of a worker's outside options requires accounting for how their individual characteristics and current employer shape access to alternative employment, compatibility with potential destinations, and the value they attach to those opportunities. Taken together, my results show that improvements in these outside options affect both mobility and wage setting: workers respond by moving between employers and by using better alternative opportunities to negotiate higher pay, even within ongoing employment spells.

References

Abowd, John M, Francis Kramarz, and David N Margolis (1999). "High wage workers and high wage firms". In: *Econometrica* 67.2, pp. 251–333.

Abowd, John M, Kevin L McKinney, and Ian M Schmutte (2019). "Modeling endogenous mobility in earnings determination". In: *Journal of Business & Economic Statistics* 37.3, pp. 405–418.

Aouad, Ali and Antoine Désir (2025). "Representing random utility choice models with neural networks". In: *Management Science* 72.8, pp. 6686–6701. DOI: 10.1287/mnsc.2023.02189.

Azar, José, Steven T. Berry, and Ioana Marinescu (2022). *Estimating Labor Market Power*. Working Paper 30365. National Bureau of Economic Research. DOI: 10.3386/w30365.

Azar, José, Ioana Marinescu, and Marshall Steinbaum (2022). "Labor market concentration". In: *Journal of Human Resources* 57.S, S167–S199.

Barrela, Rodrigo (Oct. 2026). "The Value of Outside Options: Measurement, Impact and Inequality". Job Market Paper, Nova School of Business and Economics. URL: https://rodrigobarrela.github.io/assets/pdf/Barrela_JMP.pdf.

Bartik, Timothy J. (1991). *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. DOI: 10.17848/9780585223940. URL: https://research.upjohn.org/up_press/77/.

Bartolucci, Cristian, Francesco Devicienti, and Ignacio Monzón (2018). "Identifying Sorting in Practice". In: *American Economic Journal: Applied Economics* 10.4, pp. 408–438. DOI: 10.1257/app.20170068.

Beaudry, Paul, David A. Green, and Benjamin Sand (2012). "Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory". In: *Econometrica* 80.3, pp. 1063–1104. DOI: 10.3982/ECTA8659.

Bengio, Yoshua and Jean-Sébastien Sénécal (2003). "Quick training of probabilistic neural nets by importance sampling". In: *Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics (AISTATS)*, pp. 17–24. URL: https://proceedings.mlr.press/r4/bengio03a.html.

Berger, David, Kyle Herkenhoff, Jaehun Jeong, et al. (2026). *Labor Market Power with Worker and Firm Heterogeneity*. Working Paper 35640. National Bureau of Economic Research. DOI: 10.3386/w35640.

Berger, David, Kyle Herkenhoff, and Simon Mongey (2022). "Labor Market Power". In: *American Economic Review* 112.4, pp. 1147–1193. DOI: 10.1257/aer.20191521.

Blanc, Guy and Steffen Rendle (2018). "Adaptive Sampled Softmax with Kernel Based Sampling". In: *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. Proceedings of Machine Learning Research. PMLR, pp. 590–599. URL: https://proceedings.mlr.press/v80/blanc18a.html.

Bocquet, Leonard (2024). "The network origin of slow labor reallocation". In: *Working Paper*.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019). "A distributional framework for matched employer employee data". In: *Econometrica* 87.3, pp. 699–739.

— (2022). "Discretizing Unobserved Heterogeneity". In: *Econometrica* 90.2, pp. 625–643. DOI: 10.3982/ECTA15238.

Brandily, Paul, Camille Hémet, and Clément Malgouyres (2022). "Understanding the reallocation of displaced workers to firms". In: *IZA Discussion Paper n. 15105*.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin (2006). "Wage bargaining with on-the-job search: Theory and evidence". In: *Econometrica* 74.2, pp. 323–364.

Caldwell, Sydnee and Oren Danieli (2024). "Outside Options in the Labour Market". In: *Review of Economic Studies* 91.6, pp. 3286–3315.

Caldwell, Sydnee, Ingrid Haegele, and Jörg Heining (2025). *Bargaining and Inequality in the Labor Market*. Working Paper 25-413. W.E. Upjohn Institute for Employment Research. DOI: 10.17848/wp25-413.

Caldwell, Sydnee and Nikolaj Harmon (2019). "Outside options, bargaining, and wages: Evidence from coworker networks". In: *Unpublished manuscript, Univ. Copenhagen*, pp. 203–207.

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

Card, David, Francesco Devicienti, and Agata Maida (2014). "Rent-sharing, holdup, and wages: Evidence from matched panel data". In: *Review of Economic Studies* 81.1, pp. 84–111.

Chernozhukov, Victor et al. (2018). "Double/debiased machine learning for treatment and structural parameters". In: *The Econometrics Journal* 21.1, pp. C1–C68. DOI: 10.1111/ectj.12097.

Datta, Nikhil (2024). *Local Monopsony Power*. Discussion Paper 2012. Centre for Economic Performance, London School of Economics. URL: https://cep.lse.ac.uk/pubs/download/dp2012.pdf.

Del Prato, Francesco (2024). ""The Importance of Working for Earnest": Climbing the Job Ladder through Firms' Connectivity". In: *working paper*.

Di Addario, Sabrina et al. (2023). "It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages". In: *Journal of Econometrics* 233.2, pp. 340–374.

Dodini, Samuel et al. (2024). "Monopsony, Job Tasks and Labour Market Concentration". In: *The Economic Journal* 134.661, pp. 1914–1949.

Ek, Simon et al. (2025). *Outside Options and the Sharing of Match-Specific Rents*. Working Paper 2025:17. Institute for Evaluation of Labour Market and Education Policy. URL: https://www.ifau.se/globalassets/pdf/se/2025/wp-2025-17-outside-options-and-the-sharing-of-match-specific-rents.pdf.

Friederiszick, Theodor, Axel Malmcrona, and Nelson Sunuwar (2026a). "Competition in the Labor Market: Evidence from Swedish Pharmacies". Unpublished manuscript.

— (2026b). "The Job Diversion Matrix: How Workers' Alternatives Determine the Harm from Layoffs and Mergers". Unpublished manuscript.

Hirsch, Boris et al. (2022). *The Wage Elasticity of Recruitment*. Discussion Paper 1883. Centre for Economic Performance, London School of Economics. URL: https://cep.lse.ac.uk/pubs/download/dp1883.pdf.

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock (2002). "Latent Space Approaches to Social Network Analysis". In: *Journal of the American Statistical Association* 97.460, pp. 1090–1098. DOI: 10.1198/016214502388618906.

Hruschka, Harald, Werner Fettes, and Markus Probst (2004). "An empirical comparison of the validity of a neural net based multinomial logit choice model to alternative model specifications". In: *European Journal of Operational Research* 159.1, pp. 166–180. DOI: 10.1016/S0377-2217(03)00410-7.

Huang, Zhexue (1998). "Extensions to the k-means algorithm for clustering large data sets with categorical values". In: *Data mining and knowledge discovery* 2.3, pp. 283–304.

Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan (1993). "Earnings losses of displaced workers". In: *The American economic review*, pp. 685–709.

Jäger, Simon, Christopher Roth, et al. (2024). "Worker beliefs about outside options". In: *The Quarterly Journal of Economics* 139.3, pp. 1505–1556.

Jäger, Simon, Benjamin Schoefer, et al. (2020). "Wages and the Value of Nonemployment". In: *The Quarterly Journal of Economics* 135.4, pp. 1905–1963.

Jarosch, Gregor, Jan Sebastian Nimczik, and Isaac Sorkin (2024). "Granular search, market structure, and wages". In: *Review of Economic Studies* 91.6, pp. 3569–3607.

Kim, Jin-Hwa et al. (2017). "Hadamard Product for Low-rank Bilinear Pooling". In: *International Conference on Learning Representations*. URL: https://arxiv.org/abs/1610.04325.

Kline, Patrick (2024). "Firm Wage Effects". In: *Handbook of Labor Economics*. Vol. 5. Elsevier, pp. 115–181. DOI: 10.1016/bs.heslab.2024.11.005.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2020). "Leave-out estimation of variance components". In: *Econometrica* 88.5, pp. 1859–1898.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, et al. (2022). "Wage posting or wage bargaining? a test using dual jobholders". In: *Journal of Labor Economics* 40.S1, S469–S493.

Lachowska, Marta, Alexandre Mas, and Stephen A Woodbury (2020). "Sources of displaced workers' long-term earnings losses". In: *American Economic Review* 110.10, pp. 3231–3266.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (2022). "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market". In: *American Economic Review* 112.1, pp. 169–212. DOI: 10.1257/aer.20190790.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet (2021). "Gender differences in job search: Trading off commute against wage". In: *The Quarterly Journal of Economics* 136.1, pp. 381–426. DOI: 10.1093/qje/qjaa033.

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin (2023). "The anatomy of sorting—evidence from danish data". In: *Econometrica* 91.6, pp. 2409–2455.

Manning, Alan (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

— (2011). "Imperfect Competition in the Labor Market". In: *Handbook of Labor Economics*. Vol. 4. Elsevier.

Manning, Alan and Barbara Petrongolo (2017). "How local are labor markets? Evidence from a spatial job search model". In: *American Economic Review* 107.10, pp. 2877–2907.

Marinescu, Ioana, Ivan Ouss, and Louis-Daniel Pape (2021). "Wages, Hires, and Labor Market Concentration". In: *Journal of Economic Behavior & Organization* 184, pp. 506–605. DOI: 10.1016/j.jebo.2021.01.033.

McFadden, Daniel (1978). "Modeling the Choice of Residential Location". In: *Transportation Research Record* 673, pp. 72–77. URL: https://onlinepubs.trb.org/Onlinepubs/trr/1978/673/673-012.pdf.

— (1981). "Econometric Models of Probabilistic Choice". In: *Structural Analysis of Discrete Data with Econometric Applications*. Ed. by Charles F. Manski and Daniel McFadden. MIT Press, pp. 198–272.

Nimczik, Jan Sebastian (2020). "Job mobility networks and data-driven labor markets". In: *Unpublished Working Paper.[648]*.

Postel-Vinay, Fabien and Jean-Marc Robin (2002). "Equilibrium wage dispersion with worker and employer heterogeneity". In: *Econometrica* 70.6, pp. 2295–2350.

Roussille, Nina and Benjamin Scuderi (2025). *Bidding for Talent: A Test of Conduct in a High-Wage Labor Market*. Working Paper 33848. National Bureau of Economic Research. DOI: 10.3386/w33848.

Schmutte, Ian M. (2014). "Free to Move? A Network Analytic Approach for Learning the Limits to Job Mobility". In: *Labour Economics* 29, pp. 49–61.

Schubert, Gregor, Anna Stansbury, and Bledi Taska (2024). "Employer concentration and outside options". In: *Available at SSRN 3599454*.

Schubert, Gregor, Anna Stansbury, Bledi Taska, and Matthew Dey (May 2026). *Employer Concentration and Wages: Evidence from Large Firms' Hiring Shocks*. Research Paper 7334-25. MIT Sloan School of Management. DOI: 10.2139/ssrn.5374463. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5374463.

Serafinelli, Michel (2019). ""Good" Firms, Worker Flows, and Local Productivity". In: *Journal of Labor Economics* 37.3, pp. 747–792. DOI: 10.1086/702628.

Sorkin, Isaac (2018). "Ranking firms using revealed preference". In: *The quarterly journal of economics* 133.3, pp. 1331–1393.

Sorkin, Isaac and Lucas Warwar (Oct. 2026). "Paired Movers". Working paper, Stanford University. URL: https://lucaswarwar.github.io/sw_logadditivity.pdf.

Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge University Press. URL: https://eml.berkeley.edu/books/choice2.html.

Wang, Shenhao, Qingyi Wang, and Jinhua Zhao (2020). "Deep neural networks for choice analysis: Extracting complete economic information for interpretation". In: *Transportation Research Part C: Emerging Technologies* 118, p. 102701. DOI: 10.1016/j.trc.2020.102701.

Appendix

A Table of contents

1 Introduction **1**

2 Related Literature **4**
2.1 Measuring Outside Options 4
2.2 Pass-Through of Outside Options to Wages 5
2.3 Bargaining, Wage Posting, and the Role of Firms 7
2.4 Discrete Choice Models of Labor Markets 8
2.5 Mobility Networks and Labor Market Boundaries 10

3 Data and Sample **10**

4 Conceptual Framework **17**

5 Estimating Outside Options **20**

6 Inside the Outside Option **28**

7 Validation Using Mass Layoffs **34**

8 Outside Options, Mobility, and Wages **37**

9 Identification Using Large-Firm Openings **45**

10 Portugal: Replication and Extensions **49**

11 Conclusion **54**

Appendices **61**

A Table of contents **61**

B List of figures **62**

C List of tables **63**

D Data, Worker Types, and Geography **64**

E Measurement and Model Diagnostics **68**
E.1 Estimation of the Neural Destination-Choice Model 68
E.2 Utility Approximation and the Inclusive Value 71
E.3 Estimation of the Benchmark Destination-Choice Models 72
E.4 Alignment of Utility Scores Across Worker Folds 74

F Displacement Validation **84**
F.1 Construction of the Matched Displacement Sample 84

G Mobility, Wages, and Heterogeneity **87**

H Opening-Based Instruments **96**

I Portugal: Data and Empirical Specifications **97**
I.1 Worker Panel and Estimation Sample 97
I.2 Inputs and Neural Utility Specification 98
I.3 Latent Mobility Distance and Destination Pools 100
I.4 Training, Cross-Fitting, and Predictive Evaluation 101
I.5 Firm Accounts and Rent-Sharing Specifications 103

B List of figures

Figure 1 Nested Structure of Employer Choice 17
Figure 2 How We Estimate: the DeepMNL Architecture 26
Figure 3 Destination-Choice Validation 27
Figure 4 Outside Options Across Veneto 31
Figure 5 What's Inside the Inclusive Value? 32
Figure 6 Local Sensitivity of Destination Utilities 33
Figure 7 Labor-Market Outcomes Around Mass Layoffs 36
Figure A1 Employer Size and Employment Concentration 65
Figure A2 Residence and Workplace Geography 66
Figure A3 Destination-Choice Prediction by Worker Type 79
Figure A4 Effective Number of Alternative Employers 80
Figure A5 Outside Options and Nearby Employers 81
Figure A6 Who Has Good Outside Options? 82
Figure A7 Outside Options and Employer Transitions 83
Figure A8 Nonemployment and Re-employment After Displacement 85
Figure A9 Re-employment After Mass Layoffs by Pre-Displacement Outside Options . . 86
Figure A10 Variation in Outside Options Within and Between Workers 88
Figure A11 Outside Options and Wages Net of AKM Fixed Effects 89

C List of tables

Table 1	Summary Statistics by Worker Type	15
Table 2	Mobility by Worker Type: Destination-Model Training Sample	16
Table 3	Outside Options and Leaving the Current Employer	42
Table 4	Outside Options and Current Wages	43
Table 5	Heterogeneity in Wage and Mobility Responses to Outside Options	44
Table 6	Instrumental-Variable Estimates Using Large-Firm Openings	48
Table 7	Portugal: the Bargaining Gradient Replicates	52
Table 8	Rent Sharing: VA Pass-Through Rises with Outside Options	53
Table A1	Employment by Industry: Ten Largest Industries, 1995–1999	64
Table A2	Sample Construction	67
Table A3	Inputs to the Neural Destination-Choice Model	76
Table A4	Outside Options and Mobility: Linear Probability Models	87
Table A5	Wages and Mobility by Gender	90
Table A6	Wages and Mobility by Age	91
Table A7	Wages and Mobility by Employer Tenure	92
Table A8	Wages and Mobility by Employment Grade	93
Table A9	Wages and Mobility by Collective Agreement	94
Table A10	Wages and Mobility by Employer Growth	95
Table A11	Firm Openings and Employment, 1995–1999	96
Table A12	Destination-Choice Prediction in Veneto and Portugal	103
Table A13	Portugal: First Stages for Rent Sharing	105

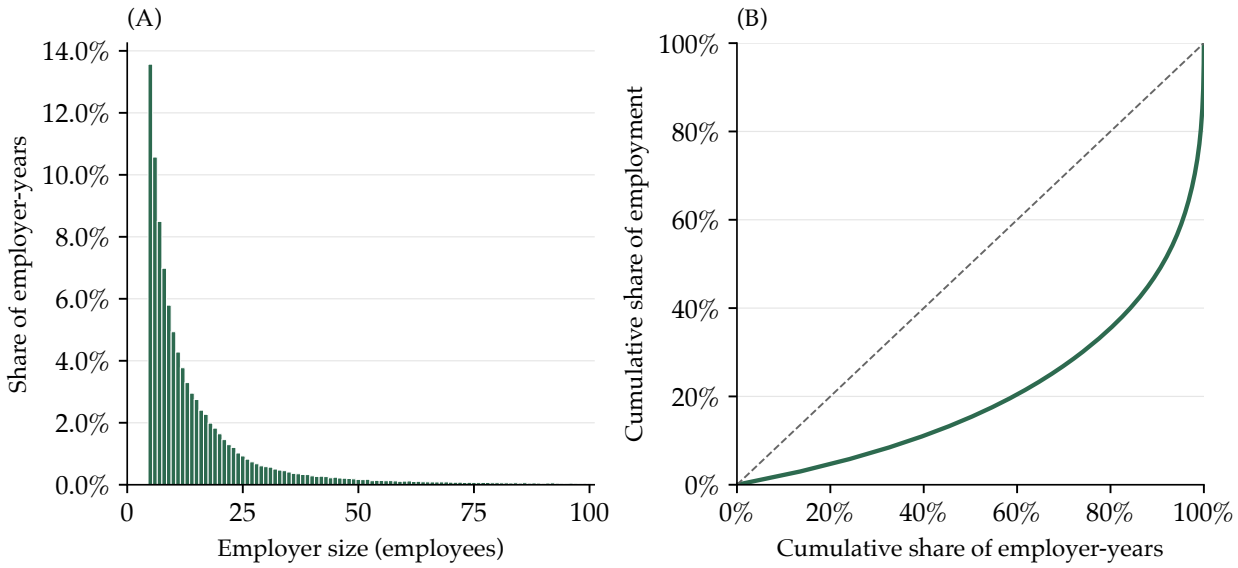
D Data, Worker Types, and Geography

Table A1: Employment by Industry: Ten Largest Industries, 1995–1999

ATECO 81	Industry	Average annual employment	Employment share (%)	Employer-years	Mean employer size
31	Metal products (excluding machinery)	132,602	13.2	27,151	24.4
45	Footwear and clothing	88,223	8.8	20,553	21.5
50	Construction	59,460	5.9	23,492	12.7
66	Hotels and catering	51,719	5.1	15,999	16.2
46	Wood products and wooden furniture	50,709	5.0	13,441	18.9
43	Textiles	47,646	4.7	7,114	33.5
32	Mechanical machinery and equipment	47,211	4.7	7,166	32.9
83	Business services	41,717	4.2	9,608	21.7
34	Electrical machinery and equipment	39,010	3.9	6,383	30.6
61	Wholesale trade	37,835	3.8	11,704	16.2
–	Other industries	408,884	40.7	80,356	25.4
–	All industries	1,005,016	100.0	222,967	22.5

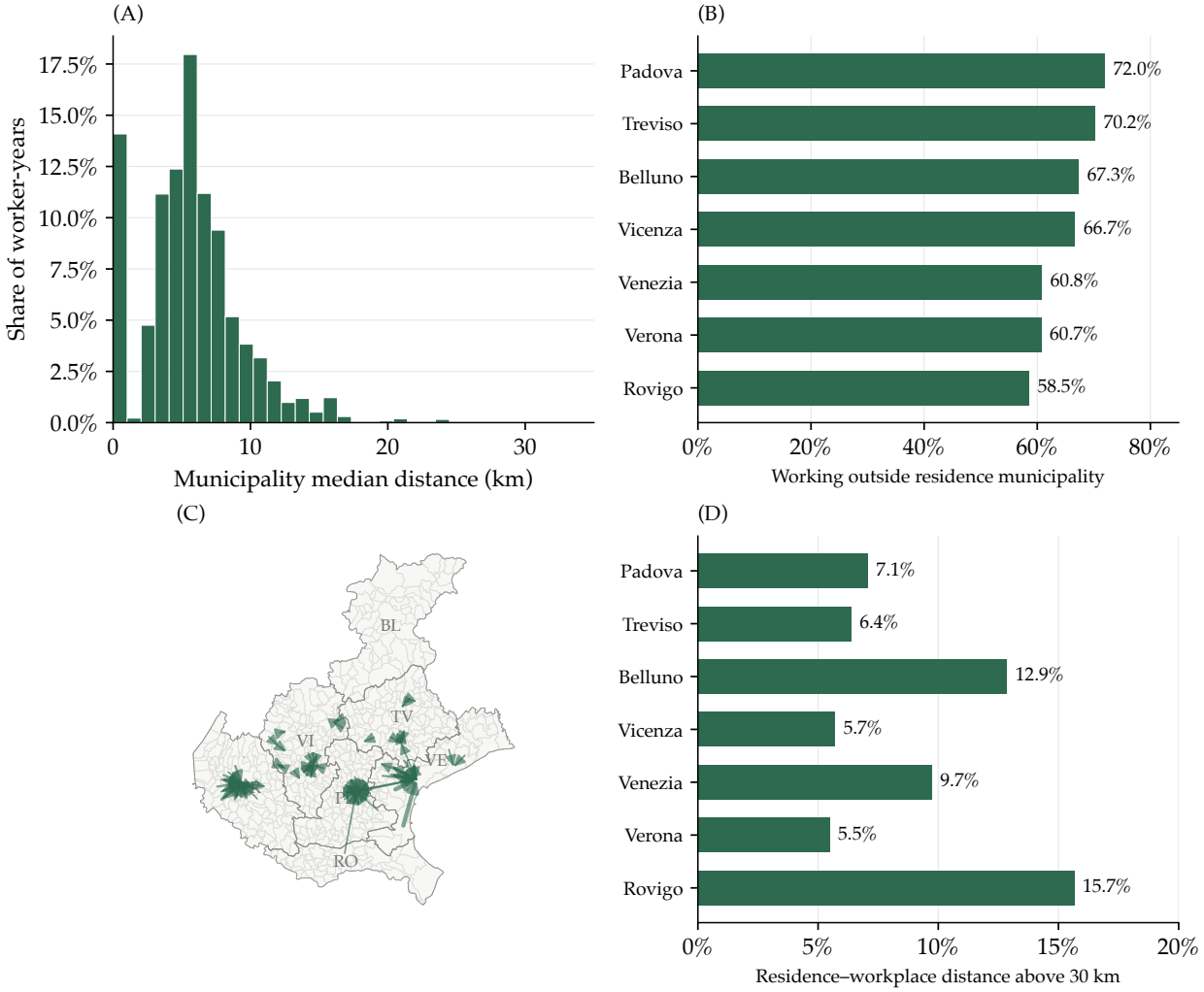
Notes: Industries are ranked by employment using the two-digit ATECO 1981 classification; industry descriptions are abbreviated. The sample comprises eligible employer-year observations in the seven Veneto provinces during 1995–1999, with at least five employees and a valid industry code, matching the sample used in Appendix Figure A1. Employer size counts distinct workers recorded at an employer during the year. Average annual employment is the sum of these annual headcounts divided by five; employment shares use the corresponding sum across all industries as the denominator. Workers employed at multiple firms within a year can enter multiple employer headcounts. Employer-years count annual employer observations; mean employer size is total headcount divided by employer-years. Figures are not full-time equivalents or employment on a particular reference date. Other industries groups all industries outside the top ten. Average annual employment is rounded to the nearest worker.

Figure A1: Employer Size and Employment Concentration



Notes: Panel A shows the distribution of annual employer headcounts, expressed as shares of all 222,967 employer-year observations. The sample includes active employers in the seven Veneto provinces in 1995–1999. Employer size counts distinct workers recorded during the year. Panel A displays sizes up to 100; the 2.7% of employer-years above 100 account for 33.1% of employment. Panel B includes all employers and plots the cumulative employment share against the cumulative employer-year share, ordered by size. The dashed diagonal represents equal employer sizes.

Figure A2: Residence and Workplace Geography



Notes: Veneto-resident worker-year observations in 1995–1999. Panel A shows municipality-level median residence-to-workplace distances, weighted by the number of worker-years in each municipality. It is a distribution of municipality medians. Panel B reports the share working outside their municipality of residence, among observations with an identified workplace municipality. Panel C shows the 75 largest inter-municipality flows; arrows point from residence to workplace and become thicker with the number of worker-years. Boundaries are municipal and provincial boundaries. Panel D reports the worker-year-weighted share with distance above 30 km. Distances are straight-line distances. The province ordering is the same in panels B and D.

Table A2: Sample Construction

Stage	Rows	Workers
Raw contracts, 1990–2001	23,134,946	2,663,270
Reference employment records, 1995–2000	7,386,393	1,502,093
Annual primary jobs, 1995–2000	6,240,752	1,502,093
Strict E2E transitions, raw	617,417	–
Valid origin characteristics	546,138	–
Residence in Veneto	518,710	–
Eligible destination in both years	449,335	–
Training transitions, 1995–1999	358,412	–
Validation transitions, 2000	90,923	–
Worker-year panel with IV	4,143,521	1,134,186
Employed at t and $t+1$	3,807,791	–
Change-on-change sample	2,944,142	–
Displaced worker-events, employer size ≥ 20	28,338	–
Displaced worker-events, employer size ≥ 10	45,018	–

Notes: Rows describe distinct samples rather than a single sequential funnel. Raw contracts are annual employment records used in preparation. Reference employment records retain multiple jobs within a worker-year; annual primary jobs select the highest weekly wage, breaking ties by tenure and then employer identifier. Strict employer-to-employer (E2E) transitions are employer changes between adjacent calendar months, without intervening months of observed nonemployment. Valid origin characteristics require an eligible origin firm-year and matching worker characteristics; usable worker coordinates and residence in Veneto are then required. Eligible destinations must be present in the candidate-firm data in both the origin and arrival years. The resulting transitions are split into training and validation samples by arrival year. These transition counts are separate from the annual primary-job counts above. The worker-year panel with IV matches the estimated inclusive value to the worker's annual primary employer in 1995–1999. Employment at t and $t+1$ requires an observed job in both years, including stayers and employer changers; the change-on-change sample additionally requires the index at both dates. The final two rows count displaced worker-events before restrictions on index or outcome availability. They use minimum pre-layoff employer sizes of 20 and 10, respectively, an employment decline of at least 50%, and a screen requiring that the most common first post-layoff employer receive less than 30% of displaced workers. They are not the matched event-study samples. The Workers column counts distinct individuals within the indicated sample; dashes denote counts not reported.

E Measurement and Model Diagnostics

E.1 Estimation of the Neural Destination-Choice Model

Estimation sample and destination pools. I use the 358,412 employer-to-employer transitions with arrival years 1995–1999 described in Section 3. Transitions with arrival year 2000 are reserved for predictive validation. Worker and origin-employer characteristics refer to the origin year, while destination attributes refer to the arrival year. Appendix Table A3 gives the complete input definitions and transformations. Standardization moments are fitted on the 1995–1999 observations and then applied to validation observations.

The empirical pool contains employers with at least five employees. Their two-digit industry must either belong to the origin employer's one-digit industry group or receive observed transitions from the origin industry during 1990–1994. This industry restriction defines the empirical pool before computational sampling. Sampling employers from that pool during training is a separate operation; the 200 alternatives used in a training observation do not define the set subsequently used for prediction and inclusive-value scoring.

Prediction objective and sampled alternatives. Index observed transitions by n, let j_n denote the realized destination, and let $v_{nj}(\theta)$ denote the neural utility assigned to employer j. Without sampling and regularization, the negative conditional log-likelihood over the empirical pools $\mathcal{C}_n$ would be

$$\mathcal{L}_{\text{full}}(\theta) = \frac{1}{N_{\text{tr}}} \sum_{n=1}^{N_{\text{tr}}} \left[\log \left(\sum_{j \in \mathcal{C}_n} \exp(v_{nj}(\theta)) \right) - v_{nj_n}(\theta) \right]. \tag{24}$$

Evaluating every employer for every transition at each parameter update is computationally expensive. I therefore always include the realized destination and draw $M = 199$ alternative employers with replacement. The draw probability is proportional to destination employment plus ten. Conditional on drawing an employer other than the realized destination, the proposal probability is

$$q_n(j) = \frac{N_{j,t_n^d} + 10}{\sum_{k \in \mathcal{C}_n \setminus \{j_n\}} (N_{k,t_n^d} + 10)}, \qquad j \in \mathcal{C}_n \setminus \{j_n\}, \tag{25}$$

where t_n^d is the arrival year and N_{j,t_n^d} is employer size. The positive offset gives smaller employers positive sampling mass while directing more computation toward larger employers. Repeated draws of an employer remain separate sampled entries.

Sampling alternatives for logit estimation has a long history: McFadden (1978) establishes the conditional-likelihood approach with adjustments for the sampling protocol. The neural implementation here uses an importance-adjusted sampled softmax, following the computational approach of Bengio and Sénécal (2003). Sampled negative scores are adjusted using their log proposal probabilities and the number of draws, while the realized destination is always included. This accounts for the unequal frequency with which employers enter the numerical sample. With finitely many draws, sampled softmax is an approximation to the full-softmax objective rather than an exact evaluation of equation (24); see Blanc and Rendle (2018).

Let $j_{n1}, \ldots, j_{nM}$ denote the sampled alternatives. The adjusted scores are

$$\widetilde{v}_{n0}(\theta) = v_{nj_n}(\theta), \qquad \widetilde{v}_{nc}(\theta) = v_{nj_{nc}}(\theta) - \log(Mq_n(j_{nc})), \quad c = 1, \ldots, M. \tag{26}$$

Applying softmax to these scores gives the sampled probability assigned to the realized destination:

$$\widetilde{p}_{n0}(\theta) = \frac{\exp(v_{nj_n}(\theta))}{\exp(v_{nj_n}(\theta)) + \frac{1}{M}\sum_{c=1}^{M} \frac{\exp(v_{nj_{nc}}(\theta))}{q_n(j_{nc})}}. \tag{27}$$

The realized destination enters the denominator exactly, while the remaining sum is approximated by the average of the importance-weighted draws. The sampling adjustment is used for training, not for the utility scores subsequently aggregated into the inclusive value.

Training loss and optimization. Let $\widetilde{p}_{nc}(\theta) = \exp(\widetilde{v}_{nc}(\theta)) / \sum_{d=0}^{M} \exp(\widetilde{v}_{nd}(\theta))$ denote the sampling-adjusted softmax probability for entry c, with $c = 0$ denoting the realized destination and $c = 1, \ldots, M$ the sampled alternatives. To discourage excessively confident predictions, I use label smoothing: the training target assigns 99% of its weight to the observed destination and distributes the remaining 1% uniformly across all 200 entries in the sampled choice set. Writing y_{nc}^{η} for these training targets, with $\eta = 0.01$, the cross-entropy contribution is

$$\ell_n(\theta) = -\sum_{c=0}^{M} y_{nc}^{\eta} \log \widetilde{p}_{nc}(\theta). \tag{28}$$

Gradients are computed by backpropagation through both projection layers and the final utility network, so all network parameters are updated jointly. I use Adam with a learning rate of 10^{-3}, weight decay of 10^{-4}, batches of 128 transitions, and gradient-norm clipping at one. Weight decay discourages large parameter values, while clipping limits unusually large optimization steps. The architecture has rank twelve, two hidden layers of twenty units with exponential linear unit activations, and a scalar output, giving 2,021 trainable parameters.

Validation and selection of the fitted model. After each training epoch, I evaluate destination rankings on a fixed sample of 500 year-2000 validation transitions. For these observations, the realized destination is ranked against the full empirical pool rather than 199 sampled employers; no training-sampling adjustment is applied to the predicted utility scores. Recall@K is the fraction of observations for which the realized destination is among the K highest-ranked employers. I retain the parameter values with the highest Recall@10. Training allows at most fifty epochs and stops after eight epochs without improvement in this criterion. This validation rule selects the fitted parameters; it differs from minimizing the training cross-entropy loss.

E.2 Utility Approximation and the Inclusive Value

Fix a worker, current employer, and year, and a common nonempty finite destination set $\mathcal{J}_{it}$. Suppress the worker, origin-employer, and time subscripts, and write v_j and $\widehat{v}_j$ for two vectors of destination utilities. Define $I = \log \sum_j \exp(v_j)$ and $\widehat{I} = \log \sum_j \exp(\widehat{v}_j)$. Then

$$|\widehat{I} - I| \leq \max_{j \in \mathcal{J}_{it}} |\widehat{v}_j - v_j|. \tag{29}$$

Proof. Let $\delta = \max_{j \in \mathcal{J}_{it}} |\widehat{v}_j - v_j|$. For every destination, $v_j - \delta \leq \widehat{v}_j \leq v_j + \delta$. Exponentiating and summing over the common destination set gives

$$e^{-\delta} \sum_j e^{v_j} \leq \sum_j e^{\widehat{v}_j} \leq e^{\delta} \sum_j e^{v_j}.$$

Taking logarithms yields $I - \delta \leq \widehat{I} \leq I + \delta$, which proves equation (29). □

E.3 Estimation of the Benchmark Destination-Choice Models

The two benchmarks in Figure 3 use the same source transition files, annual covariate definitions, and historical-industry destination-pool construction as the neural model. Training observations have arrival years 1995–1999; year-2000 observations are used for validation. The model-specific inputs and numerical training procedures are described below.

Type-specific multinomial logit. Let $c(i)$ denote the worker's type, as defined in Section 3. The benchmark assigns each type a coefficient vector on the nineteen destination and five pair characteristics:

$$v_{i\ell jt}^{\mathrm{MNL}} = \beta_{c(i)}' \begin{bmatrix} x_{jt} \\ m_{i\ell jt} \end{bmatrix}. \tag{30}$$

The conditional choice probabilities are the softmax of these utilities. Each type can therefore attach different weights to employer size, distance, industry, and the other inputs. Individual characteristics enter through type membership and the worker-specific pair characteristics; the forty-four worker/current-employer inputs do not enter the linear score separately.

Each training observation contains the realized destination and 199 alternatives drawn with replacement, with probabilities proportional to employer size plus ten. I use the sampling correction in equation (26) and label smoothing of 0.01. Coefficients are initialized at zero and updated using AdamW, with learning rate 10^{-3}, weight decay 10^{-4}, batches of 256 transitions, and gradient-norm clipping at one. Alternative draws are refreshed on successive passes through the data.

XGBoost. The tree-based benchmark assigns a score $s_{i\ell jt} = g(z_{i\ell t}, x_{jt}, m_{i\ell jt})$ using all sixty-eight observed inputs. The function g is an additive ensemble of decision trees, which can capture nonlinearities and interactions without the neural model's bilinear matching restriction. Scores are converted to probabilities by softmax within each transition's candidate group.

I construct one sampled training dataset containing the realized destination and fifty alternatives for each usable transition, using the same size-weighted sampling proposal. This gives 358,196 transition groups and 18,267,996 worker–destination rows. The observed destination receives target one and the sampled alternatives target zero. Trees are

fitted using a custom grouped-softmax cross-entropy loss,

$$\ell_n^{\text{XGB}} = -s_{nj_n} + \log\left(\sum_{j \in \mathcal{D}_n} \exp(s_{nj})\right), \tag{31}$$

Here, $\mathcal{D}_n$ denotes the sampled group; repeated draws remain separate entries. Scores enter this grouped loss without a proposal-probability offset or label smoothing. The tree benchmark is thus trained directly on the sampled prediction groups.

I use histogram-based tree construction with a maximum depth of eight and a learning rate of 0.05. The row and feature subsampling fractions are both 0.8. The minimum child weight is ten and the leaf-weight penalty is one. Training allows up to 800 boosting rounds, with early stopping after fifty rounds without improvement in grouped validation cross-entropy, evaluated on 50,000 sampled year-2000 transition groups.

Full-pool evaluation. For the comparison in Figure 3, I evaluate both fitted benchmarks on the same 4,932 year-2000 transitions used for the neural-model comparison. Each model scores eligible employers in the corresponding arrival-year pool, requiring destination coordinates for construction of the geographic inputs.

E.4 Alignment of Utility Scores Across Worker Folds

Cross-fitting produces five separately estimated networks. Their utility scores may differ in their average level and in how widely they vary across employers. Before combining their predictions, I adjust each network's scores to match a common reference: the model estimated on the full training sample. The adjustment has two parts. Multiplication changes the spread of the scores; addition changes their average level.

To calculate these adjustments, I draw a sample of 10,000 worker–employer–year observations from the reference model's scoring panel. For each observation, I obtain utility predictions from two models: the reference model and the network trained without that worker's group. Both models score the same employers for that observation. Let $\mathcal{S}_f$ be the observations in this sample whose workers belong to group f. I use $h = 0$ to denote the reference model and $h = -f$ to denote the network trained without group f.

First, consider one observation r. For each of the two models, I calculate the average utility across employers and the standard deviation around that average:

$$\mu_{rh} = \frac{1}{|\mathcal{C}_r|} \sum_{j \in \mathcal{C}_r} v_{rj}^h, \qquad s_{rh} = \left[\frac{1}{|\mathcal{C}_r|} \sum_{j \in \mathcal{C}_r} (v_{rj}^h - \mu_{rh})^2 \right]^{1/2}. \tag{32}$$

Here, v_{rj}^h is model h's predicted utility for employer j, and $\mathcal{C}_r$ is the employer pool scored for observation r. This calculation includes the current employer. The mean μ_{rh} measures the level of the scores, while s_{rh} measures how much they differ across employers for this worker-year.

Next, I average these worker-year means and standard deviations within each group:

$$\bar{\mu}_{hf} = \frac{1}{|\mathcal{S}_f|} \sum_{r \in \mathcal{S}_f} \mu_{rh}, \qquad \bar{s}_{hf} = \frac{1}{|\mathcal{S}_f|} \sum_{r \in \mathcal{S}_f} s_{rh}.$$

Thus, I first measure the spread across employers separately for each worker-year, then average those spreads. I do not calculate a single standard deviation by pooling all workers' destination scores.

These averages determine one multiplier b_f and one shift a_f for each group:

$$b_f = \frac{\bar{s}_{0f}}{\bar{s}_{-f,f}}, \qquad a_f = \bar{\mu}_{0f} - b_f \bar{\mu}_{-f,f}. \tag{33}$$

The multiplier makes the group's average standard deviation equal to the reference model's. After this rescaling, the shift makes the average utility equal as well.

Finally, I apply these two adjustments to the predictions used to construct the inclusive value:

$$\widetilde{v}_{i\ell jt} = a_{f(i)} + b_{f(i)} v_{i\ell jt}(\widehat{\theta}^{(-f(i))}). \tag{34}$$

Every worker in group f receives the same multiplier and shift, for all years and all destination employers. I then exclude the current employer and aggregate the adjusted utilities into the inclusive value using the log-sum formula. The later standardization of I_{it} for the wage and mobility regressions is a separate step, applied to the resulting worker-year index.

Table A3: Inputs to the Neural Destination-Choice Model

Input	Construction	No.
A. Destination-employer characteristics: x_{jt} (19 inputs)		
AKM firm wage effect	Firm component from the worker–firm fixed-effect fit to log weekly earnings residualized on year indicators and cubic age, using 1995–2000. Missing effects are replaced by the training firm-year mean. Standardized.	1
Employer size	$\log(1 + N_{jt})$, where N_{jt} is the number of distinct workers recorded at the employer during the year. Standardized.	1
Mean workforce tenure	Mean reconstructed tenure in years across the employer's annual panel records. Standardized.	1
Employment growth	$\log(N_{jt}/N_{j,t-1})$, requiring consecutive years. Winsorized at the training 1st and 99th percentiles; missing values replaced by the training median. Standardized.	1
Industry indicators	One-digit ATECO 81 categories 1–9, derived from the first digit of the two-digit industry code, plus an unknown-category indicator.	10
Female employment share	Fraction of annual employer panel records classified as female; missing shares use the reconstructed panel share, then 0.5 if unavailable.	1
Growth-category indicators	Indicators for the bottom and top training growth terciles; the middle tercile is omitted. Cutoffs use nonmissing training firm-years. These are relative growth categories.	2
Missing-value indicators	Separate indicators for unavailable employment growth and unavailable AKM firm effects before imputation.	2

Continued on next page

Table A3: Inputs to the Neural Destination-Choice Model (continued)

Input	Construction	No.
B. Worker and current-employer characteristics: $z_{i\ell t}$ (44 inputs)		
Age and age squared	Age in years, and its square, each standardized separately. Age is reconstructed as calendar year minus birth year, with a lower bound of 16.	2
Current-employer tenure	$\log(1 + \text{tenure in years})$, with tenure equal to the current year minus the first observed year at this employer, bounded below by zero. Standardized.	1
Prior employment experience	$\log(1 + \text{recorded employed months})$ accumulated before January of the first observed year at the current employer. Standardized; see history timing below.	1
Current-employer earnings[a]	Log real weekly earnings in the origin worker–employer year, standardized. Annual earnings in 2003 prices are divided by paid weeks, with the wage treatment described in Section 3.	1
Sex and nationality indicators	Female identifies women. Foreign nationality equals one for a reported nationality other than I (Italy). Blank fields are grouped with Italian nationality because an unreported nationality is not evidence of foreign nationality: the indicator measures recorded foreign nationality. Both indicators are zero for unmatched demographic records.	2
Any recent nonemployment	Indicator for at least one month without recorded employment in the 24 months before January of the first observed year at the current employer.	1
Recent nonemployment duration	$\log(1 + \text{months without recorded employment})$ in the corresponding preceding 60-month window. Standardized.	1
History-coverage indicators	Two indicators flag insufficient coverage before the start of the monthly dataset to reconstruct the 24- and 60-month histories. A third flags entry at the current employer before any pre-entry history is observable in that dataset. Unavailable earlier cumulative counts are set to zero; the indicators identify this coverage limitation.	3
Qualification indicators	First-character indicators pool contract variants within broad employment categories (e.g., 1FI and 1FD both enter category 1). Codes 1/6 denote blue-collar workers; 2, white-collar workers; 3/Q, managers; 5/R/W, apprentices; and O/Y, part-time blue-/white-collar workers. Other pools remaining or missing prefixes; a separate indicator covers categories absent from training. These describe employment status, not detailed occupations.	12
Current-employer attributes	Origin-employer counterparts of the four standardized continuous inputs in panel A, female share, two growth-category indicators, growth-missing indicator, and ten industry indicators, using the same transformations. Every eligible employer has an AKM estimate in this sample, so an origin AKM-missing indicator would be constant at zero. Unmatched origin-feature entries are set to zero.	18
Own-type share at current employer	Share of employment records associated with the worker's cluster at the origin employer, constructed as described below; unavailable matches are zero.	1
Own-type employment at current employer	$\log(1 + n_{\ell t, c(i)})$, where n counts annual panel records of the worker's type. Unavailable matches are zero before standardization over training worker–employer years.	1

Continued on next page

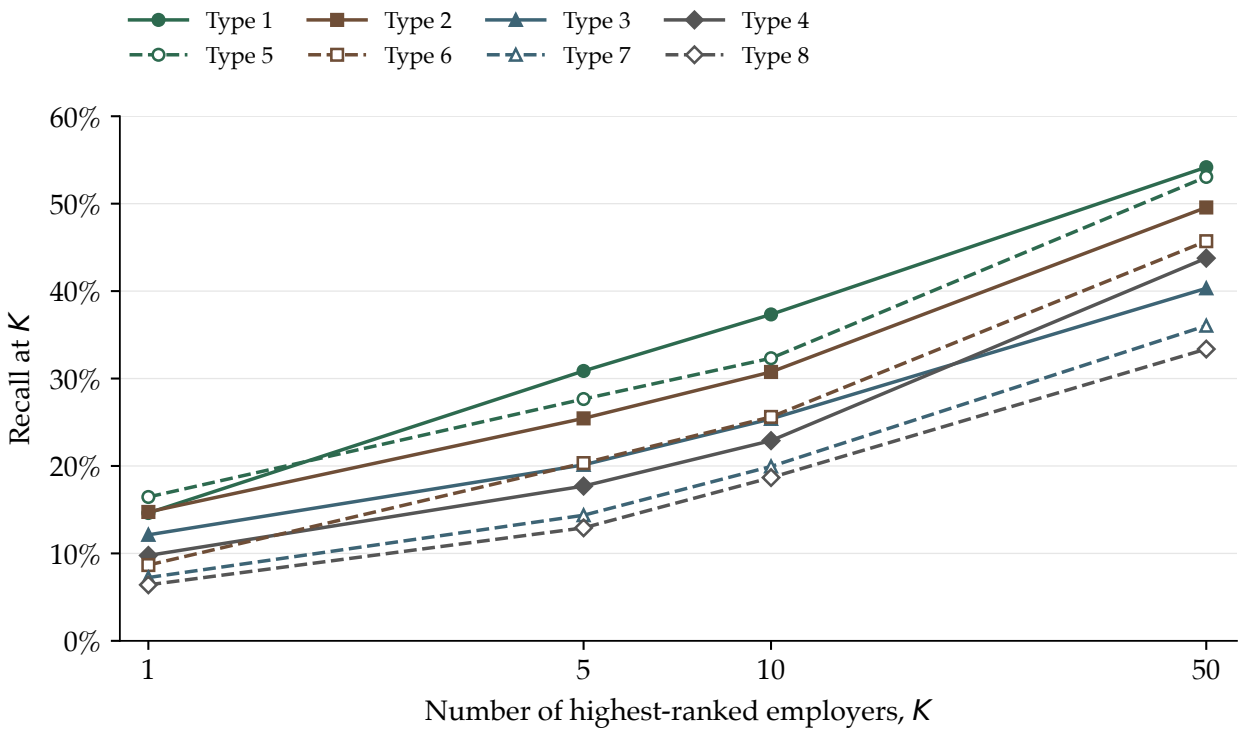
Table A3: Inputs to the Neural Destination-Choice Model (continued)

Input	Construction	No.
C. Worker–destination characteristics: $m_{i\ell jt}$ (5 inputs)		
Geographic distance	$\log(1 + d_{ij})$, where d_{ij} is great-circle distance in kilometres between reconstructed worker and destination-employer coordinates. Worker coordinates use residence municipality/postal matches. If residence information is missing or cannot be matched to coordinates after name repairs and municipality-code matching, I use the average coordinates of the worker's geocoded workplaces, weighted by paid days across employment records. Missing or nonpositive day weights are replaced by one.	1
Same province	Indicator that the origin employer and destination employer have the same province code.	1
Historical industry mobility	$\log(1 + F_{s(\ell),s(j)})$, where F counts observed employer-to-employer transitions between the origin and destination two-digit ATECO 81 industries in 1990–1994. Unobserved industry pairs have count zero.	1
Own-type employment at destination	$\log(1 + n_{jt,c(i)})$, using the annual panel-record count for the worker's type at the destination. No additional standardization; unavailable firm–year–type matches or unavailable worker types are assigned zero.	1
Own-type share at destination	$\min\{1, \max\{0, n_{jt,c(i)}/N_{jt}\}\}$, with the numerator counting annual panel records and the denominator distinct workers. Unavailable matches or worker types are assigned zero.	1

Notes: The table covers all 68 scalar inputs of the deep neural network model: 19 destination inputs, 44 worker/current-employer inputs, and 5 pair inputs. Indicator and share inputs are not standardized. "Standardized" means subtracting the 1995–1999 training mean and dividing by its sample standard deviation: firm features use eligible firm-years; personal features and origin own-type employment use eligible worker–employer years. Origin counterparts retain the firm-feature scaling. Pair inputs receive only the transformations shown.
For training transitions, worker and current-employer attributes refer to the origin year and destination attributes to the arrival year; annual inclusive-value scoring evaluates destination attributes in the year being scored. The employment-history inputs are anchored before January of the first observed year at the employer and remain fixed within that worker–employer spell, rather than being rolling histories. Months without recorded employment indicate absence from covered employment, not registered unemployment.
Type composition uses the clustering described in Section 3. The code counts panel records by employer, year, and type, whereas total firm size counts distinct workers; shares are bounded to $[0, 1]$. The current-employer share uses the same construction as the destination share. Worker type is a lookup key for these composition measures, not a separate input indicator. Province indicators constructed upstream are excluded from this specification.

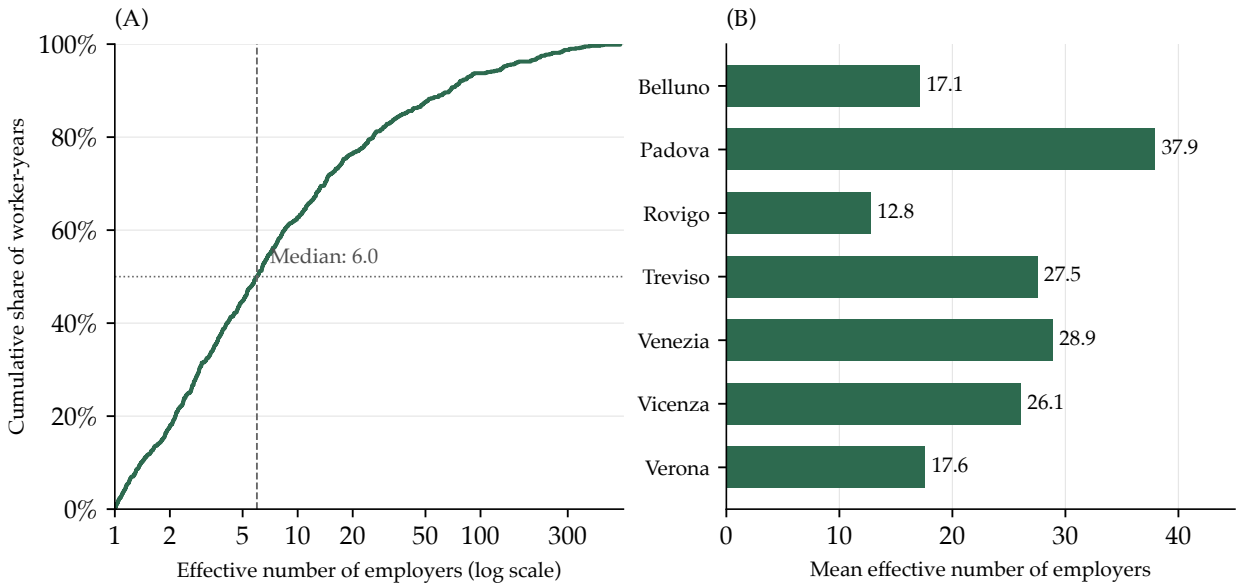
[a] In unreported specifications, I exclude workers' own earnings from the construction of the inclusive value to avoid potential leakage into wage regressions. The wage-regression results remain qualitatively unchanged.

Figure A3: Destination-Choice Prediction by Worker Type



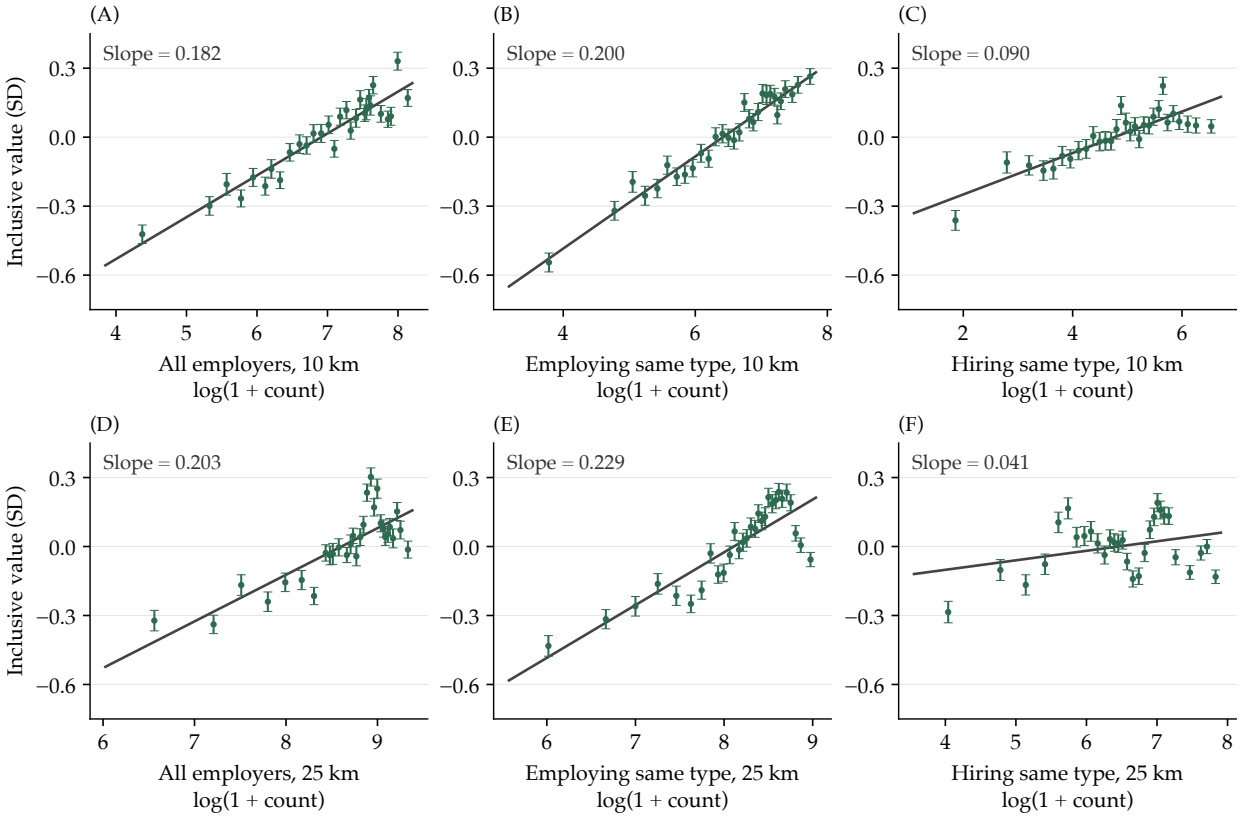
Notes: Recall at K is the share of validation transitions whose realized destination appears among the K highest-ranked outside employers. The curves reproduce the saved worker-type diagnostic for the neural destination model, based on 4,932 transitions arriving in 2000. For each worker type, the plotted values are unweighted averages of recall across the seven origin-province cells; they therefore use a different aggregation from the pooled recall in Figure 3. Types are the eight clusters described in Section 3. Markers show the observed cutoffs $K = 1, 5, 10, 50$; the horizontal axis is logarithmic.

Figure A4: Effective Number of Alternative Employers



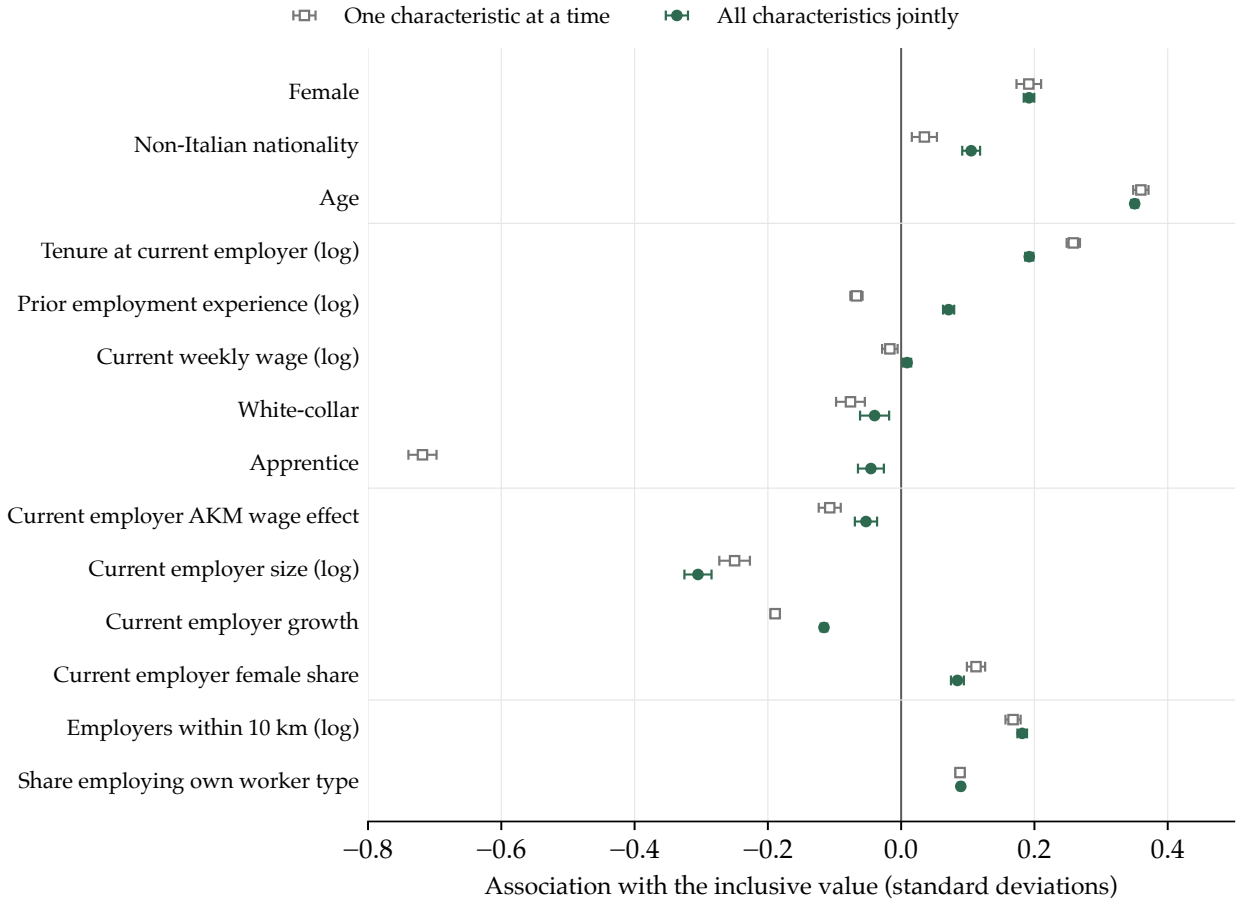
Notes: For each worker-year, the effective number of employers is $N_{it}^{\text{eff}} = \left[\sum_{j \in \mathcal{J}_{it}} \widehat{P}^2_{i\ell jt|\text{Leave}}\right]^{-1}$, using predicted destination probabilities and excluding the current employer. It equals N when N employers each receive probability $1/N$. The figure uses the original diagnostic sample of 800 worker-years and the full-training reference network. Panel A plots the empirical cumulative distribution on a logarithmic horizontal axis, retaining the complete right tail. The median is 6.0 and 62.8% of observations have an effective number no greater than ten. Panel B shows arithmetic means by province of residence. Province means are means of individual inverse concentration indices, rather than inverses of average concentration.

Figure A5: Outside Options and Nearby Employers



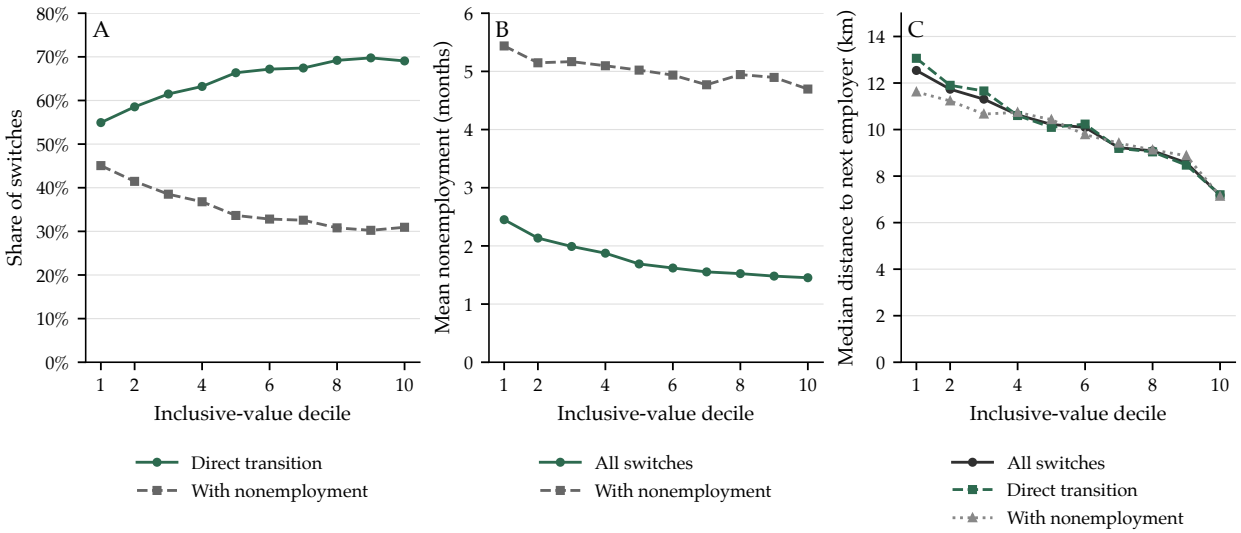
Notes: The sample contains 75,000 worker-year observations from the saved cross-fitted diagnostic. The inclusive value is standardized using the mean and population standard deviation of this diagnostic sample. Horizontal axes measure log one plus the number of employers within 10 km (panels A–C) or 25 km (D–F) of residence. Panels A and D count all employers; B and E count employers with at least one employee of the worker's type; C and F count employers with at least one observed direct transition hiring that type in the corresponding year. Counts include the incumbent when it meets these conditions. Green points are means within up to 30 quantile bins; capped vertical bars show 95% confidence intervals, calculated as the bin mean plus or minus 1.96 times its standard error. Dark lines are the original full-sample linear fits, with their slopes reported in each panel. All panels use the same vertical scale. These are descriptive relationships.

Figure A6: Who Has Good Outside Options?



Notes: Coefficients from regressions of the standardized cross-fitted inclusive value on worker, current-employer, and local-market characteristics. Grey squares include one characteristic at a time; green circles include all displayed characteristics jointly. Every specification includes year fixed effects. Continuous characteristics are standardized, so coefficients correspond to a one standard deviation difference; indicators enter in their original units. Whiskers show 95% confidence intervals, with standard errors clustered by current employer. Local-market measures count employers within 10 km of residence and the share employing workers of the individual's type. The sample contains 4,124,450 worker-years, 1,128,941 workers, and 63,314 employers during 1995–1999. The joint specification has $R^2 = 0.33$.

Figure A7: Outside Options and Employer Transitions



Notes: Workers changing employers are grouped into deciles of their cross-fitted inclusive value; decile 1 has the lowest values. The sample contains 95,444 switches for which the intervening monthly employment history can be reconstructed. Panel A reports the shares moving directly to another employer or experiencing intervening months without covered employment. Panel B reports average months without covered employment before the next employer, including direct switches as zeros and separately conditioning on intervening nonemployment. Panel C reports the median distance from the worker's residence to the next employer, for all switches and separately by transition path.

F Displacement Validation

F.1 Construction of the Matched Displacement Sample

I identify mass layoffs from changes in employers' annual employment, measured by the number of workers for whom the firm is their primary employer. A candidate event occurs when a firm located in Veneto employs at least ten workers in year T and its employment falls by at least 50% between T and $T+1$. Workers employed at the firm in T are classified as displaced if it is no longer their primary employer in $T+1$, including when they have no recorded employment. Event year T therefore denotes the year preceding the observed employment contraction.

To distinguish displacement from changes in firm identifiers or collective transfers, I examine where departing workers subsequently find employment. Using the monthly histories, I identify each worker's first observed employer after their last month at the origin firm in year T. I exclude an event if at least 30% of its displaced workers subsequently appear at the same employer, including returns to the origin firm. The denominator includes all displaced workers, whether or not subsequent employment is observed. This screen reduces 6,683 candidate events to 2,127 events, involving 45,018 displaced worker-events over 1995–1999. An inclusive value is available for 38,958 of these worker-events.

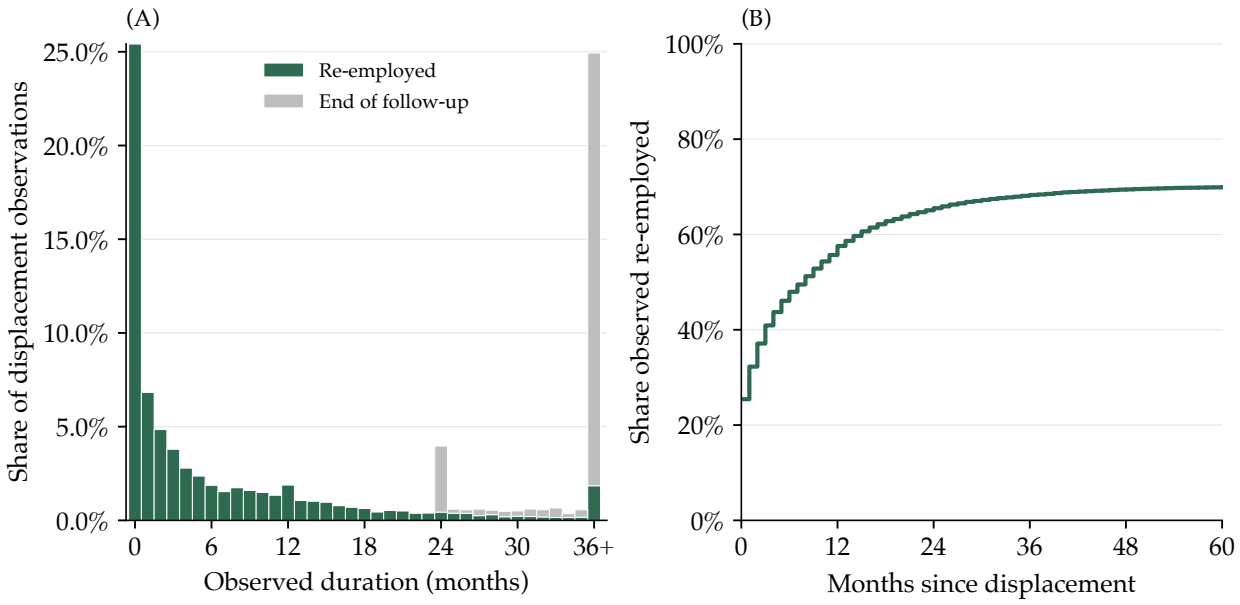
For the matched event study, I retain each worker's earliest displacement with an observed inclusive value and restrict event years to 1996–1998. This leaves 23,955 workers. I require non-missing matching characteristics in $T-1$ and observed employment in both $T-2$ and $T-1$, leaving 12,405 eligible displaced workers. The comparison pool comprises workers who remain with the same employer in $T-1$, T, and $T+1$, are observed in the annual reference panel throughout $T-2$ to $T+2$, and are not identified as displaced in the relevant year. These restrictions yield 1,103,705 eligible comparison worker-years.

I match each displaced worker to one comparison worker in the same event year, worker type, gender and administrative qualification category. Within these groups, matching selects the nearest neighbour in age, tenure, log weekly wages and the employer's AKM wage effect, all measured in $T-1$. Each continuous variable is divided by its standard deviation in the pooled eligible displaced and comparison samples. A match is retained when the Euclidean distance between these standardized characteristics does not exceed 0.4. Matching permits replacement, so a comparison worker can be matched to more than one displaced worker.

The resulting sample contains 11,595 matched pairs, covering 93.5% of eligible displaced workers and using 10,218 distinct comparison worker-years. The final displacement cohorts are 1997 and 1998, with 6,700 and 4,895 displaced workers respectively, drawn from 856 employers. I divide displaced workers according to whether their inclusive value in T lies above or below the median among the 23,955 candidates, before imposing the matching restrictions. The matched sample contains 5,982 workers above the median and 5,613 below it; each comparison worker is assigned to the group of their matched displaced worker.

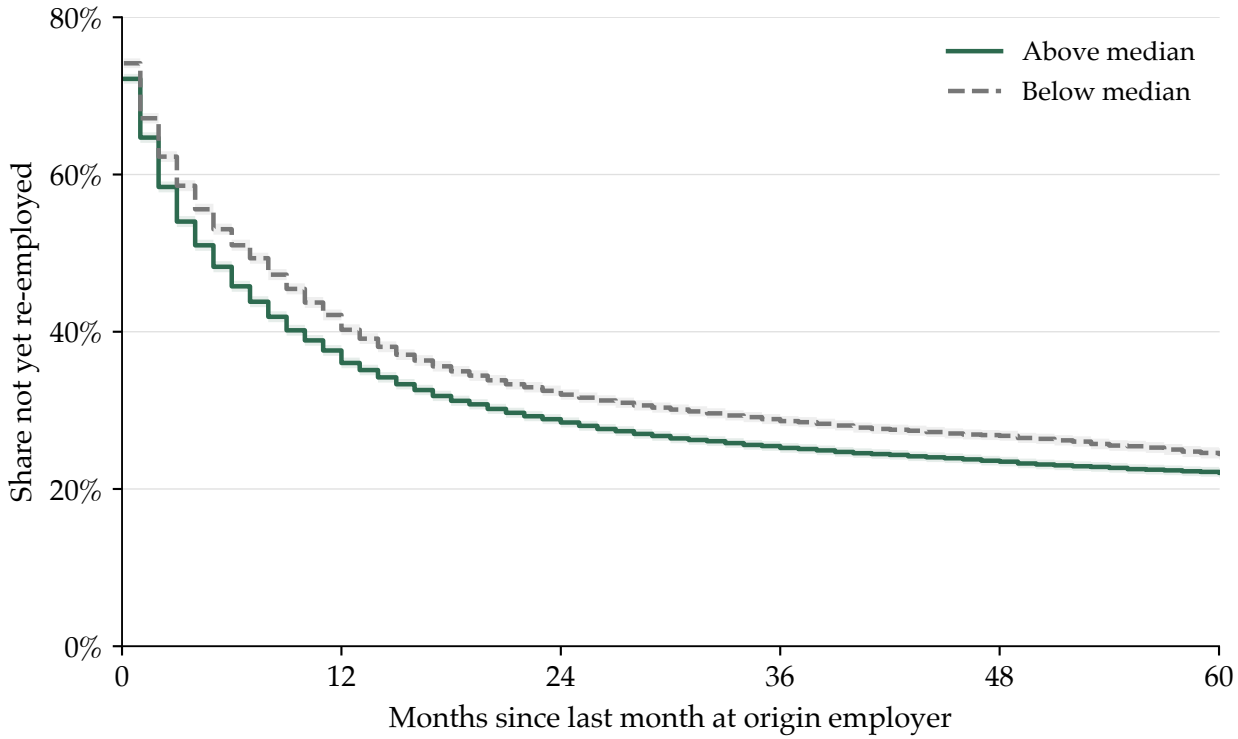
I follow both members of each pair from $T-3$ through $T+3$, producing 162,330 pair-member-year observations. Displaced workers are retained regardless of subsequent employment. Years without recorded employment contribute zero earnings and days worked, while wage and employer-characteristic outcomes use observations for which those measures are available.

Figure A8: Nonemployment and Re-employment After Displacement



Notes: The figure uses the original descriptive displacement sample of 23,769 observations dated 1996–1999. This sample is distinct from the matched event-study sample. Panel A distinguishes observed re-employment from the end of follow-up without observed re-employment. Durations are grouped into one-month bins; the final bin pools observed durations of 36 months or more. Direct transitions appear in the first bin. Both components use all displacement observations as the denominator. Panel B reports the unadjusted cumulative share observed re-employed within each number of months, also using all displacement observations as the denominator. It does not adjust for right-censoring. Nonemployment denotes absence from covered employment, rather than registered unemployment. Kaplan–Meier estimates by pre-displacement outside options are shown separately in Figure A9.

Figure A9: Re-employment After Mass Layoffs by Pre-Displacement Outside Options



Notes: Kaplan–Meier estimates of the share of displaced workers not yet observed in re-employment, split at the median pre-displacement inclusive value in the full displacement sample. The sample comprises 38,958 worker-events from 2,119 mass-layoff events at employers with at least ten employees over 1995–1999, before the restrictions defining the matched event-study sample. Each group contains 19,479 worker-events. Duration is measured from the last month at the origin employer to the first subsequent observed employment spell; workers without observed re-employment are right-censored at the end of their observation window. Shaded bands show pointwise 95% confidence intervals. Curves are displayed through 60 months.

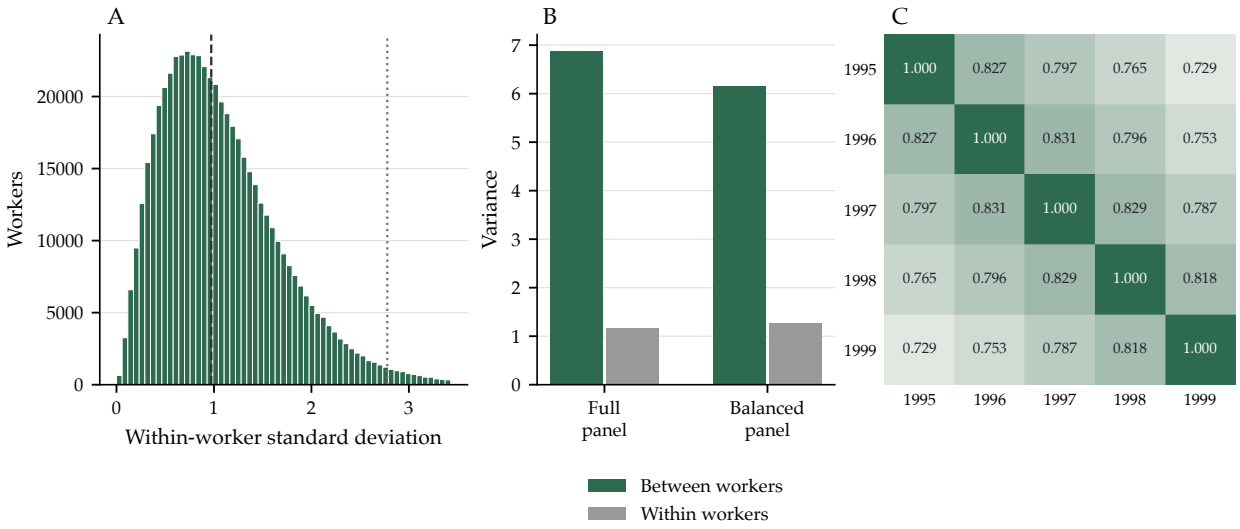
G Mobility, Wages, and Heterogeneity

Table A4: Outside Options and Mobility: Linear Probability Models

	(1) Employer change	(2) Broad leaving
Inclusive value (std.)	0.0098***	0.0082***
	(0.0004)	(0.0005)
Current log weekly wage	0.1505***	0.1642***
	(0.0014)	(0.0014)
Age squared (std.)	-0.1366***	0.4257***
	(0.0048)	(0.0057)
Log tenure (std.)	0.0708***	0.1117***
	(0.0008)	(0.0008)
Log prior employment experience (std.)	-0.0802***	-0.0491***
	(0.0011)	(0.0011)
Worker FE	Yes	Yes
Employer FE	Yes	Yes
Year FE	Yes	Yes
Eligible observations	3,807,791	4,143,521
Estimation observations	3,639,030	3,968,053
Workers	878,884	961,225
Employers	58,126	59,344
Outcome mean (estimation sample)	0.136	0.192

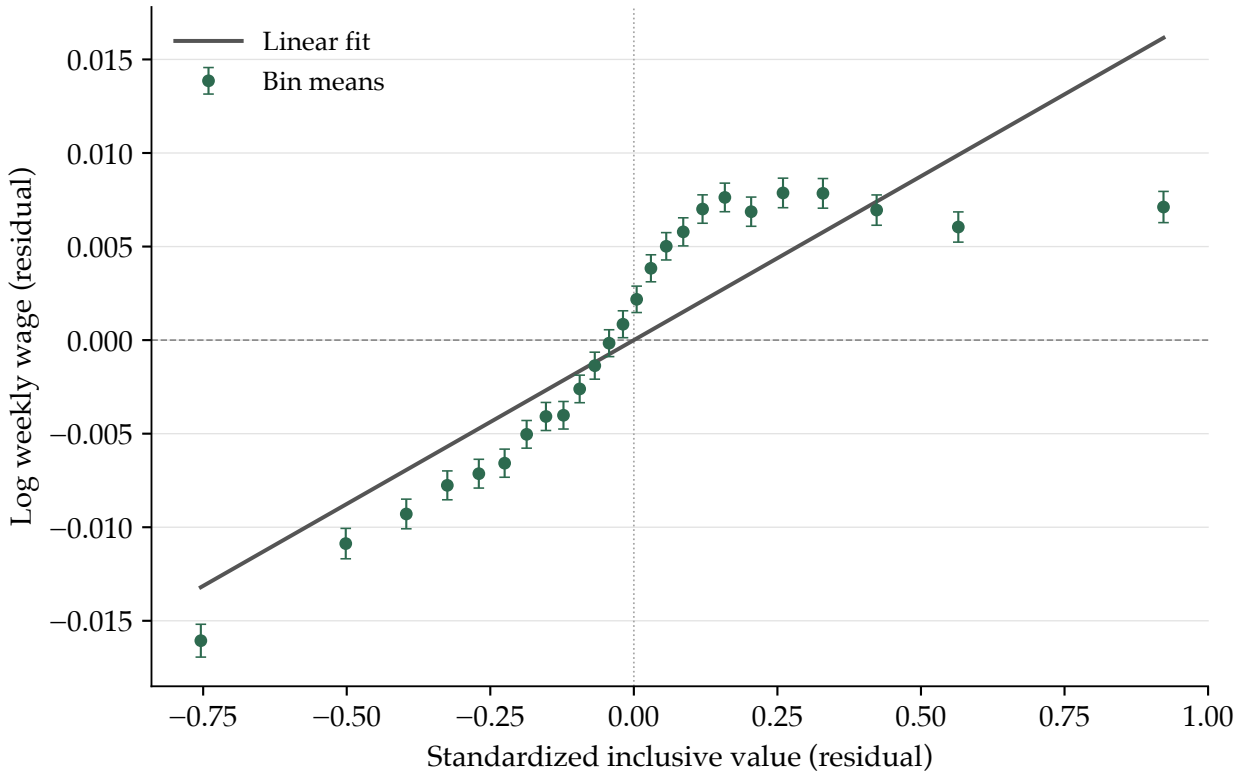
Notes: Linear probability models estimated by OLS. An employer change equals one when annual primary employers differ between t and $t+1$, conditional on observed employment in both years. Broad leaving also includes no recorded employment in $t+1$. Years are 1995–1999. The inclusive value is cross-fitted and standardized over the annual worker-year panel. The coefficient on the inclusive value is the probability change associated with a one standard deviation increase in the index. Age squared, log tenure and log prior employment experience use the stored standardized covariates. Worker, employer and year fixed effects are included. Standard errors in parentheses are heteroskedasticity-robust. Only recursive singleton removal reduces the eligible sample; observations with perfectly predicted binary outcomes are retained. The samples therefore differ from the logit samples in Table 3. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A10: Variation in Outside Options Within and Between Workers



Notes: The figure uses the cross-fitted inclusive value in its original units during 1995–1999. Multiple observations for a worker in the same year are averaged. Panel A shows the distribution of within-worker standard deviations among the 558,258 workers observed in all five years: each is the sample standard deviation of one worker's five annual inclusive values. The histogram ends at the 99.5th percentile. The dashed line marks the median within-worker standard deviation (0.97), approximately 35% of the pooled standard deviation of the inclusive value across all worker-years in the full panel (2.78, dotted line). Panel B reports the variance of worker-specific mean inclusive values and the mean squared deviation from those means, for the full and balanced panels. The first gives equal weight to workers, while the second gives equal weight to worker-years; the components need not sum to the pooled variance in the unbalanced panel. Panel C reports Pearson correlations across years among balanced-panel workers. The full panel contains 4,182,765 worker-years and 1,139,984 workers.

Figure A11: Outside Options and Wages Net of AKM Fixed Effects



Notes: Both log real weekly wages and the standardized, cross-fitted inclusive value are first residualized on worker, employer and year fixed effects, using the same 3,968,053 worker-year observations from 1995–1999. Points report mean residual wages against the mean residual inclusive value in 25 equally sized bins of the latter variable. Vertical bars equal the bin mean plus or minus 1.96 times its within-bin standard error, calculated as the standard deviation of residual wages divided by the square root of the number of observations in the bin. The solid line uses the linear regression slope estimated on the full, unbinned sample (0.0175), corresponding to column (3) of Table 4.

Table A5: Wages and Mobility by Gender

Subsample	(1) Log weekly wages	(2) Employer change (LPM)	(3) Wage observations	(4) Mobility observations
Pooled	0.0133***	0.0116***	3,968,053	3,639,030
	(0.0010)	(0.0028)		
Men	0.0192***	0.0153***	2,427,855	2,233,121
	(0.0013)	(0.0044)		
Women	0.0055***	0.0094***	1,534,330	1,400,226
	(0.0010)	(0.0020)		

Notes: Columns (1) and (2) report the coefficient on the standardized inclusive value from separate OLS regressions within each subsample. The wage outcome is log real weekly wages in t; mobility is an annual employer change between t and $t+1$, conditional on employment in both years, estimated as a linear probability model. Both regressions include worker, employer and year fixed effects and worker covariates; the mobility regressions in these tables do not include current wages. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Regressions are estimated separately for men and women; the gender indicator is omitted. The pooled row uses the full sample, including observations outside any displayed subgroup. Columns (3) and (4) report estimation observations for the respective outcomes. Standard errors in parentheses are clustered by the origin employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6: Wages and Mobility by Age

Subsample	(1) Log weekly wages	(2) Employer change (LPM)	(3) Wage observations	(4) Mobility observations
Pooled	0.0133***	0.0116***	3,968,053	3,639,030
	(0.0010)	(0.0028)		
Younger	0.0180***	0.0095***	1,286,631	1,179,496
	(0.0010)	(0.0018)		
Middle	0.0095***	0.0155***	1,283,818	1,193,116
	(0.0012)	(0.0031)		
Older	0.0096***	0.0157***	1,256,105	1,134,488
	(0.0015)	(0.0046)		

Notes: Columns (1) and (2) report the coefficient on the standardized inclusive value from separate OLS regressions within each subsample. The wage outcome is log real weekly wages in t; mobility is an annual employer change between t and $t+1$, conditional on employment in both years, estimated as a linear probability model. Both regressions include worker, employer and year fixed effects and worker covariates; the mobility regressions in these tables do not include current wages. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Age groups are terciles of standardized age in the annual panel. The pooled row uses the full sample, including observations outside any displayed subgroup. Columns (3) and (4) report estimation observations for the respective outcomes. Standard errors in parentheses are clustered by the origin employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A7: Wages and Mobility by Employer Tenure

Subsample	(1) Log weekly wages	(2) Employer change (LPM)	(3) Wage observations	(4) Mobility observations
Pooled	0.0133***	0.0116***	3,968,053	3,639,030
	(0.0010)	(0.0028)		
$<$ 2 years	0.0245***	0.0058	876,319	769,133
	(0.0018)	(0.0038)		
2–4 years	0.0094***	0.0138***	879,231	813,299
	(0.0014)	(0.0046)		
5 or more years	0.0075***	0.0155***	1,719,685	1,603,622
	(0.0015)	(0.0048)		

Notes: Columns (1) and (2) report the coefficient on the standardized inclusive value from separate OLS regressions within each subsample. The wage outcome is log real weekly wages in t; mobility is an annual employer change between t and $t+1$, conditional on employment in both years, estimated as a linear probability model. Both regressions include worker, employer and year fixed effects and worker covariates; the mobility regressions in these tables do not include current wages. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Groups are defined by tenure at the current employer. The pooled row uses the full sample, including observations outside any displayed subgroup. Columns (3) and (4) report estimation observations for the respective outcomes. Standard errors in parentheses are clustered by the origin employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A8: Wages and Mobility by Employment Grade

Subsample	(1) Log weekly wages	(2) Employer change (LPM)	(3) Wage observations	(4) Mobility observations
Pooled	0.0133***	0.0116***	3,968,053	3,639,030
	(0.0010)	(0.0028)		
Blue-collar	0.0089***	0.0115***	2,516,488	2,302,576
	(0.0012)	(0.0032)		
White-collar	0.0067***	0.0179***	953,714	893,595
	(0.0020)	(0.0046)		
Apprentice	0.0106***	0.0107*	102,614	91,279
	(0.0026)	(0.0058)		
Other	0.0057**	0.0271***	226,265	201,741
	(0.0022)	(0.0047)		

Notes: Columns (1) and (2) report the coefficient on the standardized inclusive value from separate OLS regressions within each subsample. The wage outcome is log real weekly wages in t; mobility is an annual employer change between t and $t+1$, conditional on employment in both years, estimated as a linear probability model. Both regressions include worker, employer and year fixed effects and worker covariates; the mobility regressions in these tables do not include current wages. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Groups use administrative qualification codes: 1 (blue-collar), 2 (white-collar), 5 (apprentice), and all other codes. These are employment grades, not occupations. The pooled row uses the full sample, including observations outside any displayed subgroup. Columns (3) and (4) report estimation observations for the respective outcomes. Standard errors in parentheses are clustered by the origin employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A9: Wages and Mobility by Collective Agreement

Subsample	(1) Log weekly wages	(2) Employer change (LPM)	(3) Wage observations	(4) Mobility observations
Pooled	0.0133***	0.0116***	3,968,053	3,639,030
	(0.0010)	(0.0028)		
Metalworking	0.0101***	0.0025	723,442	678,350
	(0.0017)	(0.0070)		
Commerce / Services	0.0110***	0.0204*	414,966	384,157
	(0.0024)	(0.0104)		
Textile / Clothing	0.0020	0.0114***	302,205	277,406
	(0.0016)	(0.0043)		
Wood / Furniture	0.0118***	0.0134***	165,843	154,920
	(0.0019)	(0.0038)		
Construction	0.0263***	0.0240***	118,433	105,532
	(0.0023)	(0.0049)		
Footwear	0.0046	0.0221	51,610	47,925
	(0.0032)	(0.0189)		
Other	0.0125***	0.0178***	1,997,589	1,819,848
	(0.0018)	(0.0047)		

Notes: Columns (1) and (2) report the coefficient on the standardized inclusive value from separate OLS regressions within each subsample. The wage outcome is log real weekly wages in t; mobility is an annual employer change between t and $t+1$, conditional on employment in both years, estimated as a linear probability model. Both regressions include worker, employer and year fixed effects and worker covariates; the mobility regressions in these tables do not include current wages. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Collective agreement groups use the administrative contract code of the worker-employer-year record with the most reported months: metalworking (113, 116), commerce/services (042, 051), textile/clothing (003, 001), wood/furniture (100, 101), construction (068, 071), footwear (026), and other codes. Unmatched contract records are excluded from subgroup regressions. The pooled row uses the full sample, including observations outside any displayed subgroup. Columns (3) and (4) report estimation observations for the respective outcomes. Standard errors in parentheses are clustered by the origin employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A10: Wages and Mobility by Employer Growth

Subsample	(1) Log weekly wages	(2) Employer change (LPM)	(3) Wage observations	(4) Mobility observations
Pooled	0.0133***	0.0116***	3,968,053	3,639,030
	(0.0010)	(0.0028)		
Shrinking	0.0125***	0.0158**	674,629	611,493
	(0.0030)	(0.0076)		
Stable	0.0137***	0.0107*	1,528,909	1,419,380
	(0.0016)	(0.0060)		
Growing	0.0251***	0.0016	798,603	725,196
	(0.0034)	(0.0042)		

Notes: Columns (1) and (2) report the coefficient on the standardized inclusive value from separate OLS regressions within each subsample. The wage outcome is log real weekly wages in t; mobility is an annual employer change between t and $t+1$, conditional on employment in both years, estimated as a linear probability model. Both regressions include worker, employer and year fixed effects and worker covariates; the mobility regressions in these tables do not include current wages. Worker controls are age, age squared, gender, non-Italian nationality, log tenure and log prior employment experience; controls absorbed by fixed effects are omitted. The inclusive value is cross-fitted and standardized over the annual worker-year panel. Groups use standardized employer employment growth: shrinking below -0.3, stable between -0.3 and 0.3, and growing above 0.3. These cutoffs are in standard deviation units, not percentage growth rates. The pooled row uses the full sample, including observations outside any displayed subgroup. Columns (3) and (4) report estimation observations for the respective outcomes. Standard errors in parentheses are clustered by the origin employer. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

H Opening-Based Instruments

Table A11: Firm Openings and Employment, 1995–1999

Year	All openings	At least 50 employees		At least 100 employees		At least 200 employees	
		Openings	Employment	Openings	Employment	Openings	Employment
1995	12,653	70	8,463	27	5,483	5	2,624
1996	12,102	59	13,191	29	11,054	16	9,244
1997	11,365	73	10,887	29	8,030	13	5,914
1998	12,667	81	12,104	34	8,940	14	6,144
1999	14,170	110	18,012	51	14,029	21	10,005
Total	62,957	393	62,657	170	47,536	69	33,931

Notes: An opening is a firm active in a given year but inactive in the preceding year; activity in 1994 is used to classify openings in 1995. Employment is measured as the number of distinct workers observed at each firm during its opening year. The all-openings column counts opening events regardless of firm size. For each size threshold, the table reports the number of qualifying opening events and the sum of employment at these firms. Thresholds are nested. The total row sums the annual figures over 1995–1999, the estimation period for the opening-based instruments.

I Portugal: Data and Empirical Specifications

I.1 Worker Panel and Estimation Sample

The Portuguese analysis combines the October employment records in Quadros de Pessoal (QP) with annual firm accounts from the Sistema de Contas Integradas das Empresas (SCIE). I use QP records from 2010 onward to construct worker histories and employer characteristics; the wage and rent-sharing analyses cover 2013–2019. I retain salaried employees working under an employment contract (employment-status code 3, *trabalhadores por conta de outrem*), with a valid worker identifier. When a worker has several October jobs, the primary job is the one with the highest total monthly earnings, breaking ties by normal monthly hours and then employer identifier. Monthly earnings include base pay, regular bonuses and allowances, and overtime remuneration. Monetary variables are expressed in nominal euros; the regressions include year fixed effects.

An employer transition is a change in primary employer between consecutive October observations. This annual definition does not distinguish a direct employer-to-employer move from one involving an intervening period without recorded employment. Destination-choice estimation uses transitions with arrival years 2014–2018, while transitions arriving in 2019 are reserved for predictive evaluation. Historical region-to-region and industry-to-industry flows are calculated separately from transitions arriving in 2011–2013.

Eligible destination firms have at least five employees and a valid two-digit industry division in the Portuguese Classification of Economic Activities, Revision 3 (CAE Rev.3). Within this classification, three-digit groups and four-digit classes identify progressively more detailed industries. For an observed transition to enter estimation, its origin must belong to the eligible employer pool in the origin year, and its realized destination must be eligible in both the origin and arrival years. I also require the worker and employer characteristics needed to construct the model inputs, including the origin establishment's region. These restrictions yield 853,643 training transitions and 240,827 validation transitions before imposing the coverage requirements of the latent-distance measure described below. The final supported samples contain 809,390 training transitions and 213,545 validation transitions.

The available geography identifies NUTS II regions—the second level of the European Nomenclature of Territorial Units for Statistics—rather than municipalities or precise workplace coordinates. The worker's location in the destination-choice model is that of the origin establishment. Employer location is the region containing the largest share of its establishments' employment, with the firm-level region used when establishment

information is unavailable. I therefore use a same-region indicator and historical regional flows, together with a distance estimated from the mobility network, to describe proximity between employers.

For each worker-year, I count changes in primary employer and years without an observed primary job in the five-year window ending in that year. This window is truncated at 2010, the beginning of the available histories, and a separate indicator records this truncation. Years without a recorded primary job measure absence from covered employment, rather than registered unemployment.

I.2 Inputs and Neural Utility Specification

The destination-choice model retains the multinomial-logit structure and bilinear matching architecture of Section 5, with inputs adapted to the information available in QP. To describe evaluation of alternative employers for a worker-year, let $z_{i\ell t}$ collect worker and origin-employer characteristics, x_{jt} destination-employer characteristics, and $m_{i\ell jt}$ dyadic characteristics. The input groups are as follows.

Worker characteristics. The worker vector has 42 entries: age, age squared, log one plus tenure, the number of employer changes and years absent from the five-year history; indicators for gender, foreign nationality, fixed-term contracts, other contract categories, part-time work, missing age, tenure or contract information, and a truncated history; and indicators for education, one-digit occupation, and administrative qualification. These three classifications measure different dimensions: education records schooling, occupation describes the work performed, and qualification records the administrative employment grade.[2]

The education indicators distinguish schooling below primary education (code 1); basic education, comprising the four-, six-, and nine-year cycles (2); upper-secondary education (3); post-secondary non-tertiary education (4); the undergraduate degree categories *bacharelato* (5) and *licenciatura* (6); master's degrees (7); and doctorates (8).

Occupation indicators use the major, one-digit groups of the Portuguese Classification of Occupations (CPP 2010): managers and senior officials (code 1); professionals (2); technicians and associate professionals (3); clerical support workers (4); service and sales workers (5); skilled agricultural, forestry, and fishery workers (6); craft and related trades workers (7); plant and machine operators and assemblers (8); and elementary occupations

[2]Category labels follow the GEP's Relatório Único 2019 code tables, especially Tables 14–16 and 21.

(9). The two-digit occupation groups used in the dyadic inputs subdivide these categories into more specific fields of work.

Qualification indicators distinguish senior staff (code 1); intermediate staff (2); supervisors, foremen, and team leaders (3); highly skilled workers (4); skilled workers (5); semi-skilled workers (6); unskilled workers (7); and trainees and apprentices (8). These are reported employment grades, not educational qualifications. Each of the three sets of indicators also has a residual category for missing values or values outside its retained codes.

Employer characteristics. Both origin and destination employers are described by the same 29 variables. Six continuous inputs measure log employment, employment growth, mean worker tenure, mean worker age, log one plus firm age, and log number of establishments. Employer size uses reported October employment, falling back to annual average employment when necessary. Further inputs record growing and shrinking employment, missing growth and firm-age information, workforce shares by gender, fixed-term and part-time status, education and occupation, foreign ownership, and broad industry. Growth is the change in log employment, clipped to $[-2, 2]$; the growing and shrinking indicators use thresholds of 0.02 and -0.02. The education-composition variable is the share with higher education—*bacharelato*, *licenciatura*, master's degree, or doctorate (education codes 5–8). The occupation-composition variable is the share in managerial, professional, or technical occupations (major occupation groups 1–3). Broad industry is represented by twelve groups and an unknown category. Adding the 29 origin-employer inputs to the 42 worker inputs gives $z_{i\ell t} \in \mathbb{R}^{71}$, while $x_{jt} \in \mathbb{R}^{29}$.

Dyadic characteristics. The eleven dyadic inputs comprise a same-region indicator at the NUTS II level; log one plus historical flows between the origin and destination regions, two-digit industries, and three-digit industries; the log number and share of destination workers sharing the worker's one-digit occupation; the corresponding number and share for two-digit occupation; the destination share with the same education group; latent mobility distance; and an indicator for an imputed origin coordinate. Education is grouped into basic education or less (codes 1–2), upper-secondary or post-secondary non-tertiary education (3–4), and higher education (5–8), with a separate residual group. Numeric age, tenure and history inputs, the six continuous firm inputs, log flows, log matched-worker counts, and log latent distance are standardized using fitted training transformations. Shares and indicators enter in levels. These transformations are held fixed for validation and scoring. There are no earnings or estimated employer wage ef-

fects among the Portuguese model inputs, and no clustered worker-type inputs: observed occupation and education provide the compatibility measures.

Writing $b_{i\ell jt} = [x'_{jt}, m'_{i\ell jt}]' \in \mathbb{R}^{40}$, the matching representation is

$$h_{i\ell jt} = \left(A_z z_{i\ell t} + a_z\right) \odot \left(A_b b_{i\ell jt} + a_b\right) \in \mathbb{R}^{24}, \tag{35}$$

where $\odot$ denotes element-wise multiplication. The predicted utility is

$$\widehat{v}_{i\ell jt} = f_\theta\Big([x'_{jt}, m'_{i\ell jt}, h'_{i\ell jt}]'\Big). \tag{36}$$

Thus, the final network receives 64 inputs: 29 destination features, eleven dyadic features, and 24 matching features. It has two hidden layers of twenty units with exponential linear unit activations and a scalar output. Including the two projections, the specification has 4,453 estimated parameters. Worker and origin characteristics enter through the matching representation; destination and dyadic inputs also enter the final network directly.

I.3 Latent Mobility Distance and Destination Pools

To capture proximity beyond the coarse regional identifiers, I estimate a latent-space model of firm-to-firm mobility. Let M_{fg} denote the number of observed transitions from firm f to firm g in the 2014–2018 training-arrival sample. The model is

$$M_{fg} \sim \text{Poisson}(\mu_{fg}), \qquad \log \mu_{fg} = c + \alpha_f + \gamma_g + q'_{fg}\delta - \|r_f - r_g\|_2, \quad r_f \in \mathbb{R}^8. \tag{37}$$

Here, r_f is the eight-dimensional latent position of firm f. I estimate all firm positions $\{r_f\}_f$ jointly with the sending and receiving effects $\{\alpha_f, \gamma_f\}_f$, the intercept c, and the coefficients δ by maximizing a penalized Poisson log-likelihood.

Origin and destination effects account for differences in sending and receiving flows. The vector q_{fg} includes indicators for the same broad and two-digit industry, together with absolute differences in log size, mean tenure, mean age, female share, education and occupation composition, and foreign ownership. These firm characteristics are summarized over the training-arrival years. The distance is symmetric, while the origin and destination effects allow flows to be directed. It summarizes proximity in observed mobility, rather than physical distance.

I retain firms with at least three incident training transitions and restrict estimation to the largest weakly connected component. The model using all training transitions contains 59,166 firms. The penalized Poisson objective includes every positive-flow dyad

and a weighted sample of zero-flow dyads. Each epoch draws two zero-flow dyads per positive dyad; their contributions are weighted by the ratio of all zero-flow dyads to sampled zero-flow dyads. The distance coefficient is fixed at minus one. Coordinates are centered in each dimension and rescaled to a global standard deviation of one after each optimization step. Estimation uses Adam for eighty epochs, with learning rate 0.03 and quadratic penalties of 10^{-4} on coordinates, 10^{-6} on origin and destination effects, and 10^{-5} on covariate coefficients.

For training observations, the distance is estimated after excluding the worker's group from a five-way partition of workers. A separate model using all training transitions supplies distances for the 2019 predictive evaluation. When an origin firm has no estimated coordinate, I assign it the mean coordinate of supported firms in the same region and two-digit industry. If that cell is unavailable, I successively use region by broad industry, two-digit industry, region, and the overall mean. The model receives an indicator for this imputation. Destination firms must have estimated coordinates; their coordinates are not imputed. The distance input is standardized log one plus Euclidean distance, with standardization moments fitted on sampled training alternatives.

Before sampling alternatives for estimation, I restrict the annual eligible employer pool to industries in the same broad CAE Rev.3 section as the origin (identified by a letter, such as section C for manufacturing), or two-digit industries receiving observed flows from the origin industry in 2011–2013. I then retain destinations covered by the relevant latent-space model. These restrictions define the empirical destination pool. In particular, the number of sampled alternatives used to train the neural network does not determine the pool over which I construct the inclusive value.

I.4 Training, Cross-Fitting, and Predictive Evaluation

For each training transition, I include the realized employer and sample 199 alternatives with replacement. Approximately half are drawn from the full empirical pool and half from employers in the origin region; when that regional pool is too small, all draws use the full pool. Within each component, sampling weights are proportional to employment plus ten. Draws equal to the realized employer are resampled from the full pool. Let $q_n(j)$ denote the mixture probability before excluding the realized employer j_n. The implementation leaves its score unchanged and subtracts

$$\log\left(\frac{199q_n(j)}{1-q_n(j_n)}\right)$$

from sampled alternative scores. This is the sampling adjustment used in the neural loss; it is not applied to utilities when scoring the inclusive value. The importance-adjusted sampled-softmax objective and its relation to sampling alternatives are discussed in Appendix E.1.

Training uses the same 1% label smoothing described there, Adam with learning rate 10^{-3} and weight decay 10^{-4}, batches of 128 transitions, and gradient-norm clipping at one. I allow up to sixty epochs, stop after ten epochs without an improvement, and retain the network with the lowest unsmoothed, sampling-adjusted negative log-likelihood on the early-stopping sample.

To construct worker-year outside options, I estimate five networks, each excluding one group of workers from parameter fitting. A sample of up to 20,000 transitions from that omitted group is used for early stopping. This neural-model partition is distinct from the five-way partition used to estimate latent distance. Before aggregation, I align each network's utility scores to those of a reference model fitted on the training sample. The reference reserves a worker-based tenth of the sample for early stopping. Alignment follows the mean-and-dispersion procedure in Appendix E.4, using 9,000 worker-year observations from 2016–2018. In particular, for fold f, $b_f = \bar{s}_{0f}/\bar{s}_{-f,f}$ and $a_f = \bar{\mu}_{0f} - b_f \bar{\mu}_{-f,f}$, and I apply $a_f + b_f \widehat{v}^{(-f)}$ before taking the log sum.

For each worker-year t, I evaluate the fitted utility function over the eligible employer pool $\mathcal{P}_{it}^{(-h(i))}$ associated with the worker's latent-distance fold, after the industry and coordinate-coverage restrictions and exclusion of the incumbent. The index is

$$I_{it} = \log \left[\sum_{j \in \mathcal{P}_{it}^{(-h(i))}} \exp\left\{ a_{f(i)} + b_{f(i)} \widehat{v}_{i\ell jt}^{(-f(i))} \right\} \right]. \tag{38}$$

Subsequent standardization absorbs the common Euler–Mascheroni constant. The index covers both movers and stayers in 2013–2019 and contains 18,858,572 worker-year observations.

Predictive performance is measured separately using 12,000 transitions sampled from the supported 2019 validation observations. Neither those arrival-year transitions nor their outcomes enter the training network used for this evaluation. Table A12 reports the ranking results alongside the Veneto comparison specification.

Table A12: Destination-Choice Prediction in Veneto and Portugal

	Recall@10	Recall@50	Median rank	Mean employer pool
Veneto comparison	0.238	0.407	105	47,600
Portugal	0.237	0.415	100	44,000

Notes: Recall@K is the share of held-out transitions whose realized destination is among the K highest-ranked employers. Portugal uses 12,000 transitions arriving in 2019, drawn from 213,545 supported validation transitions; Veneto uses 11,886 transitions arriving in 2000. Ranking uses the full empirical destination pools, excluding the incumbent. Employer-pool sizes are approximate, retaining the precision of the original comparison. Own earnings are excluded from both prediction specifications; the Veneto comparison retains its other country-specific inputs, including employer wage effects and clustered worker types. The country-specific menus and inputs differ, as detailed in Appendix I.2–I.4 and Section 5.

I.5 Firm Accounts and Rent-Sharing Specifications

Accounts linkage and sample. I link QP employers to SCIE using the common firm identifier recorded in the QP firm file, constructing the bridge by firm and year. Value added is measured at factor cost. The employment denominator is the SCIE number of paid employees; when that measure is zero or unavailable, I use the number of employees in the QP wage-bill file, followed by QP firm employment. Capital is tangible fixed assets divided by the same employment measure. Log value added and log capital per worker require positive values.

Following the rent-sharing approach of Card, Devicienti, and Maida (2014), I restrict the worker sample to firms with at least fifteen employees, require at least one year of tenure, and retain full-time workers. I exclude construction: property development and building construction (industry division 41), civil engineering (42), and specialized construction activities (43). I also exclude senior and intermediate staff (qualification grades 1 and 2), trainees and apprentices (grade 8), and observations with unreported qualification. The linked panel contains 13,773,956 worker-years, of which 6,308,982 satisfy these sample restrictions. Requiring the wage, accounts, instrument, outside-option index, firm age, tenure, and other regression covariates leaves 6,007,659 observations for the specification with year effects and controls. After fixed-effect sample selection, columns (2)–(4) contain 5,643,155; 5,641,338; and 5,464,130 observations, respectively.

For rent sharing, the outcome is log total monthly earnings from the linked pay panel, without the winsorization used in the wage-comparison table. The outside-option index is standardized within origin year over all finite scores, before imposing the accounts linkage and rent-sharing sample restrictions. This normalization differs from the pooled normalization in Table 7; the index need not have mean zero or unit standard deviation in the final rent-sharing sample.

Regression and instruments. Table 8 reports OLS and 2SLS estimates of equation (23) in Section 10, using the same observations for both estimators in each column. In that equation, $V_{\ell t} = \log(\mathrm{VA}_{\ell t}/N_{\ell t})$, $K_{\ell t}$ denotes tangible capital per worker, and I_{it} is the inclusive value, standardized within year for these regressions. The centering mean $\bar{V}$ is calculated over the common sample formed before fitting the separate specifications. Controls include age and age squared, tenure and tenure squared, an indicator for managerial, professional, technical, clerical, and service or sales occupations (major occupation groups 1–5), firm age and its square, and log employment and its square. The fixed effects $\mathcal{F}_{i\ell t}$ are, in columns (1)–(4), year; worker and year; worker, employer, and year; and match and year. Controls and capital are included in all four columns, where not absorbed.

I instrument value added per worker with revenue per worker at firms in the same four-digit industry but outside the incumbent's NUTS II region. To construct the instrument, I retain firm-years with accounts, positive employment and turnover, and observed industry and region. Revenue per employee is winsorized at the 0.5th and 99.5th percentiles separately by year. Writing $\mathcal{H}_{\ell t}$ for these outside-region firms in the same four-digit industry, the instrument is

$$Z_{\ell t} = \log\left[\frac{1}{|\mathcal{H}_{\ell t}|}\sum_{g\in\mathcal{H}_{\ell t}}\left(\frac{\widetilde{\mathrm{Revenue}}_{gt}}{N_{gt}}\right)\right], \qquad |\mathcal{H}_{\ell t}| \geq 10. \tag{39}$$

The tilde denotes winsorization. This is the log of an unweighted arithmetic mean of revenue per worker, not an average of its logarithm. The donor firms are selected on accounts availability and the conditions above; they need not satisfy the restrictions imposed on the worker regression sample.

There are two endogenous regressors, $V_{\ell t} - \bar{V}$ and $(V_{\ell t} - \bar{V})I_{it}$. Their excluded instruments are $Z_{\ell t} - \bar{Z}$ and $(Z_{\ell t} - \bar{Z})I_{it}$, respectively; both enter both first-stage equations along with the included covariates and fixed effects. The outside-option index itself is an included regressor. Identification therefore requires the external industry-revenue shifter, and its interaction with outside options, to affect wages through incumbent value added, conditional on these covariates and fixed effects.

The implied wage response to log value added per worker is $\beta_1 + \beta_3 I_{it}$. The interaction coefficient β_3 measures how this pass-through changes with outside options. Standard errors, clustered by four-digit industry, are in parentheses. Table A13 reports the first-stage estimates.

Table A13: Portugal: First Stages for Rent Sharing

	(1)	(2)	(3)	(4)
Panel A: Log VA/worker, $V_{\ell t} - \bar{V}$				
$Z_{\ell t} - \bar{Z}$	0.2827***	0.1659***	0.1226***	0.1228***
	(0.0490)	(0.0173)	(0.0294)	(0.0296)
$(Z_{\ell t} - \bar{Z}) \times I_{it}$	0.0735**	0.0076	-0.0018	-0.0018
	(0.0341)	(0.0075)	(0.0061)	(0.0076)
Joint excluded-instrument F	18.59	46.54	8.70	8.59
Panel B: Log VA/worker $\times I_{it}$, $(V_{\ell t} - \bar{V})I_{it}$				
$Z_{\ell t} - \bar{Z}$	0.1048***	0.0816***	0.0958**	0.0892**
	(0.0361)	(0.0263)	(0.0400)	(0.0438)
$(Z_{\ell t} - \bar{Z}) \times I_{it}$	0.4895***	0.4599***	0.4659***	0.4507***
	(0.0909)	(0.0859)	(0.0980)	(0.1120)
Joint excluded-instrument F	16.60	15.80	14.07	10.82
Worker and employer covariates	Yes	Yes	Yes	Yes
Worker FE	No	Yes	Yes	No
Employer FE	No	No	Yes	No
Worker–employer match FE	No	No	No	Yes
Year FE	Yes	Yes	Yes	Yes
Industry clusters	424	423	422	420
Observations	6,007,659	5,643,155	5,641,338	5,464,130

Notes: First-stage estimates for the 2SLS specifications in Panel B of Table 8. The dependent variable is centered log value added per worker in Panel A and its interaction with the outside-option index I_{it} in Panel B. The excluded instruments are centered $Z_{\ell t}$, defined in equation (39), and its interaction with I_{it}. Both instruments enter both equations. All columns include the outside-option index, log tangible capital per employee, and the worker and employer controls used in the outcome equation, where not absorbed by fixed effects. The index is standardized within year before sample selection; value added and the revenue instrument are centered at their common-sample means. Standard errors in parentheses are clustered by four-digit industry. The F-statistic is the clustered joint Wald test of the two excluded instruments in the corresponding first stage. Stars use two-sided cluster-t p-values: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.