Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach

Charly Porcher* Georgetown University

Eduardo Morales* Princeton University

Thomas Fujiwara* Princeton University

April 2024

Abstract

We investigate the role of information frictions in migration. We develop novel moment inequalities to estimate worker preferences while allowing for unobserved worker-specific information sets, migration costs, and location-specific amenities and prices. Using data on internal migration in Brazil, we find that common estimation procedures underestimate the importance of expected wages in migration choices, and that workers face substantial and heterogeneous information frictions. Model specification tests indicate that workers living in regions with higher internet access and larger populations have more precise wage information, and that information precision decreases with distance. Our estimated model predicts that information frictions play a quantitatively important role in reducing migration flows and worker welfare, and limit the welfare gains from reductions in migration costs.

JEL Classifications: C13, J61, R23

Keywords: migration, moment inequalities, information frictions

*We thank Treb Allen, Stephane Bonhomme, Lorenzo Casaburi, Rafael Dix-Carneiro, Dave Donaldson, Cécile Gaubert, Gabriel Kreindler, Melanie Morten, Diego Puga, and James Tybout for insightful conversations, and seminar participants at Aarhus, Carlos III, CEMFI, Dartmouth, Johns Hopkins SAIS, MIT, Penn State, Princeton, Stanford, University of Toronto, UPF, Yale, University of Virginia, and University of Zurich, for helpful comments. We especially thank Mark Bamba, Juanma Castro-Vincenzi, Pedro Cabra-Acela, and Anais Galdin for providing excellent research assistance. All errors are our own. Emails: charly.porcher@georgetown.edu, ecmorale@princeton.edu, and fujiwara@princeton.edu.

1 Introduction

Migration is among the most effective ways for people to improve their economic conditions. Yet, even within countries, migration rates are low compared to the spatial dispersion in wages (Jia et al., 2023). Both information frictions and migration costs can depress migration, but research to date has been unable to disentangle their roles because migrants' information is rarely observed. As a result, most studies place strong assumptions on workers' information and focus on estimating migration costs, sometimes interpreting them as accounting for information frictions. Estimates of migration costs tend to be high, suggesting that reducing physical barriers is key to facilitating migration. However, information frictions affect migration differently than migration costs. Costs affect how beneficial a move is, while information frictions affect how people choose a location and can lead to mistakes. Hence, information frictions may limit the benefits of reducing migration costs, and increasing access to information may be an important policy lever to improve workers' location choices.

We introduce a new method to separately identify the role of information frictions and migration costs in workers' location choices. Our method is applicable even when the researcher does not observe workers' information. We use it to address several questions. What do workers know about wages in different locations? How does allowing for incomplete information affect estimates of how workers value economic opportunities elsewhere, and of the distribution of migration costs? How would workers' location choices change if their information changed? How does workers' information mediate the impact of changes in migration costs?

Our answers rely on a methodological contribution that allows both information frictions and migration costs to vary flexibly between individuals, locations, and over time. We extend the treatment of information sets in the moment inequality literature (Pakes, 2010; Pakes et al., 2015) to multinomial logit models, static or dynamic, with arbitrarily large choice sets and choice-specific fixed effects. We then apply our moment inequality estimator to administrative data from Brazil, obtaining four main results. First, workers only have coarse information about wages in most local markets. However, workers in areas with higher internet access and larger populations have better information. Second, our estimates of the migration elasticity to expected wages are three times larger than those obtained using common estimation procedures, whereas our migration cost estimates are, on average, 21% lower. Third, counterfactuals from our estimated model indicate workers' migration rates and welfare increase as their wage information becomes more precise. Fourth, the welfare impacts of changes in migration costs (e.g., due to improvements in transportation infrastructure) are biased upwards when the researcher assumes workers' information is better than it truly is.

Our baseline analysis relies on a static migration model that incorporates workers' expected real wages, migration costs, amenities, and worker-by-location specific idiosyncratic preferences as the drivers of their location choices. Workers differ in observable types that impact their wages in each location. Types may be exogenous (e.g., determined by demographics) or chosen by the worker simultaneously with their location (e.g., sector or occupation). Within each type, workers are heterogeneous in both their expectations of location-specific wages and their idiosyncratic preferences for locations. We assume these preferences follow a type I extreme value distribution but impose no restriction on the distribution of workers' wage expectations, beyond assuming these are rational.[1] While our baseline analysis focuses on a static model, we show how to extend our moment inequality estimator to a dynamic setting with one-time migration costs and forward-looking workers.

We model migration costs, amenities, and prices as terms unobserved by the researcher that vary flexibly by origin, destination, and period, but not by worker type. They thus can be represented by choice-specific fixed effects that vary over time and across workers depending on their prior location. This approach, in combination with the large set of labor markets workers may choose from, implies a large number of parameters to estimate in most applications. Specifically, our empirical analysis features 50 locations, leading to 2,500 parameters per year to account for migration costs alone. Estimating high-dimensional parameter vectors using moment inequalities is computationally challenging when using standard inference procedures (Canay et al., 2023). Our key methodological contribution is to show how to calculate confidence intervals on each parameter in multinomial logit models featuring large choice sets, choice-specific fixed effects, and unrestricted variation in information sets.

We derive a new type of moment inequality that we call *bounding* inequality. By comparing workers' expected utility in two locations, we derive a moment *equality* that depends on a concave function of the worker's expected utility difference between both locations. We derive an *inequality* from this moment by bounding this concave function from above by its tangent at any chosen point. This inequality is *linear* in the worker's expected utility difference between the two locations, offering two advantages to our estimation procedure. First, it allows to substitute workers' unobserved wage expectations with their observed ex-post wage realizations. While this introduces workers' expectational errors in the moment function, under the rational expectations assumption, they have mean zero conditional on any variable in the worker's information set, leaving the inequality unaffected. Second, it provides a two-step estimation procedure that is feasible even in the presence of many choice-specific fixed effects.[2]

[1] The model with endogenous types is a nested logit model, with each type corresponding to a nest. Thus, the cross-location distribution of idiosyncratic preferences within a type follows a multinomial logit model.

[2] An existing literature has derived linear moment inequalities in the context of models with limited

The first step in our estimation procedure provides bounds for the parameter determining how responsive migration is to expected wage differences. We derive a moment inequality that compares the expected utility in two destinations for two workers with the same origin but distinct wage types, in the spirit of Ho and Pakes (2014). This comparison is, thus, specific to that origin, period, and two destinations, allowing to difference out migration costs, amenities, and price levels. The resulting inequality depends only on idiosyncratic preferences and the difference in expected wages between the two destinations for each type, multiplied by a preference parameter. We prove this moment inequality partially identifies the true parameter value as long as it conditions only on variables that belong to workers' information set, and we characterize the conditions under which point identification is achieved.

In the second step, we use the first-step estimates to derive moment inequalities that bound the choice-specific fixed effects. We combine our bounding inequalities with odds-based moment inequalities of the type initially introduced in Dickstein and Morales (2018) and recently applied in Bombardini et al. (2023) and Dickstein et al. (2023). Our inequalities yield bounds on the choice-specific fixed effects *one at a time*. Thus, instead of aiming to estimate a joint confidence set for all choice-specific fixed effects, which is infeasible in settings with many choices, we estimate one-dimensional confidence intervals for each fixed effect.

Intuitively, the second step uses moments expressed for a chosen origin and period that compare the worker's expected utility in two destinations. We thus infer the destinations' relative attractiveness from the number of workers who choose them, conditional on the expected wage difference. Intuitively, a migration choice that many workers make despite low expected wage gains must be due to the destination offering higher amenities or lower prices. Conversely, a migration choice expected to offer high wage gains but that is rarely observed implies the destination has lower amenities or higher prices, or that migration costs specific to that origin-destination are high. We prove our estimator generally partially identifies the location-specific fixed effects and characterize the stricter conditions under which it point identifies them.

We use a simulation exercise to illustrate our estimators' performance relative to other typical estimators. Consistent with our theoretical results, our moment inequalities successfully bound the true parameter values as long as the researcher correctly identifies a variable that belongs to the agent's information set. Furthermore, in some settings, our moment inequalities point identify the parameters. Conversely, maximum likelihood estimators are biased unless the researcher perfectly observes every agent's entire information set. When the maximum likelihood estimates are biased, these are often outside of the bounds of the

heterogeneity in idiosyncratic preferences (Pakes, 2010; Morales et al., 2019) or that assume their distribution is bounded (Eizenberg, 2014). We do so in the context of a multinomial logit model.

identified sets defined by our moment inequalities.

The second part of the paper uses our moment inequality estimator to study internal migration across labor markets in Brazil. We use linked employer-employee data from the *Relação Anual de Informações Sociais* (RAIS) and focus on a sample of workers with a high school degree identified as male and white. We define local labor markets as sector-region pairs and study the 1,000 markets formed by the largest 50 regions and 20 sectors. We identify worker types as their sector of employment and build a proxy for sector-by-region specific wages by estimating a wage function that depends flexibly on observed worker characteristics, an unobserved individual ability in each sector (individual-by-sector specific fixed effects), and a sector-by-location-by-period specific unobserved aggregate labor market shifter. This wage function approximates realized wages well, with a median R^2 of 0.83 across sectors. Under our assumption that worker characteristics are portable across locations, the migration decision only requires predicting the aggregate labor market shifter.[3]

Our empirical analysis provides three main conclusions. First, workers face substantial and heterogeneous information frictions. We do not reject that workers know the wage quartile to which each labor market belonged the previous year, but reject that workers *in general* have more precise information than that. That is, we cannot reject that workers can determine whether the previous year's wages in a labor market are in the top 25% of markets by its aggregate labor market-specific shifter, in the 50-75% bracket, in the 25-50% bracket, or in the bottom 25%, but we can reject that they can classify markets according to finer partitions. However, workers located in regions with a higher level of internet access or larger populations have more precise information overall, and workers in general have more accurate information about wages in nearby areas.

Second, the assumptions on workers' information imposed in the prior literature yield downward-biased estimates of the migration wage elasticity, and upward-biased migration cost estimates. Our moment inequality confidence interval for the elasticity of migration to expected wages is centered at 1.52, and does not include the Poisson Pseudo-Maximum Likelihood (PPML) estimate of 0.5. In addition, our migration cost estimates (measured in utility terms) are centered around values 21% lower than the PPML estimates.[4]

Third, the information frictions we infer also affect the partial-equilibrium welfare gains

[3]While our wage equation allows for worker-specific comparative advantage across sectors (as in Dix-Carneiro, 2014), it does not include an unobserved (to the researcher) location-specific individual effect that may be known by workers prior to determining their migration decisions. Kennan and Walker (2011) evaluate the importance of such a term and conclude that its impact on migration decisions in their setting is negligible.

[4]We implement the PPML estimator in Artuç and McLaren (2015), which yields consistent estimates of worker preference parameters if their expectations are rational and information sets are common to all workers. The difference between the PPML estimates and our moment inequality estimates is thus consistent with our finding that information sets are heterogeneous by worker locations.

from counterfactual policies that reduce migration costs.[5] We show this by simulating our model calibrated to the 1,000 labor markets in our sample, making the additional assumption that wages follow an AR(1) process with parameters we estimate. A 10% decline in migration costs increases welfare by about 5% if workers are fully informed about current wages, but only by 2.5% if they only observe the quartiles of the previous year's wages. Finally, we estimate substantial gains from improving workers' information. Welfare increases by about 3% when the precision of workers' information improves from only discerning quartiles of lagged wages to having perfect information on current wages.

Our paper is related to three strands of the literature. First, it contributes to work studying workers' mobility decisions in the face of large differences in income levels across locations. Our baseline static model incorporates worker location-specific idiosyncratic preferences and fixed migration costs, as in Tombe and Zhu (2019) and Morten and Oliveira (2024). In the dynamic extension to our model, we further allow for forward-looking workers and one-time migration costs, as in Kennan and Walker (2011).[6] Our contribution is to show how to estimate static and dynamic migration models without fully specifying workers' information, and to illustrate the impact on model estimates and counterfactual predictions of misspecifying workers' information sets.

Second, we contribute to the literature on the relevance of information frictions for migration. Recent work uses either randomized or natural experiments to evaluate the impact of migrants' information sets on their location decisions; e.g., Bryan et al. (2014); Bergman et al. (2020, 2023); Wilson (2021); Baseler (2023). In the absence of exogenous variation in information sets, other studies follow a structural approach. Kaplan and Schulhofer-Wohl (2017) introduce a model in which workers acquire information on location characteristics through a Bayesian learning process with specific parametric assumptions on the distribution of priors and signals. Porcher (2022) extends this approach by endogenizing the information acquisition process of workers who are rationally inattentive. Our contribution is to infer the importance of information frictions for migration choices while neither observing exogenous determinants of agents' information sets nor imposing any parametric restriction on the stochastic process determining these information sets.

Third, our paper contributes to studies using choice data to identify agents' preferences when their expectations of choice characteristics are rational but unobserved. In the absence of direct measures of agents' subjective expectations (Manski, 2004), it is common to assume that the researcher observes agents' entire information sets and can thus construct a perfect proxy of their expectations (Manski, 1991). An alternative approach estimates discrete-

[5] Examples of such policies include infrastructure improvements as in Morten and Oliveira (2024).

[6] Our analysis is partial equilibrium. For general equilibrium models of labor mobility, see, e.g., Artuç and McLaren (2015); Desmet et al. (2018); Caliendo et al. (2019, 2021).

choice models under the assumption that agents' information sets are identical across agents of the same observable type, while allowing the content of these information sets to be partly unobserved to the researcher (Traiberman, 2019). We allow information sets to vary across agents in unobservable ways, and use moment inequalities to partially identify agents' preferences in a multinomial logit model, static or dynamic, with arbitrarily large choice sets and choice-specific fixed effects. To deal with the large number of parameters to estimate, we introduce a novel way of deriving moment inequalities, the bounding inequalities, and show how to combine them with odds-based inequalities (Dickstein and Morales, 2018; Bombardini et al., 2023; Dickstein et al., 2023) to provide bounds that are informative and simple to compute.

The paper is organized as follows. Section 2 develops a static model of workers' location choices with incomplete information. Section 3 describes our two-step moment inequality estimator, and Section 4 illustrates its properties on simulated data. Section 5 discusses our empirical application. Section 6 shows how to extend our estimator to a dynamic model.

2 Model of Migration with Incomplete Information

We model the static location choice of workers in a population of interest. Workers are classified into S types indexed by s or r, and are indexed by i or j within each type.[7] While the model in this section assumes the worker's type is exogenous, we introduce in Appendix G a model in which their type is a choice the worker determines jointly with their location.

Define a variable y^l_{is} that equals one if worker i of type s chooses location l (and zero otherwise). We assume that

$$y^l_{is} \equiv \mathbb{1}\{l = \underset{l'=1,\ldots,L}{\operatorname{argmax}}\, \mathbb{E}[\mathcal{U}^{l'}_{is}|\mathcal{J}_{is}]\} \qquad \text{for } l = 1, \ldots, L, \tag{1}$$

where $\mathbb{1}\{A\}$ is an indicator function that equals 1 if A is true, $\mathcal{U}^l_{is} \in \mathbb{R}$ denotes the worker's utility of choosing l, $\mathcal{J}_{is}$ is their information set, and $\mathbb{E}[\cdot|\mathcal{J}_{is}]$ is a conditional expectation operator reflecting their beliefs. We impose the following assumptions on the worker's expected utility of choosing any $l = 1, \ldots, L$.

First, we assume the worker's expectations are rational. That is, for any vector $\mathcal{X}_{is}$, denoting $F(\cdot|\mathcal{J}_{is})$ as the distribution of $\mathcal{X}_{is}$ conditional on $\mathcal{J}_{is}$, it holds that

$$\mathbb{E}[\mathcal{X}_{is}|\mathcal{J}_{is}] = \int_x x dF(x|\mathcal{J}_{is}). \tag{2}$$

[7] In our empirical application (see Section 5), we focus on populations of interest defined by the worker's demographic characteristics and prior location, and identify the worker's type with the sector of employment.

Second, the utility of choosing location l for worker i of type s is:

$$\mathcal{U}^l_{is} = u^l_{is} + \varepsilon^l_{is}, \tag{3a}$$
$$u^l_{is} = \kappa^l + \alpha w^l_{is}, \tag{3b}$$

where w^l_{is} is the natural logarithm of the nominal wage that worker i would earn if they chose location l, and α captures the relative importance of wages in workers' utility. The terms κ^l and ε^l_{is} are the common and idiosyncratic components of all other determinants of a worker's location choice.[8] For simplicity, we refer to κ^l and ε^l_{is} as location l's amenity level and idiosyncratic preferences, respectively.

Third, defining the vector of idiosyncratic preferences $\varepsilon_{is} = (\varepsilon^1_{is}, \ldots, \varepsilon^L_{is})$ and the vector of amenity levels $\kappa = (\kappa^1, \ldots, \kappa^L)$, we assume that, for worker i of type s, it holds that

$$(\varepsilon_{is}, \alpha, \kappa) \subseteq \mathcal{J}_{is}, \tag{4}$$

where, for random vectors $\mathcal{X}$ and $\mathcal{X}'$, $\mathcal{X} \subseteq \mathcal{X}'$ indicates that the distribution of $\mathcal{X}$ conditional on $\mathcal{X}'$ is degenerate. Equation (4) imposes that when making their location choice, worker i knows their idiosyncratic preferences ε_{is}, the wage sensitivity α, and the amenity vector κ. It does not restrict which other variables may also belong to the vector $\mathcal{J}_{is}$.

Fourth, for any types s and r, locations l and l', and worker indices i and j, it holds that

$$\mathbb{E}[\Delta w^{ll'}_{is} | \mathcal{J}_{is}, \mathcal{J}_{jr}] = \mathbb{E}[\Delta w^{ll'}_{is} | \mathcal{J}_{is}] = \mathbb{E}[\Delta w^{ll'}_{is} | \mathcal{W}_{is}] = \mathbb{E}[\Delta w^{ll'}_{s} | \mathcal{W}_{is}], \tag{5}$$

where $\Delta w^{ll'}_{is} \equiv w^l_{is} - w^{l'}_{is}$ and $\mathcal{W}_{is}$ is a vector that includes all elements of $\mathcal{J}_{is}$ other than idiosyncratic preferences ε_{is}. The first equality in equation (5) imposes that a worker of a type s has at least as much information as any worker of a different type r about the differences in the wages they would earn in other locations. The second equality imposes that, once we condition on all other elements of the worker's information set, idiosyncratic preferences do not further help the worker forecast wages. Finally, the third equality imposes that the worker's expected wage difference between locations l and l' only depends on the expected difference between two terms that are common to all workers of type s.

The first equality in equation (5) naturally holds if we assume all workers in the population of interest have the same information; i.e., if $\mathcal{J}_{is} = \mathcal{J}_{jr}$ for any i, j, s, and r. While assuming homogeneous information sets is common in location-choice models (e.g., Diamond et al.,

[8] By defining the population of interest by workers' prior location, our framework allows κ^l to vary across workers with different origin locations. Thus, κ^l may account for several unobserved determinants of a worker's location choice, such as migration costs and (log) prices.

2019), this equality also holds if workers of a given type s know more about their type-specific wage differences than any worker of a different type r. This maintains the possibility for unobservable variation in the information sets of workers of different types and leaves the variation in information sets between workers of the same type completely unrestricted.

The second equality in equation (5) is an implication of the exogeneity of individual-specific shocks often assumed in migration models. These models often assume—besides perfect information on wages—that wages and preference shocks are independent, which implies the second equality. Importantly, our model allows for any pattern of correlation between location-specific amenities and wages, as we do not restrict the correlation between κ and the wage vector $w_{is} = (w_{is}^1, \ldots, w_{is}^L)$.

The third equality in equation (5) is imposed by data limitations. Our estimation approach allows us to model flexibly the information workers have about payoff-relevant variables that we either observe or for which we have consistent estimates. Our data allows us to estimate several individual-specific factors affecting wages (such as demographics or experience in a sector) that are common across locations and, therefore, do not enter the wage difference $\Delta w_{is}^{ll'}$ between any two locations l and l'. It does not, however, allow us to estimate, for every worker and location, a wage component that is location- and worker-specific. Hence, we impose that workers' wages may contain idiosyncratic location components, but workers do not make location choices based on those.[9] While equation (5) limits the worker's information about their own idiosyncratic location-specific wage components, it does not restrict their information about their type-specific comparative advantage. Thus, when types are equated to sectors or occupations, it is consistent with the findings in Dix-Carneiro (2014) and Traiberman (2019), which show that workers choosing their sector or occupation do so based on worker-specific wage components.

Fifth, we assume ε_{is} is *iid* across workers and, denoting as $F_\varepsilon(\cdot)$ the cumulative distribution function of ε_{is}, we assume that, for any workers i and j of types s and r, respectively, it holds

$$F_\varepsilon(\varepsilon_{is}|\mathcal{W}_{is}, \mathcal{W}_{jr}) = F_\varepsilon(\varepsilon_{is}) = \exp\Big(-\sum_{l=1}^{L} \exp(-\varepsilon_{is}^l)\Big). \tag{6}$$

The first equality in equation (6) imposes that worker i's idiosyncratic preferences are independent of all other elements of their, and of worker j's, information sets. The second equality imposes that ε_{is}^l is *iid* across all $l = 1, \ldots, L$, and follows a type I extreme value distribution with location parameter equal to zero and scale parameter equal to one.

[9] In our empirical application, 83% of the variation in individual wages across locations can be explained by the contribution of sector-location effects and individual-sector effects, therefore limiting how much residual variation at the individual-sector level could effectively influence migration choices.

Equations (1) to (6) are the only assumptions we impose on agents' behavior and information sets. Hence, we not only let workers' wage information vary flexibly across workers and locations in unobservable ways, but we also leave the data-generating process of wages completely unrestricted. In contrast, most studies of location choice estimating preference parameters assume wages are known to agents (Redding and Rossi-Hansberg, 2017) or, alternatively, specify both a stochastic process for wages and workers' information sets (Kennan and Walker, 2011), or assume these information sets arise from a particular learning process (Kaplan and Schulhofer-Wohl, 2017; Porcher, 2022). As workers' information is likely heterogeneous, unobserved to researchers and difficult to model correctly, we view the absence of restrictions on the wage process and information sets as an important advantage.

We consider a setting in which the researcher observes a random sample of workers of size I_s of each type s. For the i-th sampled worker of type s, the researcher observes the location choice, $y_{is} = (y_{is}^1, \dots, y_{is}^L)$. Additionally, for every type s, the researcher observes the vector of wage components $w_s = (w_s^1, \dots, w_s^L)$ and a vector of covariates $z_s = (z_s^1, \dots, z_s^L)$, with z_s^l a vector that may be used to predict w_s^l. In practice, w_s may not be directly observed, but estimated; see Appendix E.1 for details. Crucially, we do not assume the researcher observes $\mathcal{W}_{is}$ for any worker i of any sector s.

Only pairwise differences between the elements of κ are identified. Thus, without loss of generality, we impose the normalization $\kappa^1 = 0$. Given this normalization, the goal of estimation is to recover $(\kappa^2, \dots, \kappa^L)$ and α, and to learn about the content of workers' information sets. To acquire knowledge about workers' information, we test the null hypothesis that, for a subset of locations $\mathcal{L}$, a particular wage predictor z_s^l belongs to the information set for every worker of type s in a group of interest. We denote by $\theta \equiv (\theta_\alpha, \theta_2, \dots, \theta_L)$ the unknown parameter vector whose true value is $\theta^* \equiv (\alpha, \kappa^2, \dots, \kappa^L)$. We denote by Θ_α the set of possible values of θ_α, Θ_l the set of possible values of θ_l, $l = 2, \dots, L$, and Θ the parameter space; i.e., $\Theta = \Theta_\alpha \times \Theta_2 \times \cdots \times \Theta_L$.

3 Estimation Through Moment Inequalities

We describe here the estimation of θ. Given the large number L of choices in our setting, θ has high dimensionality. Commonly applied inference procedures for moment inequalities rely on computing a confidence set at each point in a grid covering the parameter space, limiting their applicability to settings with a small number of parameters. We propose a novel two-step procedure that circumvents these computational limitations and produces a confidence interval for each parameter individually.

In the first step, we compute a confidence interval for θ_α using inequalities that difference out the parameters in $(\theta_2, \ldots, \theta_L)$. In the second step, for each $l = 2, \ldots, L$, we derive inequalities that depend exclusively on the parameters θ_α and θ_l, and combine these inequalities with the first-step confidence interval for θ_α to compute a confidence interval for θ_l.

The moment inequalities we use in the first step combine the inequalities we use in the second step. Thus, for exposition purposes, we first describe the second-step inequalities in Section 3.1. We then describe how we build the first-step inequalities in Section 3.2. Section 3.3 explains how we use these inequalities to estimate confidence sets.

3.1 Second-Step Moment Inequalities

Given knowledge of the wage coefficient α, we use two types of inequalities to identify bounds on the amenity parameter κ^l for every $l = 2, \ldots, L$. In Section 3.1.1, we introduce a new type of moment inequalities that we name *bounding* inequalities. In Section 3.1.2, we describe how to apply odds-based inequalities (Dickstein and Morales, 2018; Bombardini et al., 2023; Dickstein et al., 2023) to our estimation problem.

3.1.1 Bounding Moment Inequalities

For any two locations l and l', we denote by $\Delta\theta_{ll'} \equiv \theta_l - \theta_{l'}$ the unknown parameter whose true value is $\Delta\kappa^{ll'} \equiv \kappa^l - \kappa^{l'}$, and as $\Theta_{ll'}$ the set of possible values of $\Delta\theta_{ll'}$. Then, for any worker i of type s, random vector z_s with support $\mathcal{Z}_s$, locations l and l', and deterministic function $h_{is}^{ll'} \colon \mathcal{Z}_s \times \Theta_{ll'} \to \mathbb{R}$, we define the moment

$$\begin{aligned} &\mathbb{m}_{is}^{ll'}(z_s, \Delta\theta_{ll'}, h_{is}^{ll'}(\cdot)) \equiv \\ &\mathbb{E}[y_{is}^{l'} - y_{is}^{l} \exp(-h_{is}^{ll'}(z_s, \Delta\theta_{ll'}))(1 + h_{is}^{ll'}(z_s, \Delta\theta_{ll'}) - (\Delta\theta_{ll'} + \alpha\Delta w_s^{ll'}))|z_s]. \end{aligned} \tag{7}$$

This moment is derived by comparing the expected utilities of choosing locations l and l' for a worker i of type s. The indices ll' reflect that the moment $\mathbb{m}_{is}^{ll'}(\cdot)$ varies across location pairs as it depends on the distribution of the wage difference $\Delta w_s^{ll'}$ conditional on z_s. Since the model in Section 2 does not specify the stochastic process for wages, this distribution may vary across location pairs l and l' for any type s.

Similarly, the indices is reveal that the moment $\mathbb{m}_{is}^{ll'}(\cdot)$ may vary across workers, even within a type, as it depends on the distribution of y_{is}^l and $y_{is}^{l'}$ conditional on z_s. By equation (1), these choice variables are functions of the information set of worker i of type s, $\mathcal{J}_{is}$. Since the list of variables in the worker's information set may be heterogeneous across workers, its distribution, and that of y_{is}^l and $y_{is}^{l'}$, conditional on z_s, may also vary across workers.

Theorem 1 establishes a key property of the moment $\mathbb{m}_{is}^{ll'}(\cdot)$ if evaluated at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$.

Theorem 1 *Assume equations (1) to (6) hold and $z_s \subset \mathcal{J}_{is}$. Then, $\mathbb{m}_{is}^{ll'}(z_s, \Delta\kappa^{ll'}, h_{is}^{ll'}(\cdot)) \geqslant 0$ for any locations l and l', worker i of type s, $z_s \in \mathcal{Z}_s$, and deterministic function $h_{is}^{ll'}\colon \mathcal{Z}_s \times \Theta_{ll'} \to \mathbb{R}$.*

We prove Theorem 1 in Appendix A.1. Theorem 1 states that, given equations (1) to (6) and a vector z_s that belongs to the information set of worker i of type s, the moment in equation (7) is positive when evaluated at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$. For any pair of locations l and l', we may then compute the set of values of $\Delta\theta_{ll'}$ for which

$$\mathbb{m}_{is}^{ll'}(z_s, \Delta\theta_{ll'}, h_{is}^{ll'}(\cdot)) \geqslant 0, \tag{8}$$

and if equations (1) to (6) hold and $z_s \subset \mathcal{J}_{is}$, $\Delta\kappa^{ll'}$ will belong to this set regardless of the value of z_s on which the moment conditions. Given locations l and l', a worker i of type s, and a value of z_s, the set of values of $\Delta\theta_{ll'}$ other than $\Delta\kappa^{ll'}$ that satisfy the inequality in equation (8) depend on the function $h_{is}^{ll'}(\cdot)$ entering the moment $m_{is}^{ll'}(\cdot)$. Appendix B.1 shows that, for any $z_s \in \mathcal{Z}_s$, the function $h_{is}^{ll'}(z_s, \Delta\theta_{ll'})$ that minimizes the set of values of $\Delta\theta_{ll'}$ that satisfy the inequality in equation (8) is:

$$h_{is}^{ll'}(z_s, \Delta\theta_{ll'}) = \Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'} | z_s, y_{is}^l = 1]. \tag{9}$$

The subindices is in this function reflect that, for any two locations l and l', the expectation of $\Delta w_s^{ll'}$ conditional on z_s and y_{is}^l may vary across workers, with this variation being due to potential cross-worker heterogeneity in the content of their information sets.

Appendix B.1 shows the inequality implied by equations (7) to (9) can be written as

$$\frac{\mathbb{E}[y_{is}^l | z_s]}{\mathbb{E}[y_{is}^{l'} | z_s]} \exp(-\alpha\mathbb{E}[\Delta w_s^{ll'} | z_s, y_{is}^l = 1]) \leqslant \exp(\Delta\theta_{ll'}). \tag{10}$$

By swapping the identity of locations l and l' in equations (7) to (9), we obtain the inequality

$$\frac{\mathbb{E}[y_{is}^l | z_s]}{\mathbb{E}[y_{is}^{l'} | z_s]} \exp(-\alpha\mathbb{E}[\Delta w_s^{ll'} | z_s, y_{is}^{l'} = 1]) \geqslant \exp(\Delta\theta_{ll'}). \tag{11}$$

Equations (10) and (11) identify lower and upper bounds on $\Delta\theta_{ll'}$, respectively. Intuitively, equation (10) reveals that if workers are more likely to choose l over l', as represented by a high ratio $\mathbb{E}[y_{is}^l | z_s]/\mathbb{E}[y_{is}^{l'} | z_s]$, even when they expect low wages in l relative to l', as represented by a high value of $-\alpha\mathbb{E}[\Delta w_s^{ll'} | z_s, y_{is}^l = 1]$, then it must be that the amenity difference between l and l', $\Delta\theta_{ll'}$, is sufficiently large. The intuition for equation (11) is analogous.

One may derive bounds similar to those in equations (10) and (11) for any $z_s \in \mathcal{Z}_s$.

The parameter $\Delta\theta_{ll'}$ will generally be partially identified. However, the following corollary describes a case where the lower and upper bounds for $\Delta\theta_{ll'}$ in equations (10) and (11) coincide, implying that this parameter is point identified.

Corollary 1 *Assume equations (1) to (6) hold, $z_s \subset \mathcal{J}_{is}$, and $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_{is}]$. Then, the bounds in equations (10) and (11) imply $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$.*

We prove Corollary 1 in Appendix B.2. Corollary 1 strengthens the assumptions in Theorem 1 by requiring that $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_{is}]$. This additional requirement implies the vector z_s contains all the information worker i has about the wage difference between locations l and l'. When this additional assumption holds, the moment inequality defined by equations (7) to (9) only holds when the unknown parameter $\Delta\theta_{ll'}$ equals its true value $\Delta\kappa^{ll'}$.

Theorem 1 holds for any two locations l and l'. Given the normalization $\kappa^1 = 0$, choosing $l' = 1$ yields $\Delta\theta_{l1} = \theta_l$, and equations (10) and (11) identify bounds on the amenity component θ_l, with point identification if the assumptions in Corollary 1 hold for $l' = 1$.

3.1.2 Odds-based Moment Inequalities

For any locations l and l', worker i of type s, and value $z_s \in \mathcal{Z}_s$, we define the moment

$$\mathbb{m}_{is,o}^{ll'}(z_s, \Delta\theta_{ll'}) \equiv \mathbb{E}[y_{is}^l \exp(-(\Delta\theta_{ll'} + \alpha\Delta w_s^{ll'})) - y_{is}^{l'}|z_s]. \tag{12}$$

The index o differentiates this odds-based moment from the bounding moment in equation (7). Theorem 2 establishes a key property of this moment when evaluated at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$.

Theorem 2 *Assume equations (1) to (6) hold and $z_s \subset \mathcal{J}_{is}$. Then, $\mathbb{m}_{is,o}^{ll'}(z_s, \Delta\kappa^{ll'}) \geqslant 0$ for any locations l and l', worker i of type s, and $z_s \in \mathcal{Z}_s$.*

We prove Theorem 2 in Appendix A.2. Theorem 2 states that, given equations (1) to (6) and a vector z_s that belongs to the information set of worker i of type s, the moment in equation (12) is positive when evaluated at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$. For any locations l and l', we may then compute the set of values of $\Delta\theta_{ll'}$ for which

$$\mathbb{m}_{is,o}^{ll'}(z_s, \Delta\theta_{ll'}) \geqslant 0, \tag{13}$$

and if equations (1) to (6) hold and $z_s \subset \mathcal{J}_{is}$, $\Delta\kappa^{ll'}$ will belong to this set for any of z_s. As we show in Appendix B.3, the inequality (13) can be written as

$$\frac{\mathbb{E}[y_{is}^l|z_s]}{\mathbb{E}[y_{is}^{l'}|z_s]}\mathbb{E}[\exp(-\alpha\Delta w_s^{ll'})|z_s, y_{is}^l = 1] \geqslant \exp(\Delta\theta_{ll'}). \tag{14}$$

By swapping the identity of locations l and l' in equation (14), we obtain the inequality

$$\frac{\mathbb{E}[y_{is}^{l}|z_s]}{\mathbb{E}[y_{is}^{l'}|z_s]}(\mathbb{E}[\exp(-\alpha\Delta w_s^{l'l})|z_s, y_{is}^{l'} = 1])^{-1} \leqslant \exp(\Delta\theta_{ll'}). \tag{15}$$

Equations (14) and (15) identify upper and lower bounds on $\Delta\theta_{ll'}$, respectively, with a similar intuition as the bounding inequality; see Section 3.1.1. These equations generally partially identify $\Delta\theta_{ll'}$. However, the following corollary describes a case where the lower and upper bounds for $\Delta\theta_{ll'}$ in equations (14) and (15) coincide, implying they point identify $\Delta\theta_{ll'}$.

Corollary 2 *Assume equations (1) to (6) hold, $z_s \subset \mathcal{J}_{is}$, and $\Delta w_s^{ll'} = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]$. Then, the bounds in equations (14) and (15) imply $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$.*

We prove Corollary 2 in Appendix B.4. Corollary 2 strengthens the assumptions in Theorem 2 by requiring that worker i has perfect information on $\Delta w_s^{ll'}$, $\Delta w_s^{ll'} = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]$. When this is satisfied, the inequality in equation (13) only holds if $\Delta\theta_{ll'}$ equals its true value $\Delta\kappa^{ll'}$.

Theorem 2 and Corollary 2 hold for any locations l and l'. Setting $l' = 1$, they indicate when the moment in equation (13) partially or point identifies location l's amenity term, θ_l.

3.2 First-Step Moment Inequalities

For any locations l and l', any workers i of type s and j of type r, any vectors $z_s \in \mathcal{Z}_s$ and $z_r \in \mathcal{Z}_r$, and any deterministic function $g_{ijsr}^{ll'}\colon \mathcal{Z}_s \times \mathcal{Z}_r \times \Theta_\alpha \to \mathbb{R}$, we define the moment

$$\mathbb{M}_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha, g_{ijsr}^{ll'}(\cdot)) \equiv \tag{16}$$
$$\mathbb{E}[y_{is}^{l}y_{jr}^{l} + y_{is}^{l'}y_{jr}^{l'} - y_{is}^{l}y_{jr}^{l'}\exp(-g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha))(2 + 2g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha) - \theta_\alpha(\Delta w_s^{ll'} + \Delta w_r^{l'l}))|z_s, z_r].$$

This moment is derived from a comparison of the expected utilities of choosing locations l and l' for workers i and j. For reasons analogous to those discussed when describing equation (7), the expectation in equation (16) may vary across pairs of workers is and jr, and across pairs of locations l and l'; hence the indices in the moment definition. Theorem 3 establishes a key property of this moment when evaluated at $\theta_\alpha = \alpha$.

Theorem 3 *Assume equations (1) to (6) hold, $z_s \subset \mathcal{J}_{is}$ for worker i of type s, and $z_r \subset \mathcal{J}_{jr}$ for worker j of type r. Then, $\mathbb{M}_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha, g_{ijsr}^{ll'}(\cdot)) \geqslant 0$ for any l and l', workers i of type s and j of type r, $z_s \in \mathcal{Z}_s$, $z_r \in \mathcal{Z}_r$, and deterministic function $g_{ijsr}^{ll'}\colon \mathcal{Z}_s \times \mathcal{Z}_r \times \Theta_\alpha \to \mathbb{R}$.*

We prove Theorem 3 in Appendix A.3. Theorem 3 states that, given equations (1) to (6), the assumption that z_s belongs to the information set of worker i of type s, and the assumption that z_r belongs to the information set of worker j of type r, the moment in equation (16) is positive when evaluated at $\theta_\alpha = \alpha$. We may then compute the set of values of θ_α that satisfy

$$\mathbb{M}^{ll'}_{ijsr}(z_s, z_r, \theta_\alpha, g^{ll'}_{ijsr}(\cdot)) \geqslant 0, \tag{17}$$

and if equations (1) to (6) hold, $z_s \subset \mathcal{J}_{is}$, and $z_r \subset \mathcal{J}_{jr}$, then α will belong to this set. By minimizing the moment in equation (16) with respect to $g^{ll'}_{ijsr}(\cdot)$, we compute the function that minimizes the set of values of θ_α that satisfy the inequality in equation (17). Appendix B.5 shows that, for any $z_s \in \mathcal{Z}_s$ and $z_r \in \mathcal{Z}_r$ this function is

$$g^{ll'}_{ijsr}(z_s, z_r, \theta_\alpha) = \theta_\alpha \mathbb{E}[0.5(\Delta w^{ll'}_s + \Delta w^{l'l}_r)|z_s, z_r, y^l_{is} y^{l'}_{jr} = 1]. \tag{18}$$

As shown in Appendix B.5, the inequality represented in equation (17) with the function in equation (18) can be written as

$$\frac{\mathbb{E}[y^l_{is} y^{l'}_{jr}|z_s, z_r]}{\mathbb{E}[0.5(y^l_{is} y^l_{jr} + y^{l'}_{is} y^{l'}_{jr})|z_s, z_r]} \leqslant \exp(\theta_\alpha \mathbb{E}[0.5(\Delta w^{ll'}_s + \Delta w^{l'l}_r)|z_s, z_r, y^l_{is} y^{l'}_{jr} = 1]). \tag{19}$$

This inequality yields lower and upper bounds on θ_α when its right-hand side is increasing or decreasing in θ_α, respectively. Intuitively, if worker i of type s is likely to choose location l whereas worker j of type r is likely to choose l', as represented by a high value of the ratio on the left-hand side of equation (19), and both workers expect the wage to be high in their chosen location relative to the alternative, as represented by a positive value of $\mathbb{E}[0.5(\Delta w^{ll'}_s + \Delta w^{l'l}_r)|z_s, z_r, y^l_{is} y^{l'}_{jr} = 1]$, then the wage coefficient θ_α cannot be too low. Conversely, if both workers is and jr are still likely to choose locations l and l', respectively, but now the workers expect the wage to be low in their chosen locations relative to the alternative, as represented by a negative value of $\mathbb{E}[0.5(\Delta w^{ll'}_s + \Delta w^{l'l}_r)|z_s, z_r, y^l_{is} y^{l'}_{jr} = 1]$, then θ_α cannot be too high. Hence, positive and negative wage differences between locations identify opposite bounds.

The parameter θ_α is generally partially identified. The following corollary describes a case where moment inequalities of the type in equation (19) point identify θ_α.

Corollary 3 *Assume equations (1) to (6) hold, $\Delta\kappa^{ll'} = 0$, and $\mathbb{E}[\Delta w^{ll'}_s|z_s] = \mathbb{E}[\Delta w^{ll'}_s|\mathcal{J}_{is}] = \mathbb{E}[\Delta w^{l'l}_r|z_r] = \mathbb{E}[\Delta w^{l'l}_r|\mathcal{J}_{ir}] = \Delta\bar{w}$, where $\Delta\bar{w}$ is a common constant. Then, the combination of two moment inequalities of the type in equation (19), one with $\Delta\bar{w} > 0$ and the other one with $\Delta\bar{w} < 0$, point identifies θ_α.*

We prove Corollary 3 in Appendix B.6. According to Corollary 3, the conditions that must be satisfied for inequalities of the type in equation (19) to point identify θ_α are even stronger than those required in Corollary 1 to point identify $\Delta\theta_{ll'}$ for locations l and l' given knowledge of α. Intuitively, the moment inequality in equation (16) results from adding two inequalities, one for worker i of type s that compares their expected utility of location l to that of location l', and one for worker j of type r that compares their expected utility of l' to that of l; see Appendix A.3 for details. Combining the two inequalities is useful because the resulting inequality does not depend on the amenity parameters κ^l and $\kappa^{l'}$, which cancel since they are the same for workers of types s and r. However, the combined inequality is naturally weaker than each of the two inequalities that are combined into it, and the conditions required for it to point identify the parameter of interest are stronger.

3.3 Using the Inequalities for Estimation

Given a confidence set for θ_α, we describe here how we use the moment inequalities introduced in Section 3.1 to compute a confidence set for the amenity term θ_l for a location l. The procedure we follow to compute a confidence set for θ_α using the inequalities in Section 3.2 is analogous; see Appendix B.7. To simplify the exposition, we describe our inference procedure in terms of the bounding inequalities in equation (8); in practice, we combine these with the odds-based inequalities in equation (13).

The moment inequality in equation (8) is conditional on a vector z_s. In our setting, z_s includes predictors of the wage difference $\Delta w_s^{ll'}$ between any two locations. When computing confidence sets, we use unconditional moments that multiply the corresponding moment function by an instrument vector. We denote by $\Delta x_s^{ll'}$ the instrument vector we use to build moment inequalities that compare the utility of workers of type s between locations l and l'.

We compute $\Delta x_s^{ll'}$ as follows. Given $q \in \mathbb{N}$ and a predictor $\Delta z_s^{ll'}$ of $\Delta w_s^{ll'}$, we compute a vector $(Q_0^q, \ldots, Q_q^q)$ of q-quantiles of the distribution of $\Delta z_s^{ll'}$ across all types and location pairs. The vector $\Delta x_s^{ll'}$ then groups observations with values of $\Delta z_s^{ll'}$ between the same two quantiles, weighting each observation by the absolute value of $\Delta z_s^{ll'}$ to the power of $d \in \mathbb{Z}$:

$$\Delta x_s^{ll'} = (\Delta x_{s,1}^{ll'}, \ldots, \Delta x_{s,q}^{ll'})' \text{ with } \Delta x_{s,k}^{ll'} \equiv \mathbb{1}\{Q_{k-1}^q < \Delta z_s^{ll'} \leqslant Q_k^q\}|\Delta z_s^{ll'}|^d \text{ for } k \in [1, q]. \quad (20)$$

If, for example, $q = 2$, $\Delta x_s^{ll'}$ partitions all combinations of location pairs and types (l, l', s) into two sets depending on whether the $\Delta z_s^{ll'}$ is above or below its median across all combinations.

The elements of $\Delta x_s^{ll'}$ in equation (20) are weakly positive and a function of z_s. Thus, Theorem 1 and the Law of Iterated Expectations (LIE) imply that, for any locations l and l', worker i of type s, and deterministic function $h_{is}^{ll'}\colon \mathcal{Z}_s \times \Theta_{ll'} \to \mathbb{R}$, the $q \times 1$ vector of unconditional moment inequalities

$$\mathbb{E}[\mathrm{m}_{is}^{ll'}(z_s, \Delta\theta_{ll'}, h_{is}^{ll'}(\cdot))\Delta x_s^{ll'}] \geqslant 0, \quad (21)$$

holds for $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$ if $z_s^{ll'} \subset \mathcal{J}_{is}$, where $\mathbb{m}_{is}^{ll'}(\cdot)$ is defined in equation (7).

The empirical setting described in Section 2 includes one observation per worker. The sample analogue of the moment inequality in equation (21) thus averages over only one observation. However, as this inequality is valid for every worker i of every type s, it holds that, for any locations l and l' and function $h_{is}^{ll'}\colon \mathcal{Z}_s \times \Theta_{ll'} \to \mathbb{R}$, the $q \times 1$ vector of inequalities

$$\sum_{s=1}^{S}\sum_{i=1}^{I_s} \mathbb{E}[\mathbb{m}_{is}^{ll'}(z_s, \Delta\theta_{ll'}, h_{is}^{ll'}(\cdot))\Delta x_s^{ll'}] \geqslant 0 \tag{22}$$

is satisfied at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$ if $z_s \subset \mathcal{J}_{is}$ for every worker $i = 1, \ldots, I_s$ of every type $s = 1, \ldots, S$.

Given the normalization of the amenity in location $l = 1$ (i.e., $\kappa^1 = 0$), the inequality in equation (22) can be used to compute a confidence set for θ_l in any location $l \neq 1$, by fixing the pair of locations (l, l') to both $(l, 1)$ and $(1, l)$. The inequalities in equation (22) with indices $(l, 1)$ identify an upper bound on θ_l; those with indices $(1, l)$ identify a lower bound.

To compute a 95% confidence interval for θ_l, we first compute 96% confidence intervals for θ_l conditional on each value of θ_α in a 99% confidence interval for this parameter. We denote these confidence intervals as $\hat{\Theta}^l_{.96}(\theta_\alpha)$. We compute them by applying the procedure in Andrews and Soares (2010) to the sample analogue of the inequalities in equation (22) for the appropriate choice of location indices; see Appendix B.8. We then compute the 95% confidence interval for θ_l as the union of $\hat{\Theta}^l_{.96}(\theta_\alpha)$ for each value of θ_α in its 99% confidence interval. As discussed in Bei (2024), the resulting confidence interval for θ_l is valid.

4 Properties of Moment Inequalities: Simulation

This section uses simulations to illustrate the properties of the two-step estimator described in Section 3. We study its performance as the precision of workers' and the researcher's wage information varies. Three main insights emerge. First, our estimator yields an interval that contains the true parameters even when both the agents and the researcher imperfectly observe the realized wages. Second, the Maximum Likelihood Estimator (MLE) is downward biased and generally not contained in our estimator's identified set. Third, our estimator often yields empty identified sets when the researcher assumes incorrect information sets, demonstrating its potential to test for the true content of agents' information sets.

4.1 Simulation Set-up

Workers choose between three locations $l = \{1, 2, 3\}$ according to the model in Section 2. We simulate data for $6{,}000{,}000$ workers, each of them of a different type. We index observations by s, and set the wage coefficient to $\alpha = 1$ and the location-specific amenities to $\kappa^1 = \kappa^2 = 0$ and $\kappa^3 = 1$.[10] The estimator described in Section 3 requires specifying neither a stochastic process for wages nor workers' information sets. However, specifying these aspects of the model is needed to generate the model-implied choice for all sampled individuals.

We assume the wage vector $w_s = \{w_s^l\}_{l=1}^3$ is *iid*, and each element w_s^l is determined as

$$w_s^l = z_{1s}^l + z_{2s}^l + z_{3s}^l, \tag{23}$$

with z_{ks}^l independent across $l = \{1, 2, 3\}$ and $k = \{1, 2, 3\}$, and distributed uniformly with support $[\mu_k^l - \sigma_k, \mu_k^l + \sigma_k]$. We set $\mu_1^l = \mu_3^l = 0$ for $l = \{1, 2, 3\}$ and $(\mu_2^1, \mu_2^2, \mu_2^3) = (0, -0.5, -1)$; thus, mean wages decline in order from $l = 1$ to $l = 3$. We set the dispersion of z_{2s}^l to $\sigma_2 = 4$ and study cases for different values of σ_1 and σ_3.

We assume workers observe $(z_{1s}, z_{2s}) = \{(z_{1s}^l, z_{2s}^l)\}_{l=1}^3$ but not $z_{3s} = \{z_{3s}^l\}_{l=1}^3$. Therefore, $\mathbb{E}[z_{1s}, z_{2s}|\mathcal{W}_s] = (z_{1s}, z_{2s})$ and $\mathbb{E}[z_{3s}|\mathcal{W}_s] = \mathbb{E}[z_{3s}] = 0$ for all s, and, as a result, it holds for any two locations l and l' that worker s's expected wage difference is

$$\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_s] = \Delta z_{1s}^{ll'} + \Delta z_{2s}^{ll'}. \tag{24}$$

where, for all k, $\Delta z_{ks}^{ll'} = z_{ks}^l - z_{ks}^{l'}$. The variable $\Delta z_{3s}^{ll'}$ thus equals the expectational error worker s makes when forecasting $\Delta w_s^{ll'}$. By changing σ_3, we then evaluate the impact workers' imperfect wage information has on the performance of the different estimators we consider.

We assume the researcher observes (y_s^l, w_s^l, z_{2s}^l) for $l = \{1, 2, 3\}$ and $s = 1, \ldots, S$. Thus, $z_{1s} = \{z_{1s}^l\}_{l=1}^3$ belongs to worker s's information set but is unobserved by the researcher. By changing σ_1, we then evaluate the impact that unobserved (by the researcher) components of a worker's information set have on different estimators.

4.2 Simulation Results

We report the main simulation results in Table 1. We describe in Section 4.2.1 the 95% confidence intervals for the wage coefficient θ_α, displayed in the column labeled *First Step*. We describe in Section 4.2.2 the 95% confidence intervals for the amenities θ_2 and θ_3, displayed in the columns labeled *Second Step*. We provide a detailed description of the moment inequalities we use in this simulation exercise in Appendix C.1.

[10] As expected wages vary by type according to equation (5), the large S limits the impact of noise in our simulation. The confidence sets we report are thus similar to the identified sets implied by our inequalities.

4.2.1 Confidence Intervals for the Wage Coefficient

We compute confidence intervals for θ_α using sample analogues of the inequalities introduced in Section 3.2. We build separate inequalities for each pair of locations l and l'. When building each inequality, we combine each observation s with an observation r selected to reproduce the conditions under which, according to Corollary 3, the resulting inequality would point identify θ_α if the restriction $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_{is}]$ were to hold for all $s = 1, \ldots, S$.

In cases 1 and 2, we set $\sigma_1 = 0$ and, thus, $z_{1s}^l = 0$ for every location l and worker s. That is, the researcher observes all wage predictors on which the agent bases their decision. Hence, $\mathbb{E}[\Delta w_s^{ll'}|z_{2s}] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_{is}]$ for every location pair and type s. Consistently with Corollary 3, the resulting 95% confidence interval for θ_α is very tight around its true value α. This is true even in case 2, where the agent makes errors when forecasting wages in every location, indicated by a nonzero value of σ_3.

Two key characteristics of our simulation setting ensure a tight confidence interval for θ_α in cases 1 and 2. First, Corollary 3 requires for point-identification of θ_α that the locations l and l' being compared have equal amenity levels; i.e., $\Delta\kappa^{ll'} = 0$. This holds for locations 1 and 2, since we set $\kappa^1 = \kappa^2 = 0$. We illustrate the importance of this condition in Table C.1 in Appendix C.2 by varying the value of the amenity terms κ^1, κ^2 and κ^3. The confidence interval for θ_α remains tight as long as $\Delta\kappa^{ll'} = 0$ for two locations l and l', and becomes wider as the differences in κ^l across all three locations $l = \{1, 2, 3\}$ increase.

Second, Corollary 3 also requires that the observations s and r combined in equation (16) verify $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_r^{l'l}|z_r]$. In our setting, wages and their predictors are continuous variables, so the predicted wages $\mathbb{E}[\Delta w_s^{ll'}|z_s]$ and $\mathbb{E}[\Delta w_r^{l'l}|z_r]$ never coincide exactly. This explains why the confidence interval for θ_α reported in Table 1 includes values other than α even when all other sufficient conditions in Corollary 3 are satisfied. As Table C.2 in Appendix C.3 reveals, the confidence interval for θ_α becomes wider as the average difference between $\mathbb{E}[\Delta w_s^{ll'}|z_s]$ and $\mathbb{E}[\Delta w_r^{l'l}|z_r]$ for the matched types s and r increases.

Cases 3 and 4 share the feature that $\sigma_1 > 0$ and, thus, the researcher only observes part of the agent's information set—the true information set is (z_{1s}, z_{2s}) for every type s, but the researcher only observes z_{2s}. By using information on only a subset of the agent's true information set, the researcher obtains a wider confidence interval for θ_α. In case 5, the researcher wrongly assumes the agent has perfect information on wages. In this case, the confidence interval for θ_α only includes one value ($\theta_\alpha = 0.87$), different from the true value $\alpha = 1$.[11] We show in Table C.5 in Appendix C.6 that $\theta_\alpha = 0.87$ can be ruled out as the true

[11] Molinari (2020) and Andrews and Kwon (2024) point out that moment inequality models, when they are misspecified, may yield confidence sets that are tight but do not include the true parameter value.

Table 1: Simulation Results - Moment Inequality Confidence Intervals

Case	σ_1	σ_3	z_s	First Step	Second Step		
				α	Mom. Ineq.	κ^2	κ^3
1	0	0	z_{2s}	[1 , 1.02]	Bounding	[0 , 0]	[1 , 1]
					Odds-based	[0 , 0]	[1 , 1]
					Both	[0 , 0]	[1 , 1]
2	0	1	z_{2s}	[1 , 1.01]	Bounding	[0 , 0]	[1 , 1]
					Odds-based	[-0.33 , 0.32]	[0.68 , 1.33]
					Both	[0 , 0]	[1 , 1]
3	1	0	z_{2s}	[0.82 , 1.29]	Bounding	[-0.31 , 0.31]	[0.70 , 1.30]
					Odds-based	[0 , 0]	[1 , 1.01]
					Both	[0 , 0]	[1 , 1.01]
4	1	1	z_{2s}	[0.82 , 1.31]	Bounding	[-0.31 , 0.31]	[0.69 , 1.31]
					Odds-based	[-0.38 , 0.39]	[0.68 , 1.45]
					Both	[-0.31 , 0.31]	[0.69 , 1.31]
5	0	1	w_s	[0.87 , 0.87]	Bounding	[-0.05 , -0.10]	[0.85 , 0.88]
					Odds-based	$\varnothing$	$\varnothing$
					Both	$\varnothing$	$\varnothing$

The true parameter values are $\alpha = 1$, $\kappa^2 = 0$, and $\kappa^3 = 1$. The column α contains a 95% confidence interval for θ_α based on the estimator introduced in Section 3.2 and described in Appendix C.1. The columns κ^2 and κ^3 contain 95% confidence intervals for θ_2 and θ_3 based on the estimators introduced in Section 3.1. The confidence intervals for θ_2 and θ_3 in the rows labeled *Bounding* use the inequalities introduced in Section 3.1.1; those in the row labeled *Odds-based* use the inequalities introduced in Section 3.1.2; and those in the row labeled *Both* combine both inequalities. Confidence sets are computed following the procedure in Andrews and Soares (2010) and using grids over $[0.5, 1.5]$ for α, $[-0.5, 0.5]$ for κ_2 and $[0.5, 1.5]$ for κ_3.

parameter value by increasing the number of instruments.[12]

4.2.2 Confidence Intervals for Amenities

Given a confidence interval for θ_α, we then compute confidence intervals for θ_2 and θ_3 one at a time, using the inequalities introduced in sections 3.1.1 and 3.1.2.

Consistently with Corollary 1, the bounding inequalities point identify θ_2 and θ_3 when the agent's information set is perfectly observed by the researcher; i.e., when $\sigma_1 = 0$. As σ_1 increases, the confidence intervals built using the bounding inequalities alone still contain the true amenities, in line with Theorem 1, but widen to include other parameter values.

In contrast, the odds-based inequalities perform best when workers make no expectational errors (i.e., when $\sigma_3 = 0$). If this holds, the odds-based inequalities point identify the amenities as long as the first-step confidence interval for the wage coefficient equals its true

[12] Increasing the number of quantiles q into which we divide the support of the predictor $\Delta z^{ll'}_{2s}$ from $q = 2$ to $q = 4$ is enough to yield an empty confidence interval in case 5.

value α, as predicted by Corollary 2, or the value of σ_1 is also equal to zero, this being true in case 1. When the confidence interval for θ_α includes values other than α and $\sigma_1 > 0$, the odds-based inequalities may still point identify amenities, as in case 3, but they will not do so in general.[13]

Since the bounding inequalities are insensitive to agents' expectational errors (i.e., insensitive to σ_3) and the odds-based inequalities are partially insensitive to agents having information the researcher does not observe (i.e., insensitive to σ_1), there are advantages from combining both types of inequalities in estimation. Cases 2 and 3 show that, when either $\sigma_1 = 0$ or $\sigma_3 = 0$, combining bounding and odds-based inequalities point identifies amenities, although neither of these two inequalities point identifies amenities in both cases when considered in isolation.[14]

Case 4 is likely the most empirically relevant. In this setting, the agent's information set is partly unobserved (i.e., $\sigma_1 > 0$), and the agent predicts wages with error (i.e., $\sigma_3 > 0$). Even in this scenario, our estimator yields confidence intervals that contain the true parameter values. In this particular case where σ_1 and σ_3 are equal, the odds-based inequalities are redundant, as they yield larger intervals than those obtained from the bounding inequalities.

Case 5 shows the bounding inequalities may fail to produce an empty confidence set when the researcher wrongly assumes that workers have complete information—although, as discussed above, the set is empty when we use more detailed instruments. Conversely, the confidence intervals defined by the odds-based inequalities alone, or by both types used jointly, are empty.

4.3 Alternative Estimators

Maximum Likelihood. To compare our moment inequality estimator with more traditional estimation approaches, we report in Table 2 maximum likelihood estimates (MLEs) computed under the assumption that all variables in the agent's information set are observed by the researcher. Given a choice of wage predictor z_s, we compute MLEs of $(\theta_\alpha, \theta_2, \theta_3)$ assuming z_s is all information worker s has on w_s; that is,

$$\underset{(\theta_\alpha,\theta_2,\theta_3)}{\operatorname{argmax}} \left\{ \sum_{s=1}^{S} \sum_{l=1}^{3} \mathbb{1}\{y_s^l = 1\} \ln \left(\frac{\exp(\theta_l + \theta_\alpha \mathbb{E}[w_s^l | z_s])}{\sum_{l'=1}^{3} \exp(\theta_{l'} + \theta_\alpha \mathbb{E}[w_s^{l'} | z_s])} \right) \right\}, \qquad \text{with } \theta_1 = 0.$$

[13] For example, in unreported results, we observe that the confidence set for the amenities defined by the odds-based inequalities is not a singleton when $\sigma_1 = 2$ and $\sigma_3 = 0$.

[14] In Appendix C.4, we illustrate in plots why the confidence intervals for θ_2 and θ_3 obtained from combining bounding and odds-based inequalities are smaller than when each inequality is considered in isolation.

Table 2: Simulation Results - Maximum Likelihood Estimator

Case	σ_1	σ_3	z_s	α	κ^2	κ^3
1	0	0	z_{2s}	1	0	1
2	0	1	z_{2s}	1	0	1
3	1	0	z_{2s}	0.91	0	0.92
4	1	1	z_{2s}	0.91	0	0.92
5	0	1	w_s	0.87	-0.03	0.87

The true parameter values are $\alpha = 1$, $\kappa^3 = 0$, and $\kappa^2 = 1$.

If $z_s = z_{2s}$ and, thus, $\mathbb{E}[w_s|z_s] = z_{2s}$, the MLE is consistent if and only if $\sigma_1 = 0$, as in cases 1 and 2, as only then does the worker's wage expectation coincide with the researcher's assumed expectation. Conversely, if $z_s = z_{2s}$ and $\sigma_1 > 0$, as in cases 3 and 4, the worker's true expectation and the researcher's assumed one do not coincide and, as a result, the MLE of all parameters is biased towards zero. In case 5, the researcher assumes workers have perfect information (i.e., $z_s = w_s$ and, thus, $\mathbb{E}[w_s|z_s] = w_s$) but, contrary to that assumption, workers make forecasting errors (i.e., $\sigma_3 > 0$), and the MLE is also biased.[15]

Comparing the estimates in tables 1 and 2 yields two conclusions. When the researcher observes a subset of the worker's true information set (i.e., if $\sigma_1 > 0$), the MLE is biased and our moment inequality estimator yields confidence intervals that contain the true parameter value. Moreover, as illustrated by cases 3 and 5, the confidence intervals produced by our moment inequality estimator may not include the corresponding MLEs.

Alternative moment inequality estimator. In Table C.4 in Appendix C.5, we apply an alternative inference procedure that only relies on the second-step inequalities described in Section 3.1. Specifically, we compute confidence intervals for the vector $(\theta_\alpha, \theta_l)$ for each location $l = 2, \ldots, L$ using the inequalities in equations (8) and (13). We obtain in this way $L - 1$ two-dimensional confidence sets. After projecting them on each of their elements, we obtain a confidence interval for each amenity term $(\theta_2, \ldots, \theta_L)$ and $L-1$ intervals for θ_α. We can then report a randomly chosen confidence interval for θ_α among the $L-1$ available ones. This alternative procedure yields results similar to those in Table 1 except in case 3, where adding information from the odds-based inequalities for the identification of θ_α results in a tighter confidence interval for this parameter.[16]

[15] Cases 3 and 5 show that, when the difference between the worker's true expectation and the researcher's assumed expectation have identical marginal distribution (i.e., if $\sigma_1 = \sigma_3$), the bias in the MLE is smaller if the researcher assigns workers an information set that is too small than if she assigns them an information set that is too large. Dickstein and Morales (2018) find a similar pattern in their setting.

[16] Although for all cases considered in Tables 1 and C.4 this alternative estimator yields weakly smaller confidence sets, unreported simulation results show that, when the agent's information set is heterogeneous across choices, its results are sensitive to which of the $L-1$ available intervals is chosen.

5 Empirical Application

We study individual migration events across local labor markets in Brazil between 2002 and 2011. We describe the data we use in Section 5.1, present estimates of the wage coefficient and migration cost in Section 5.2, and discuss tests of the content of workers' information sets in Section 5.3. We evaluate the effect of counterfactual changes in information sets and migration costs in Section 5.4.

5.1 Data

Our main data source is the *Relação Anual de Informações Sociais* (RAIS), an administrative dataset maintained by the Brazilian Ministry of Labor. It contains information on workers and establishments in the Brazilian formal labor market. We use the establishment's location (microregion, the closest equivalent to a commuting zone) and sector (industry) to define labor markets and measure workers' wages by their average monthly earnings.

Since our application will primarily explore workers' heterogeneous information sets according to their location, we restrict our sample to workers with similar demographic characteristics. Specifically, we study workers aged 25-64 with at least a high school degree identified as male and white.[17] Since RAIS only covers formal employment, a worker may be absent from the dataset because they are out of the labor force, unemployed, or informally employed. Hence, our conclusions will only refer to formal workers, and we restrict our sample to individuals with a persistent attachment to the formal labor market, selecting only those recorded in RAIS for at least seven years during our ten-year analysis period.[18]

To ensure that we observe a large number of individuals per market, we focus on 1,000 labor markets consisting of all combinations of the 50 microregions (out of 558) and 20 sectors (out of 51) with the largest total employment reported in RAIS. We then obtain our working sample by randomly sampling one million individuals per year among those employed in the 1,000 labor markets of interest. Appendix D provides more details on the RAIS data and the construction of our sample, and reports summary statistics on migration rates, wages, and other characteristics of the selected individuals and labor markets.

5.2 Estimation of Model Parameters

In Section 5.2.1, we describe the implementation of our moment inequality estimator. In Section 5.2.2, we discuss our estimates and compare them to the estimates we obtain using

[17]These are, respectively, the largest education, gender, and race categories in RAIS.

[18]Our estimation approach only requires the researcher to observe a subset of the labor markets that workers choose from; the informal sector can thus be conceived as part of the worker's unobserved choice set.

alternative estimation procedures.

5.2.1 Implementation of Moment Inequalities

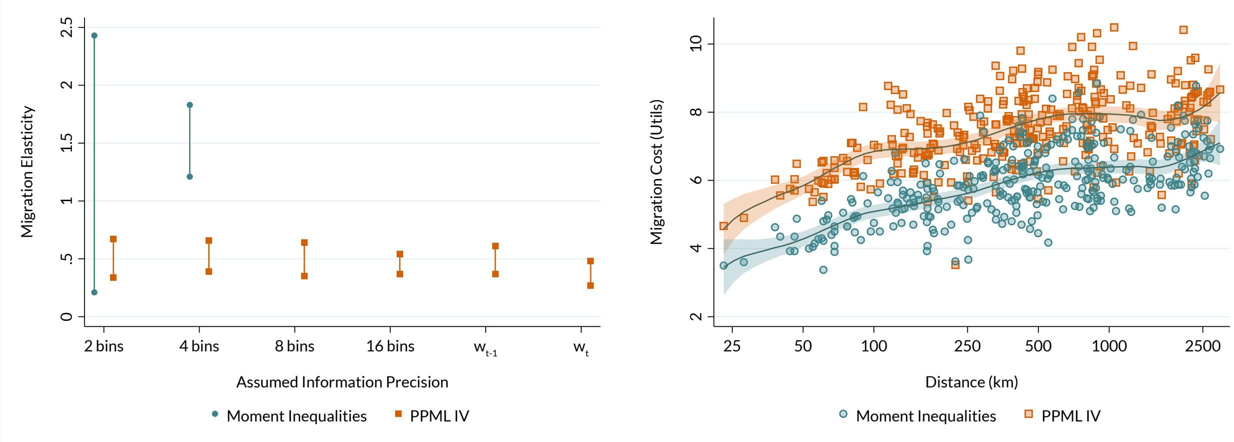
To map our empirical setting to the model described in Section 2, we consider multiple subpopulations of interest defined as the workers present in a given location n during a year $t-1$. We study their location decision at period t and equate the type s of a worker to their sector. While we assume that the wage coefficient α is constant across all sampled workers, we let the location-specific fixed effects in the vector κ vary across subpopulations. Hence, the location-specific parameter κ^l for any given l varies depending on the worker's prior location and the time period, thus accommodating time-varying unobserved migration costs, amenities, and price levels. For simplicity, we refer to these below simply as *amenities*.

We implement the two-step moment inequality estimation procedure described in Section 3. A key variable entering our moment inequalities is the wage difference $\Delta w_s^{ll'}$ for any locations l and l' and sector s. While our data contain worker- and year-specific wages for the sector-location pair in which the worker is employed, they naturally do not include any information on the wage they would have earned in other labor markets. We construct a proxy for these unobserved wages. We allow the log wage that a worker would obtain if employed in a particular sector and location to depend on individual worker characteristics (age and sectoral experience) with year-specific coefficients, on location-sector-year-specific fixed effects, on worker-sector-specific fixed effects, and a residual. Appendix E.1 provides details on the specification of the wage equation and reports the corresponding estimates. For each sector s and period t in our sample, these wage regressions yield estimates of the vector of labor market-specific wage shifters $w_{st} = (w_{st}^1, \ldots, w_{st}^L)$ entering our moment inequalities. These shifters account for labor supply and demand factors that impact the wages of all workers in a labor market. Since every other term in the wage equation is portable across locations, these shifters fully determine the differences in the wages that workers can receive in different labor markets. Thus, for simplicity, we refer to these below simply as *wages*.

As discussed in Section 3.3, building moment inequalities that bound the model parameters' true value requires specifying wage predictors $z_{st} = (z_{st}^1, \ldots, z_{st}^L)$ assumed to belong to the information set of all workers entering our inequalities. We present moment inequality estimates that use different wage predictors. Some of our estimates rely on assuming that workers know the previous year's relevant wages (i.e., $z_{st} = w_{st-1}$), while others only rely on assuming that workers are at least capable of indicating whether the lagged wages in a particular labor market belongs to one of $b \in \mathbb{N}$ quantiles, or bins, of the lagged wage distribution. To compute wage predictors consistent with these assumptions, we first compute the vector $(B_0^b, \ldots, B_b^b)$ of b-quantiles of the distribution of w_{st-1}^l across all sectors and locations,

Figure 1: Migration Elasticity and Amenities from Moment Inequalities vs. PPML-IV

(a) Migration Elasticity (95% conf. int.) (b) Amenities (95% conf. int. midpoint)



Panel (a) reports 95% confidence intervals of the wage coefficient α under different estimation procedures and informational assumptions. The blue circles delimit the moment inequality confidence intervals. The absence of circles for cases above 4 bins reflects that the associated confidence intervals are empty. The orange squares mark the PPML-IV confidence intervals. Informational assumptions are stronger as we move towards the right along the horizontal axis. In panel (b), each point equals the midpoint of a 95% confidence interval for a parameter κ^l_{nt}, for $t = 2011$. The blue circles indicate the midpoints of the moment inequality confidence interval, and the orange squares indicate the PPML-IV estimates. The fit lines are kernel-weighted local polynomial estimates, with the shaded area representing 95% confidence bands.

and then define the wage predictors as

$$z^l_{st} \equiv \sum_{k=1}^{b} \mathbb{1}\{B^b_{k-1} < w^l_{st-1} \leqslant B^b_k\}\mathbb{E}[w^l_{st-1}|B^b_{k-1} < w^l_{st-1} \leqslant B^b_k]. \tag{25}$$

We refer to b as the precision of workers' information. As b increases, the predictor in equation (25) becomes closer to w^l_{st-1}. Given a wage predictor z_{st}, we use the formula in equation (20) to compute eight instruments $\Delta x^{ll'}_{st}$ for every location pair l and l', sector s, and sample period t.[19] We provide additional details on the implementation of the moment inequalities in Appendix E.2.

5.2.2 Estimation Results

First-step estimates: wage coefficient. Panel (a) in Figure 1 reports 95% confidence intervals for the wage coefficient α under different informational assumptions. When we assume workers can only determine whether lagged wages in any given location are above or below the median of the distribution of wages across all labor markets, we obtain a 95% confidence

[19]The eight instruments result from setting $q = 4$ and $d \in \{-1, 1\}$.

interval for α that equals [0.24, 2.44]. The width of this interval reflects that the dummy variable indicating whether lagged wages are below or above the median is only loosely correlated with current wages. When we increase the assumed precision of workers' information and impose instead that workers can determine the quartile to which lagged market-specific wages belong, we obtain a tighter interval equal to [1.21, 1.83]. Assuming workers can classify locations according to more detailed quantiles of the distribution of wages, or that they know lagged or current wages without error, yields empty confidence sets. This leads us to reject that workers know lagged location-specific wages with any level of precision above quartiles. Thus, from now on, we use [1.21, 1.83] as our preferred set estimator of α.

For comparison, we also include 95% confidence intervals computed using the two-step PPML-IV estimator in Artuç and McLaren (2015). As discussed in Appendix E.3, this alternative estimation method yields point estimates of the wage coefficient at the expense of assuming that all workers employed in the same sector in a period t (regardless of their location of residence) have exactly the same information set, and consequently also the same wage expectations. This is a stronger assumption than the one required for our moment inequalities to bound the true parameter value, which requires the researcher to specify a (possibly different) variable that belongs to every worker information set, but does not restrict the additional information each worker may have, which may vary flexibly across workers and labor markets.[20] It is thus not surprising that the PPML-IV estimator yields confidence intervals for the wage coefficient that generally do not overlap with those generated by our moment inequality estimator: while the PPML-IV estimator yields confidence intervals between 0.3 and 0.6, the lowest value in our preferred moment inequality estimator is 1.21. The PPML-IV estimator of the wage coefficient is thus downward biased, underestimating the value workers assign to the expected monetary returns of migration relative to other non-income location payoffs.

Second-step estimates: amenities. Panel (b) in Figure 1 illustrates the moment inequality estimates of the parameters κ_{nt}^{l} for the year $t = 2011$ and all origin n and destination l locations in our sample. Specifically, this panel displays midpoints of the 95% moment inequality confidence interval for each amenity term κ_{nt}^{l}.[21] For comparison, we also display estimates of these amenities computed using the PPML-IV estimator. Although we estimate each parameter κ_{nt}^{l} without imposing any restriction on their variability, our estimated parameters tend to increase in the distance between locations n and l, consistently with these parameters

[20] E.g., the moment inequality confidence interval that uses quartiles of lagged wages as the relevant wage predictor is valid if workers, even within the same sector, location, and period, have different information, as long as all workers can at least classify labor markets into quartiles according to lagged market wages.

[21] For any two locations n and l and period t, we compute a 95% confidence interval for κ_{nt}^{l} as the union of 96% confidence intervals for κ_{nt}^{l} computed conditional on each value of θ_α in its 99% confidence interval; see Section 3.3. As shown in Appendix E.4, the resulting confidence intervals for the parameters κ_{nt}^{l} are tight.

accounting for migration costs in our model. The differences in levels between the PPML-IV estimates and the moment inequalities estimates are substantial, with the latter being on average 21% smaller than the former. Moreover, if we convert migration costs into their log-wage equivalents by dividing them by the estimates of the wage coefficient α produced by each estimation method, we find that the moment inequality estimates are 74% smaller. In sum, our empirical analysis indicates that estimation procedures commonly used in the migration literature yield upward biased estimates of migration costs or, more generally, of the relative importance of non-wage variables in workers' migration decisions.

5.3 Tests of Information Heterogeneity

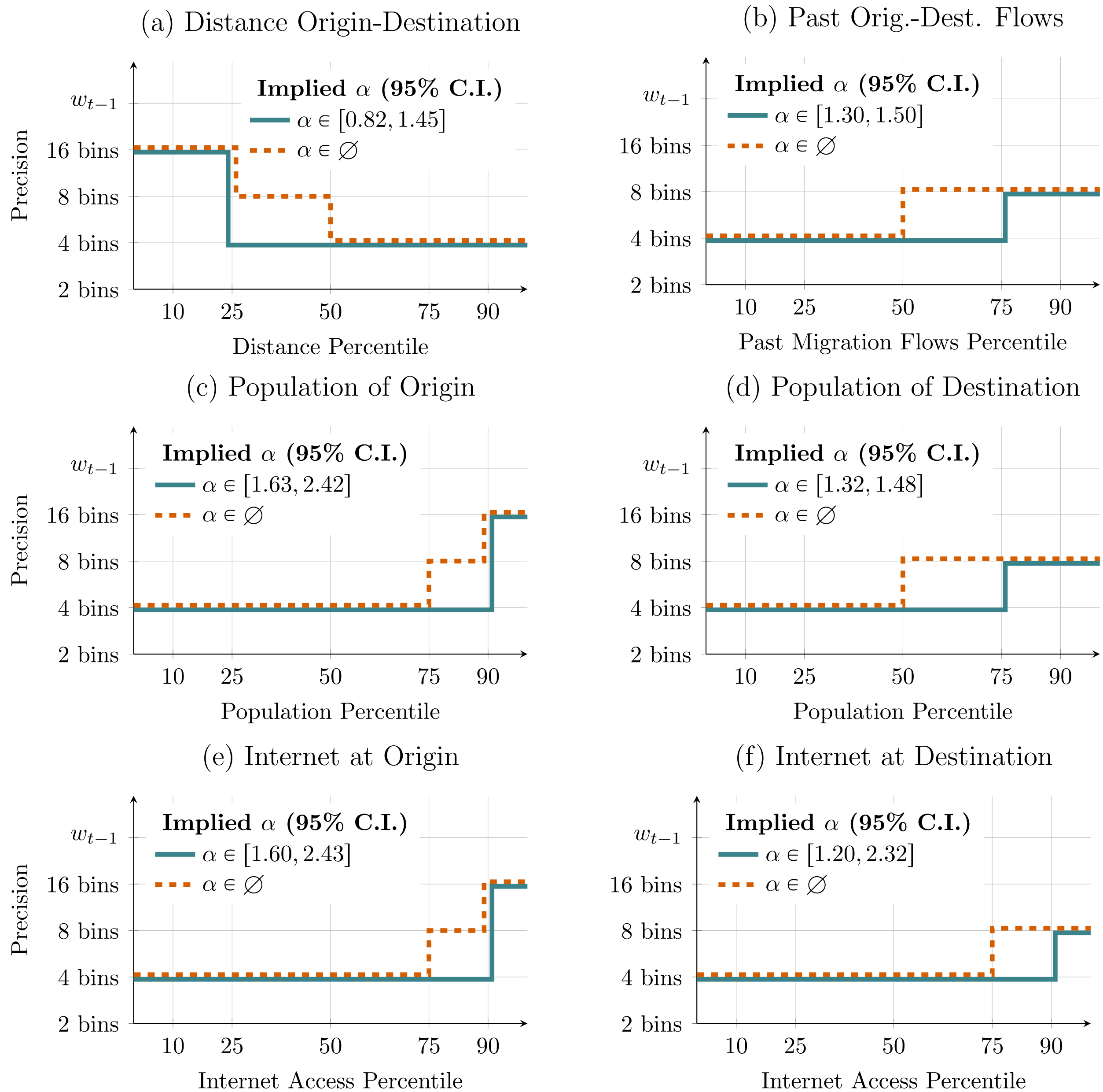
As panel (a) in Figure 1 showed, the moment inequality 95% confidence interval for the wage coefficient α is non-empty when we assume all workers know the quartile to which lagged wages belong, but becomes empty when we assume *all* workers can classify *all* markets into eight (or more) bins. It is possible, however, that some workers are more informed than others, or that they have more information about some markets than others.

In this section, we explore whether the migration patterns in our sample are consistent with workers having more precise information about wages in some, but not all, locations. We do so by testing whether the identified set for α defined by our moment inequalities becomes empty when we assume that some worker groups have additional wage information about specific groups of labor markets. We consider worker groups defined by the population and internet penetration in their prior location of residence, and groups of labor markets defined by their distance and past migration flows from the worker's location, as well as by the population and internet penetration in the market's location.[22]

We implement a similar testing approach for each dimension of heterogeneity described above. First, we classify workers and markets into six intervals delimited by the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of workers and markets along the corresponding dimension. Then, we order these intervals according to the direction along which we suspect information may be more precise. For example, we classify workers into intervals depending on the internet penetration in the location of residence and order these intervals from higher to lower internet penetration. Consistently with our finding in panel (a) in Figure 1, we start from a baseline information set according to which all workers can classify

[22]The distance between any two locations equals the geodesic distance between their centroids. Past migration flows are measured as the total number of workers recorded in RAIS as having migrated between any two locations in the three years prior to our sample period (1999-2001). The population of each location is computed as the total employment in RAIS in the period 1999-2001. The share of households with internet access in each location equals the average share of households with broadband internet access between 2007 and 2011, the period for which this information is available (see Appendix D).

Figure 2: Testing for Heterogeneous Information Sets



This figure displays patterns of information precision that can and cannot be rejected in the data. Each panel shows how these patterns vary along a key dimension, including distance, past migration flows, the population of origin and destination, and the share of households with internet access in the origin and destination. Patterns of information precision that can be rejected are shown in dotted lines, while patterns that cannot be rejected are shown in solid lines. We test each hypothesis by building an instrument function that defines wage proxies according to the assumed precision of information. These wage proxies reflect the characteristics of the origin and destination labor markets. For example, in the case of distance, the non-rejected pattern is tested by defining the wage predictor in a far-away labor market (beyond the 25^{th} quantile of distance) with lower precision (4 bins of lagged wages) than wage predictor in a nearby labor market (16 bins). We report the estimated 95% confidence interval of the migration elasticity that is consistent with the information assumption in each test.

all markets into quartiles, and test whether workers in the first interval can further classify markets into eight bins. This translates into wage predictors with different levels of precision across different combinations of workers and markets, affecting the values of the instruments in equation (20) and, thus, changing our moment inequalities. If the resulting 95% confidence interval for α is empty, we reject that assumption and end the testing procedure. If it is not empty, we increase the level of precision to 16 bins on that first interval and perform a new test. Calling B_j the maximum precision level tested and not rejected for the jth interval, the next iteration maintains B_j on that interval and searches for the maximum level of precision in the interval $j+1$, up to precision B_j. This procedure yields a weakly monotone information schedule along each dimension.

Figure 2 displays our results. Panel (a) shows we cannot reject workers are well informed about wages in labor markets that are within the 25th percentile of distance (383 km) of their location of residence. As discussed in Section 5.4.1, migration rates increase in workers' information. Thus, the well-known fact that migration rates decrease in distance (Beine et al., 2016) may be due less than previously thought to migration costs increasing in this dimension, and more to the role that information frictions play in migration choices. In panel (b), we observe that past migration flows between two locations are positively correlated with the information residents in one location have about wages in the other one. Given the impact information frictions have on the decision to migrate, this finding can partly explain the fact that workers from a particular origin tend to persistently migrate to the same destinations, providing an explanation for the impact that enclaves and social networks have on migration flows (e.g., Munshi, 2020). Panels (c) and (d) show that workers living in the five largest microregions by population have better information, and that all workers have more information about wages in the top quartile of regions by population. The information premium from living in highly populated areas adds to the skill-accumulation benefits of cities discussed in the prior literature (e.g., Glaeser and Maré, 2001; De la Roca and Puga, 2017). Finally, panels (e) and (f) provide evidence from one particular mechanism that may explain the findings in panels (c) and (d): workers living in regions with higher internet access have better information, and all workers have better information about regions with high internet access. This finding is consistent with the evidence on the informational impact of broadband internet access documented in the prior literature (e.g., Akerman et al., 2022; Grubanov-Boskovic et al., 2022).

5.4 Counterfactuals

The findings above indicate migrants face substantial information frictions. We now use our estimated model to predict how changes in the worker's information set and migration costs affect their migration choices and expected utility. In this analysis, we keep all elements of the economic environment constant and focus on workers' individual responses, ignoring the impact that widespread changes in workers' information and costs would have on migration flows through changes in equilibrium wages and prices (see Porcher, 2022, for such analysis).

We simulate the migration decisions made by agents facing wages that follow an AR(1) process with sector and location drifts. We calibrate this process using observed wages in our sample of 50 regions and 20 sectors over the period 2002-2011. Our estimates indicate wages are strongly serially correlated (with a persistence estimate of 0.93). Thus, the key potential source of workers' imperfect wage information is not the prevalence of unpredictable time-varying wage shocks but the lack of precise information on past wages. While our two-step estimation procedure makes it computationally feasible to obtain confidence intervals for all model parameters, computing model predictions that account for uncertainty in all parameter estimates is computationally very costly, as it requires building a multidimensional grid that spans all confidence intervals and evaluating our model at each point in that grid. As an alternative, we consider all points in the confidence interval for the wage coefficient α, but calibrate amenities by regressing the midpoint of the corresponding moment inequality confidence intervals (see Figure E.2 in Appendix E.4) on a constant and log distance.

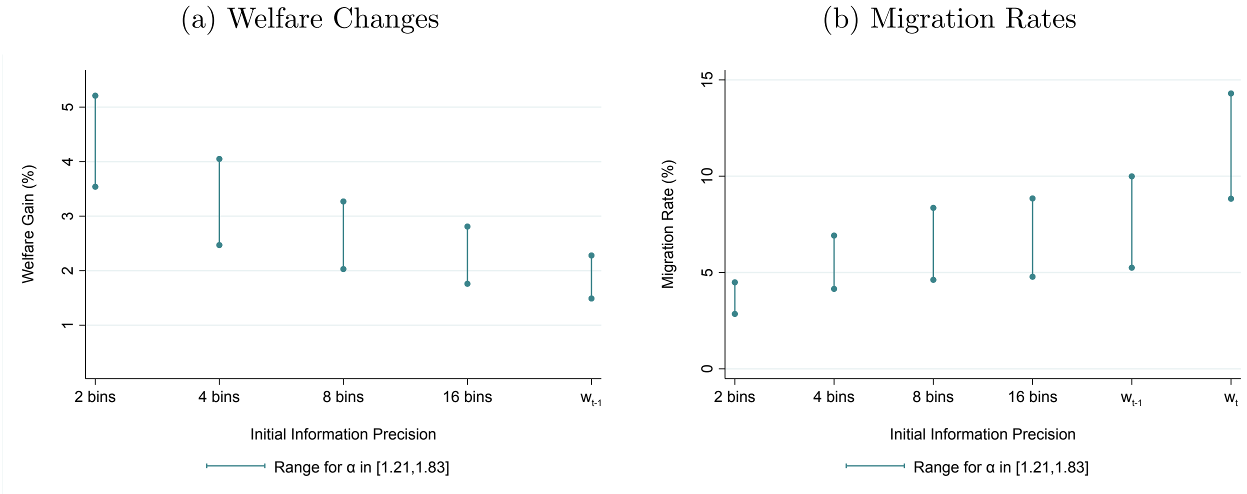
For each scenario, we evaluate the predictions of our estimated model for an economy that includes one million individuals drawn from the 2002 empirical distribution of workers across the 50 locations and 20 sectors in our estimation sample. We simulate the choices of these individuals for the period 2002-2011 and for 100 different simulation draws of the wage process. For each outcome variable of interest, we report the average value across all individuals and simulation draws.

5.4.1 Changes in Information Sets

We evaluate the impact on migration choices and welfare of giving workers information on market wages in all locations. Specifically, we focus on the impact of improving workers' information on both the migration probability, measured as the probability a worker changes locations in two consecutive periods, and welfare. We measure welfare as the average utility across simulated workers and periods, including the contribution of idiosyncratic tastes for locations and, importantly, using ex-post wages as the income measure. Hence, workers with perfect wage information choose locations exactly maximizing their utility, while workers with incomplete information maximize expected utility, and may choose locations that, ex-post, do not offer the highest utility.

The results are displayed in Figure 3. Panel (a) shows that the largest welfare gains are obtained by workers whose initial information only allows them to determine whether lagged

Figure 3: Effects of Providing Full Information About Wages



This figure displays counterfactual changes in welfare (panel (a)) and migration rates (panel (b)) as a result of giving workers perfect wage information. The intervals correspond to the range of model predictions consistent with value of α in the 95% confidence interval $[1.21, 1.83]$; see Section 5.2.2.

wages are above or below median. In this case, welfare gains are between 3.5% and 5.2%, with the highest gains generated by the model that sets the wage coefficient α to the largest value within its 95% confidence interval. The gains decrease but remain sizeable for workers who were initially better informed. Even if workers observed lagged wages perfectly—a hypothesis we reject in Panel (a) in Figure 5.2.2—the gains from observing contemporaneous wages perfectly would still range between 1.5 and 2.3%.

Panel (b) reports migration rates for workers with different information. Migration rates increase in the precision of the worker's wage information. They are below 5% for workers with the coarsest information set we consider, and between 9 and 14% when information is complete. This finding is in contradiction with the predictions of the model in Kaplan and Schulhofer-Wohl (2017), where workers improve their location-specific information by migrating, and tend to migrate more the worse their initial information is. Importantly, for our baseline information set according to which workers can discern the quartile to which lagged market wages belong, our estimated model predicts a migration rate between 4.2 and 6.9%, which aligns well with the migration rate of approximately 6% observed in the actual sample in 2002; see Figure D.1a in Appendix D.2.2.

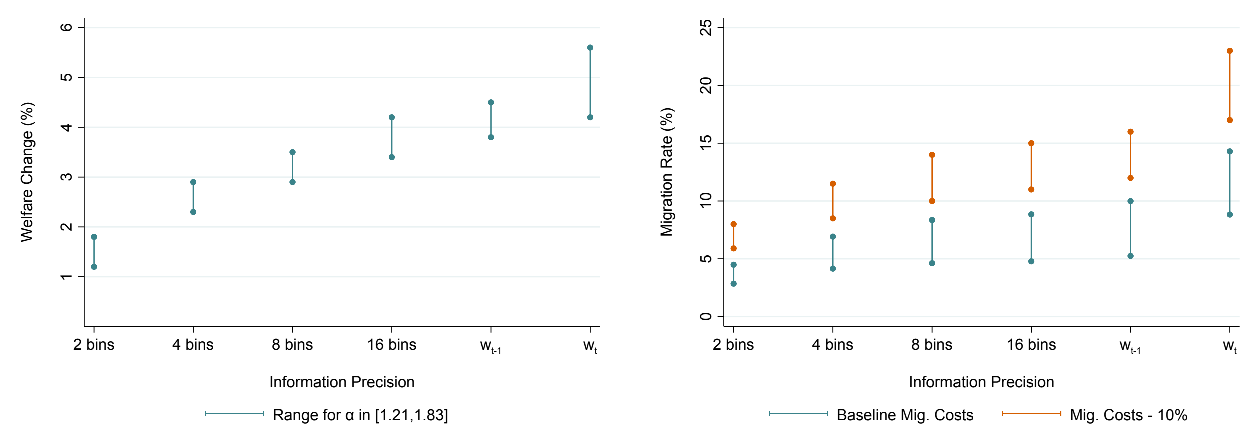
5.4.2 Reducing Barriers to Mobility

Reducing barriers to geographic mobility within a country is one of the main policy levers available to alleviate spatial misallocation (Morten and Oliveira, 2024). However, the benefits

Figure 4: Effects of Reducing Migration Costs, by Information Level

(a) Welfare Changes

(b) Migration Rates



This figure displays counterfactual changes in welfare (panel (a)) and migration rates (panel (b)) from a 10% reduction in migration costs, depending on workers' information. The intervals correspond to the range of model predictions consistent with a value of α in the 95% confidence interval $[1.21, 1.83]$; see Section 5.2.2.

of reducing mobility frictions likely depend on whether agents are well-informed about the economic opportunities in the newly accessible regions. In this section, we evaluate how the effects of such policies depend on the information available to potential migrants. Specifically, for several information sets, we compute the predictions of our estimated model for a 10% reduction in our calibrated migration costs.

Figure 4 displays the results. Panel (a) reveals that the welfare gains from a 10% reduction in migration costs increase greatly with the precision of workers' information on market wages. When workers are fully informed about wages in all locations, the welfare gains range from 4.2 to 5.7%, depending on the estimate of α. However, when workers can only discern whether lagged wages in each location are above or below median wages, the same reduction in migration costs only yields 1.2 to 1.8% welfare gains.

Panel (b) illustrates the increases in migration rates from reducing migration costs at each information level. The migration rates increase significantly for all information levels, and more so in relative terms for workers with a lower level of information precision. However, those larger increases in mobility have a higher rate of mistakes when the information precision is low, leading to the lower welfare gains reported in panel (a).

6 Extension: Dynamic Model of Location Choice

Our analysis so far allows for completely flexible migration costs, but assumes that agents are myopic. In this section, we describe how to extend our estimation method to settings with forward-looking agents facing one-time migration costs. In Section 6.1, we describe the assumptions of our dynamic migration model. In Section 6.2, we briefly describe how to adapt the procedure in Section 3 to the estimation of the parameters of the dynamic model. Appendix F provides additional details.

6.1 Theoretical Framework

The key departure from the model in Section 2 is that workers now determine their optimal choice at a period t internalizing its impact on future utility. Defining a variable y^l_{ist} that equals one if worker i of type s chooses location l at period t, and zero otherwise, we assume

$$y^l_{ist} \equiv \mathbb{1}\{l = \operatorname*{argmax}_{l'=1,\dots,L} \mathbb{E}[\mathcal{V}^{l'}_{ist}|\mathcal{J}_{ist}]\} \qquad \text{for } l = 1,\dots,L, \tag{26}$$

with $\mathcal{V}^l_{ist}$ the choice-specific value function and $\mathbb{E}[\cdot]$ defined as in equation (2). We impose:

$$\mathcal{V}^l_{ist} = v^l_{ist} + \varepsilon^l_{ist}, \tag{27a}$$

$$v^l_{ist} = \beta x^l_{nt} + \lambda^l_t + \alpha w^l_{ist} + \delta \mathcal{V}^{(lt)}_{ist+1}, \tag{27b}$$

where n indexes the location of worker i of type s and period $t-1$, and

$$\mathcal{V}^{(lt)}_{ist+1} \equiv \max_{l'=1,\dots,L} \mathbb{E}[\mathcal{V}^{(lt)l'}_{ist+1}|\mathcal{J}^{(lt)}_{ist+1}]. \tag{28}$$

Equation (27a) splits the choice-specific value function into the idiosyncratic component ε^l_{ist} and a variable v^l_{ist} that equation (27b) defines as the sum of four terms. First, the migration costs between locations n and l, modeled as a function of observed covariates x^l_{nt} and a vector of parameters β. Second, a location- and period-specific term λ^l_t, which captures a location's amenities and (log) price index. Third, the wage component αw^l_{ist}. Fourth, the product of the discount factor δ and a variable $\mathcal{V}^{(lt)}_{ist+1}$ that, according to equation (28), equals the worker's period-$t+1$ value function conditional on choosing alternative l at period t.[23,24]

[23] We assume for simplicity that types are exogenous, but one may apply our inequalities to a model in which workers optimally choose their type subject to transition costs as in, e.g., Caliendo et al. (2019, 2021).

[24] A comparison of equations (3) and (27) shows that, at the expense of assuming $\delta = 0$, the static model allows for a more flexible specification of migration costs, which may vary freely across locations and periods.

Defining $\lambda_t = (\lambda_t^1, \ldots, \lambda_t^L)$ and $\varepsilon_{ist} = (\varepsilon_{ist}^1, \ldots, \varepsilon_{ist}^L)$, we assume that

$$(\varepsilon_{ist}, \lambda_t, \alpha, \beta, \delta) \subseteq \mathcal{J}_{ist}. \tag{29}$$

Thus, when making their choice at period t, workers know the vectors of contemporaneous idiosyncratic preferences ε_{ist} and amenities λ_t, and the preference parameters α, β and δ. Equation (29) does not restrict the information workers have about the wage vector $w_{ist'} = (w_{ist'}^1, \ldots, w_{ist'}^L)$ for any $t' \geqslant t$ or the amenity vector $\lambda_{t'}$ for any $t' > t$.

While we do not specify the full content of workers' information sets, we limit the processes that determine them and assume that, for any $t' > t$,

$$\mathcal{J}_{ist'} \perp\!\!\!\perp y_{ist} | \mathcal{J}_{ist}. \tag{30}$$

Thus, conditional on the worker's information set at a period t, the worker's information set in subsequent periods does not depend on the worker's choice at t. Our framework thus does not allow for endogenous learning, understood as the process through which the worker's information set at t may depend on the history of locations visited by the worker.[25]

Defining $\Delta v_{ist}^{ll'} \equiv v_{ist}^l - v_{ist}^{l'}$, we impose that for any period t, locations l and l', types s and r, and workers i and j that share a common prior location n,

$$\mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{J}_{ist}, \mathcal{J}_{jrt}] = \mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{J}_{ist}] = \mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{W}_{ist}]. \tag{31}$$

The first equality imposes that every worker has at least as much information as any other worker of a different type r with whom it shares prior location n about differences in their own location-specific value functions. The second equality imposes that, once we condition on all other elements of the information set of worker i of type s at period t, the idiosyncratic preferences in ε_{ist} do not contain any information on $\Delta v_{ist}^{ll'}$ for any two locations l and l'. The variable $\Delta v_{ist}^{ll'}$ depends on the worker's future choices, which will depend on $\varepsilon_{ist'}$ for $t' > t$; thus, equation (31) will generally not hold unless ε_{ist} is independent over time.

As in the static model, data limitations force us to restrict the information workers have on location-specific wages. For any workers i and j of types s and r, respectively, and locations l and l', it holds that

$$\mathbb{E}[\Delta w_{ist}^{ll'} | \mathcal{W}_{ist}, \mathcal{W}_{jrt}] = \mathbb{E}[\Delta w_{st}^{ll'} | \mathcal{W}_{ist}]. \tag{32}$$

Thus, the worker's period-t expectation of the contemporaneous wage difference between two

[25] We thus do not allow for the type of learning specified in Kennan and Walker (2011), where workers are assumed to know the wage of a location if they had lived in it in the past.

locations l and l' equals the expectation of terms that do not vary across workers of the same type s. We do not restrict the information workers have about the difference in type- and location-specific wages $\Delta w_{st'}^{ll'}$ between any two locations l and l' and in any period t'.

Finally, as in equation (6), we assume that, for workers i and j of types s and r,

$$F_{\varepsilon}(\varepsilon_{ist}, \varepsilon_{jrt}|\mathcal{W}_{ist}, \mathcal{W}_{jrt}) = F_{\varepsilon}(\varepsilon_{ist})F_{\varepsilon}(\varepsilon_{jrt}) = \exp\bigg(- \sum_{l=1}^{L}(\exp(-\varepsilon_{ist}^{l}) + \exp(-\varepsilon_{jrt}^{l}))\bigg). \quad (33)$$

That is, the vectors ε_{ist} and ε_{jrt} are independent of $(\mathcal{W}_{ist}, \mathcal{W}_{jrt})$ and of each other, and each of their elements is *iid* according to a type I extreme value distribution.

The elements of λ_t are identified up to a location parameter. Thus, we normalize $\lambda_t^1 = 0$ for all t. Given this normalization, the model parameters are $(\lambda_t^2, \dots, \lambda_t^L)$ for all sample periods, the wage coefficient α, and the migration cost parameters β.

6.2 Estimation With Moment Inequalities

We provide a two-step estimation procedure. In the first step, we compute a confidence set for (α, β) using moment inequalities that difference out the amenity λ_t^l for any l and t. In the second step, for each $l = 2, \dots, L$ and sample period t, we derive inequalities that depend only on the parameters α, β, and λ_t^l, and combine these inequalities with the confidence set for (α, β) to compute a confidence interval for λ_t^l. We summarize here the approach to derive the second-step inequalities and provide further details in Appendix F.1.

We describe here how to apply bounding inequalities of the type introduced in Section 3.1.1 to the dynamic model in Section 6.1.[26] Following steps analogous to those taken to derive the static bounding inequality in equation (8), we obtain the following inequality,

$$\mathbb{E}[y_{ist}^{l'} - y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}) - (v_{ist}^{l} - v_{ist}^{l'}))|z_{st}] \geqslant 0. \quad (34)$$

This inequality cannot be used for estimation as the value function difference $v_{ist}^{l} - v_{ist}^{l'}$ is not a function only of observed covariates and parameters.

We follow Morales et al. (2019) and implement a discrete analogue of Euler's perturbation method to derive an inequality that can be used for estimation. Specifically, we substitute $v_{ist}^{l'}$ in equation (34) by a function $\tilde{v}_{ist}^{l'}$, where $v_{ist}^{l'}$ and $\tilde{v}_{ist}^{l'}$ differ in that the latter conditions on the choices that, from period $t+1$ onwards, would be optimal for worker i of type s if they had chosen alternative l at t. As our dynamic model exhibits one-period dependence, $v_{ist}^{l} - \tilde{v}_{ist}^{l'}$ is a function exclusively of the difference in static utilities at period t and the

[26] While it may be feasible to use the odds-based moment inequalities introduced in Section 3.1.2 in the context of our dynamic model, we have not found the way of doing so.

discounted difference in static utilities at period $t+1$ that are due to whether the worker chooses alternatives l or l' at period t. Specifically,

$$v_{ist}^{l} - \tilde{v}_{ist}^{l'} = u_{ist}^{l} - u_{ist}^{l'} + \delta\beta \sum_{l''=1}^{L} y_{ist+1}^{(lt)l''}(x_{lt+1}^{l''} - x_{l't+1}^{l''}), \tag{35}$$

where $y_{ist+1}^{(lt)l''}$ is the optimal choice at period $t+1$ of worker i of type s if they were to choose alternative l at t. The expression in equation (35) is a function of observed covariates and parameters. Moreover, $v_{ist}^{l'} \geqslant \tilde{v}_{ist}^{l'}$ for every worker, period, and choices l and l'. Thus, the sign of the moment inequality in equation (34) is preserved if $\tilde{v}_{ist}^{l'}$ takes the place of $v_{ist}^{l'}$. We show formally in Appendix F.1.2 that the resulting inequality bounds the amenity λ_t^l for all $l = 2, \ldots, L$ and period t. Appendix F.1.1 shows how to derive moment inequalities that are informative about (α, β).

7 Conclusion

We introduce a new method to estimate discrete choice models when agents' information sets are unobserved to the researcher and potentially heterogeneous between individuals, and when the choice set is arbitrarily large and payoffs are parameterized with choice-specific fixed effects. In the context of location choice, our method allows both information frictions and migration costs to vary flexibly between individuals, locations, and over time, and lets us separately identify the role of information frictions and migration costs in workers' location choices.

The application of our moment inequality estimator to data on internal migration choices in Brazil yields four takeaways. First, workers have coarse and heterogeneous information about wages in other local markets. In particular, workers' location choices are consistent with them observing wages in nearby labor markets with higher precision. Second, accounting for this incomplete information substantially alters the mapping from observed location choices and wages to workers' preferences for wages and non-wage attributes. Our wage preference estimates are three times larger than those from common estimation procedures, whereas our migration cost estimates are, on average, 21% lower. Third, providing workers with complete information in our estimated model would increase both migration rates and average welfare. Fourth, policies that reduce migration costs by improving transportation infrastructure or reducing regulatory barriers are much less effective at inducing mobility and improving migrants' welfare when workers face information frictions of the magnitude we recover.

Our results emphasize that information frictions affect people's behavior and, thus, observed choices do not reflect their preferences alone. Moreover, even with correct estimates of people's preferences, information frictions affect the predicted effects of policy proposals, including those that do not target information frictions directly. These insights, as well as our method, apply beyond the context of location choice. Deciding which schools to apply to and which health insurance plan to select are examples of other economically consequential decisions that most people make with limited information. We view the application of our method to study these other decisions as a promising avenue for research.

References

Akerman, Anders, Edwin Leuven, and Magne Mogstad, "Information Frictions, Internet and the Relationship between Distance and Trade," *American Economic Journal: Applied Economics*, 2022, *14* (1), 133–163. [28]

Andrews, Donald W. K. and Soonwoo Kwon, "Misspecified Moment Inequality Models: Inference and Diagnostics," *Review of Economic Studies*, 2024, *91* (3), 45–76. [18]

Andrews, Donald W.K. and Gustavo Soares, "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 2010, *78*, 119–157. [16, 19]

Artuç, Erhan and John McLaren, "Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility," *Journal of International Economics*, 2015, *97* (2), 278–294. [4, 5, 25]

Baseler, Travis, "Hidden Income and the Perceived Returns to Migration: Experimental Evidence from Kenya," *American Economic Journal: Applied Economics*, 2023, *15* (4), 321–352. [5]

Bei, Xinyue, "Inference on Union Bounds with Applications to DiD, RDD, Bunching, and Structural Counterfactuals," *mimeo*, January 2024. [16]

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga, "A Practitioners' Guide to Gravity Models of International Migration," *The World Economy*, 2016, *39* (4), 496–512. [28]

Bergman, Peter, Eric Chan, and Adam Kapor, "Housing Search Frictions: Evidence from Detailed Search Data and a Field Experiment," *mimeo*, May 2020. [5]

_, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F Katz, and Christopher Palmer, "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice," *mimeo*, August 2023. [5]

Bombardini, Matilde, Bingjing Li, and Francesco Trebbi, "Did U.S. Politicians Expect the China Shock?," *American Economic Review*, 2023, *113* (1), 174–209. [3, 6, 10]

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 2014, *82* (5), 1671–1748. [5]

Caliendo, Lorenzo, Luca David Opromolla, Fernando Parro, and Alessandro Sforza, "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement," *Journal of Political Economy*, 2021, *129* (12), 3491–3545. [5, 32]

_ , Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, *87* (3), 741–835. [5, 32]

Canay, Ivan A., Gastón Illanes, and Amilcar Velez, "A User's Guide for Inference in Models Defined by Moment Inequalities," *Journal of Econometrics*, 2023, *accepted.* [2]

De la Roca, Jorge and Diego Puga, "Learning by Working in Big Cities," *Review of Economic Studies*, January 2017, *84* (1), 106–142. [28]

Desmet, Klaus, Dáavid Kristián Nagy, and Esteban Rossi-Hansberg, "The Geography of Development," *Journal Political Economy*, 2018, *126* (3), 903–983. [5]

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "Who Benefits from Rent Control? The Equilibrium Consequences of San Francisco's Rent Control Expansion," *Working Paper*, 2019. [7]

Dickstein, Michael J. and Eduardo Morales, "What do Exporters Know?," *The Quarterly Journal of Economics*, 2018, *133* (4), 1753–1801. [3, 6, 10, 21]

_ , Jihye Jeon, and Eduardo Morales, "Patient Costs and Physicians' Information," *mimeo*, 2023. [3, 6, 10]

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885. [4, 8]

Eizenberg, Alon, "Upstream Innovation and Product Variety in the U.S. Home PC Market," *Review of Economic Studies*, 2014, *81*, 1003–1045. [3]

Glaeser, Edward L. and David C. Maré, "Cities and Skills," *Journal of Labor Economics*, 2001, *19* (2), 316–342. [28]

Grubanov-Boskovic, Sara, Sona Kalantaryan, Silvia Migali, and Marco Scipioni, "The impact of the Internet on migration aspirations and intentions," *Migration Studies*, 2022, *9* (4), 1807–1822. [28]

Ho, Katherine and Ariel Pakes, "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians," *American Economic Review*, 2014, *104* (12), 3841–3884. [3]

Jia, Ning, Raven Molloy, Christopher Smith, and Abigail Wozniak, "The Economics of Internal Migration: Advances and Policy Questions," *Journal of Economic Literature*, 2023, *61* (1), 144–180. [1]

Kaplan, Greg and Sam Schulhofer-Wohl, "Understanding the Long-run Decline in Interstate Migration," *International Economic Review*, 2017, *58* (1). [5, 9, 30]

Kennan, John and James R. Walker, "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 2011, *79* (1), 211–251. [4, 5, 9, 33]

Manski, Charles F., "Nonparametric Estimation of Expectations in the Analysis of Discrete Choice Under Uncertainty," in William Barnett, James Powell, and George Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge: Cambridge University Press, 1991. [5]

_ , "Measuring Expectations," *Econometrica*, 2004, *5* (9), 1329–1376. [5]

Molinari, Francesca, "Microeconometrics with Partial Identification," in Steven N. Durlauf, Lars Peter Hansen, James J. Heckman, and Rosa L. Matzkin, eds., *Handbook of Econometrics, Volume 7A*, Elsevier, 2020, pp. 355–486. [18]

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended Gravity," *The Review of Economic Studies*, 2019, *86* (6), 2668–2712. [3, 34]

Morten, Melanie and Jaqueline Oliveira, "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," *American Economic Journal: Applied Economics*, 2024, *16* (2). [5, 30]

Munshi, Kaivan, "Social Networks and Migration," *Annual Review of Economics*, 2020, *12* (Volume 12, 2020), 503–524. [28]

Pakes, Ariel, "Alternative Models for Moment Inequalities," *Econometrica*, 2010, *78* (6), 1783–1822. [1, 3]

_ , Jack R. Porter, Katherine Ho, and Joy Ishii, "Moment Inequalities and their Application," *Econometrica*, 2015, *83* (1), 315–334. [1]

Porcher, Charly, "Migration with Costly Information," *mimeo*, 2022. [5, 9, 29]

Redding, Stephen J. and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58. [9]

Tombe, Trevor and Xiaodong Zhu, "Trade, Migration and Productivity: A Quantitative Analysis of China," *American Economic Review*, 2019, *forthcoming.* [5]

Traiberman, Sharon, "Occupations and Import Competition," *American Economic Review*, 2019, *109* (12), 4260–4301. [6, 8]

Wilson, Riley, "Moving to Jobs: The Role of Information in Migration Decisions," *Journal of Labor Economics*, 2021, *39* (4), 1083–1128. [5]

A Main Proofs

A.1 Proof of Theorem 1

Equation (1) implies that, for any worker i of type s and locations l and l', it holds that

$$(y_{is}^{l} + y_{is}^{l'})(\mathbb{1}\{\mathbb{E}[\mathcal{U}_{is}^{l} - \mathcal{U}_{is}^{l'}|\mathcal{J}_{is}] \geqslant 0\} - y_{is}^{l}) = 0.$$

Equations (3) to (5) imply we can rewrite this equality as

$$(y_{is}^{l} + y_{is}^{l'})(\mathbb{1}\{\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \Delta\varepsilon_{is}^{ll'} \geqslant 0\} - y_{is}^{l}) = 0, \tag{A.1}$$

where $\Delta\kappa^{ll'} = \kappa^{l} - \kappa^{l'}$, $\Delta w_{is}^{ll'} = w_{is}^{l} - w_{is}^{l'}$, and $\Delta\varepsilon_{is}^{ll'} = \varepsilon_{is}^{l} - \varepsilon_{is}^{l'}$. Taking the expectation of this equality conditional on $\mathcal{W}_{is}$ and a dummy variable that equals one if worker i of type s chooses either location l or location l', we obtain

$$\mathbb{E}[\mathbb{1}\{\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \Delta\varepsilon_{is}^{ll'} \geqslant 0\} - y_{is}^{l}|\mathcal{W}_{is}, y_{is}^{l} + y_{is}^{l'} = 1] = 0.$$

Equation (6) implies that $\Delta\varepsilon_{is}^{ll'}$ follows a Logit distribution, so integrating over $\Delta\varepsilon_{is}^{ll'}$ yields

$$\mathbb{E}[(1 + \exp(-\Delta\kappa^{ll'} - \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]))^{-1} - y_{is}^{l}|\mathcal{W}_{is}, y_{is}^{l} + y_{is}^{l'} = 1] = 0.$$

Multiplying by $1 + \exp(-\Delta\kappa^{ll'} - \alpha\mathbb{E}[\Delta w_{s}^{ll'}|\mathcal{W}_{is}])$, we obtain

$$\mathbb{E}[1 - y_{is}^{l} - y_{is}^{l}\exp(-\Delta\kappa^{ll'} - \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])|\mathcal{W}_{is}, y_{is}^{l} + y_{is}^{l'} = 1] = 0.$$

Given the conditioning on the event $y_{is}^{l} + y_{is}^{l'} = 1$, we can rewrite this moment equality as

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])))|\mathcal{W}_{is}] = 0. \tag{A.2}$$

As $-\exp(-x)$ is concave in $x \in \mathbb{R}$, a linear approximation to this function at any $a \in \mathbb{R}$ bounds it from above. The formula for this approximation is $-\exp(-a)(1 + a - x)$. Thus, given a function $h_{is}^{ll'}\colon \mathcal{Z}_s \times \Theta_{ll'} \to \mathbb{R}$ and equation (A.2), we derive the following inequality

$$\mathbb{E}[y_{is}^{l'} - y_{is}^{l}\exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}) - \Delta\kappa^{ll'} - \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])|\mathcal{W}_{is}] \geqslant 0. \tag{A.3}$$

Let's consider the moment

$$\mathbb{E}[y_{is}^{l'} - y_{is}^{l}\exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}) - \Delta\kappa^{ll'} - \alpha\Delta w_{is}^{ll'})|\mathcal{W}_{is}]. \tag{A.4}$$

Define $\nu_{is}^{ll'} \equiv \Delta w_{is}^{ll'} - \mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]$. As $\mathcal{W}_{is} \subset \mathcal{J}_{is}$, we use the LIE to write equation (A.4) as

$$\mathbb{E}[\mathbb{E}[y_{is}^{l'} - y_{is}^{l} \exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}) - \Delta\kappa^{ll'} - \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] - \nu_{is}^{ll'}))|\mathcal{J}_{is}]|\mathcal{W}_{is}].$$

Equation (1) implies $\mathbb{E}[y_{is}^{l}|\mathcal{J}_{is}] = y_{is}^{l}$, equation (4) implies $(\Delta\kappa^{ll'}, \alpha) \subset \mathcal{J}_{is}$, and, by definition, $\mathcal{W}_{is} \subset \mathcal{J}_{is}$. Therefore, if $z_s \subseteq \mathcal{W}_{is}$ and $h_{is}^{ll'}(\cdot)$ is deterministic, the moment above equals

$$\mathbb{E}[y_{is}^{l'} - y_{is}^{l} \exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}) - \Delta\kappa^{ll'} - \alpha\mathbb{E}[\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'}|\mathcal{J}_{is}])|\mathcal{W}_{is}].$$

Equation (5) implies $\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{J}_{is}] = \mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]$ and equation (2) implies $\mathbb{E}[\nu_{is}^{ll'}|\mathcal{W}_{is}] = 0$, so we can rewrite the moment in equation (A.4) as

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l} \exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(-(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'})) + \Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])|\mathcal{W}_{is}].$$

However, this moment is exactly the same entering the moment inequality in equation (A.3), which implies that the following inequality involving the moment in equation (A.4) is equivalent to that in equation (A.3):

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l} \exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(-(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'})) + \Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})|\mathcal{W}_{is}] \geqslant 0.$$

Finally, if $z_s \subset \mathcal{W}_{is}$, we can use the LIE and conclude that

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l} \exp(-h_{is}^{ll'}(z_s, \Delta\kappa^{ll'}))(-(1 + h_{is}^{ll'}(z_s, \Delta\kappa^{ll'})) + \Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})|z_s] \geqslant 0. \tag{A.5}$$

The moment in this inequality is precisely the moment in equation (7) when evaluated at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$. Thus, equation (A.5) implies Theorem 1. ■

A.2 Proof of Theorem 2

To prove Theorem 2, we first consider equation (A.2). Let's consider the moment

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})))|\mathcal{W}_{is}] \tag{A.6}$$

or, equivalently,

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'} + \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'}))))|\mathcal{W}_{is}],$$

where $\nu_{is}^{ll'} \equiv \Delta w_{is}^{ll'} - \mathbb{E}[\Delta w_{s}^{ll'}|\mathcal{W}_{is}]$. Given that $\mathcal{W}_{is} \subset \mathcal{J}_{is}$, we can use the LIE to rewrite this moment as

$$\mathbb{E}[\mathbb{E}[y_{is}^{l'} + y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'} + \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'}))))|\mathcal{J}_{is}]|\mathcal{W}_{is}].$$

Equation (1) implies $\mathbb{E}[y_{is}^{l}|\mathcal{J}_{is}] = y_{is}^{l}$. Consequently, we can further rewrite

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}\mathbb{E}[-\exp(-(\Delta\kappa^{ll'} + \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'})))|\mathcal{J}_{is}]|\mathcal{W}_{is}].$$

Equation (4) and the definition of $\mathcal{W}_{is}$ as including all elements of $\mathcal{J}_{is}$ other than ε_{is} implies that $(\Delta\kappa^{ll'}, \alpha) \subset \mathcal{W}_{is}$. As $\Delta w_{is}^{ll'} = \mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'}$, equation (5) further implies that $\mathbb{E}[\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'}|\mathcal{J}_{is}] = \mathbb{E}[\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'}|\mathcal{W}_{is}]$. Thus, we rewrite the moment above,

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}\mathbb{E}[-\exp(-(\Delta\kappa^{ll'} + \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'})))|\mathcal{W}_{is}]|\mathcal{W}_{is}]. \tag{A.7}$$

As $-\exp(x)$ is concave in $x \in \mathbb{R}$, equation (2) and Jensen's inequality imply the inequality

$$\begin{gathered}\mathbb{E}[y_{is}^{l'} + y_{is}^{l}\mathbb{E}[-\exp(-(\Delta\kappa^{ll'} + \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'})))|\mathcal{W}_{is}]|\mathcal{W}_{is}] \\ \leqslant \\ \mathbb{E}[y_{is}^{l'} + y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])))|\mathcal{W}_{is}].\end{gathered}$$

The right-hand side of this inequality coincides with equation (A.2) and, thus, we conclude

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}\mathbb{E}[-\exp(-(\Delta\kappa^{ll'} + \alpha(\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}] + \nu_{is}^{ll'})))|\mathcal{W}_{is}]|\mathcal{W}_{is}] \leqslant 0.$$

As the moments in equations (A.6) and (A.7) are equivalent, this inequality rewrites as

$$\mathbb{E}[y_{is}^{l'} + y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})))|\mathcal{W}_{is}] \leqslant 0.$$

Multiplying by -1 on both sides of this equation, we obtain the following inequality

$$\mathbb{E}[y_{is}^{l}\exp(-(\Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})) - y_{is}^{l'}|\mathcal{W}_{is}] \geqslant 0.$$

Finally, if $z_s \subset \mathcal{W}_{is}$, we can use the LIE to conclude

$$\mathbb{E}[y_{is}^{l}\exp(-(\Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})) - y_{is}^{l'}|z_s] \geqslant 0. \tag{A.8}$$

The moment in this inequality is precisely the moment in equation (12) when evaluated at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$. Thus, equation (A.8) implies Theorem 2. ■

A.3 Proof of Theorem 3

For any locations l and l' and any worker i of type s, equation (A.1) implies

$$(y_{is}^{l}+y_{is}^{l'})(\mathbb{1}\{\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]+\Delta\varepsilon_{is}^{ll'}\geqslant 0\}-y_{is}^{l})=0.$$

For any locations l and l', any worker i of type s, and any worker j of type r, we thus have

$$y_{jr}^{l'}(y_{is}^{l}+y_{is}^{l'})(\mathbb{1}\{\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]+\Delta\varepsilon_{is}^{ll'}\geqslant 0\}-y_{is}^{l})=0. \tag{A.9}$$

Next, we take the expectation of this equality conditional on $\mathcal{W}_{is}$, on $\mathcal{W}_{jr}$, and on a dummy variable that equals one if worker i of type s chooses either location l or location l'; i.e.,

$$\mathbb{E}[y_{jr}^{l'}(\mathbb{1}\{\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]+\Delta\varepsilon_{is}^{ll'}\geqslant 0\}-y_{is}^{l})|\mathcal{W}_{is},\mathcal{W}_{jr},y_{is}^{l}+y_{is}^{l'}=1]=0.$$

Given equations (1) and (6), we rewrite this moment equality, after integrating over $\Delta\varepsilon_{is}^{ll'}$, as

$$\mathbb{E}[y_{jr}^{l'}((1+\exp(-\Delta\kappa^{ll'}-\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]))^{-1}-y_{is}^{l})|\mathcal{W}_{is},\mathcal{W}_{jr},y_{is}^{l}+y_{is}^{l'}=1]=0.$$

Multiplying by $1+\exp(-\Delta\kappa^{ll'}-\alpha\mathbb{E}[\Delta w_{s}^{ll'}|\mathcal{W}_{is}])$, we obtain

$$\mathbb{E}[y_{jr}^{l'}(1-y_{is}^{l}(1+\exp(-\Delta\kappa^{ll'}-\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])))|\mathcal{W}_{is},\mathcal{W}_{jr},y_{is}^{l}+y_{is}^{l'}=1]=0,$$

or, equivalently,

$$\mathbb{E}[y_{jr}^{l'}(1-y_{is}^{l}-y_{is}^{l}\exp(-\Delta\kappa^{ll'}-\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]))|\mathcal{W}_{is},\mathcal{W}_{jr},y_{is}^{l}+y_{is}^{l'}=1]=0.$$

Given that this expectation conditions on the event $y_{is}^{l}+y_{is}^{l'}=1$, we can further rewrite

$$\mathbb{E}[y_{jr}^{l'}(y_{is}^{l'}+y_{is}^{l}(-\exp(-(\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}]))))|\mathcal{W}_{is},\mathcal{W}_{jr}]=0,$$

or, equivalently,

$$\mathbb{E}[y_{is}^{l'}y_{jr}^{l'}+y_{is}^{l}y_{jr}^{l'}(-\exp(-(\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w_{is}^{ll'}|\mathcal{W}_{is}])))|\mathcal{W}_{is},\mathcal{W}_{jr}]=0.$$

As the function $-\exp(-x)$ is concave in $x\in\mathbb{R}$, given any deterministic function $g_{ijsr}^{ll'}\colon\mathcal{Z}_s\times\mathcal{Z}_r\times\Theta_\alpha\to\mathbb{R}$, we can derive the following inequality

$$\mathbb{E}[y_{is}^{l'}y_{jr}^{l'}+y_{is}^{l}y_{jr}^{l'}\exp(-g_{ijsr}^{ll'}(z_s,z_r,\alpha))\times$$

$$(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}])|\mathcal{W}_{is},\mathcal{W}_{jr}]\geqslant 0. \tag{A.10}$$

Let's consider the moment

$$\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha\Delta w^{ll'}_{is})|\mathcal{W}_{is},\mathcal{W}_{jr}], \tag{A.11}$$

or, equivalently,

$$\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))\times$$
$$(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha(\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}]+\nu^{ll'}_{is}))|\mathcal{W}_{is},\mathcal{W}_{jr}],$$

where $\nu^{ll'}_{is}\equiv\Delta w^{ll'}_{is}-\mathbb{E}[\Delta w^{ll'}_{s}|\mathcal{W}_{is}]$. Equation (5) implies we can also write $\nu^{ll'}_{is}\equiv\Delta w^{ll'}_{is}-\mathbb{E}[\Delta w^{ll'}_{s}|\mathcal{W}_{is},\mathcal{W}_{jr}]$ and conclude that $\mathbb{E}[\nu^{ll'}_{is}|\mathcal{W}_{is},\mathcal{W}_{jr}]=0$. Given that $\mathcal{W}_{is}\subset\mathcal{J}_{is}$ and $\mathcal{W}_{jr}\subset\mathcal{J}_{jr}$, we can use the LIE to rewrite the moment inequality above as

$$\mathbb{E}[\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))\times$$
$$(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha(\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}]+\nu^{ll'}_{is}))|\mathcal{J}_{is},\mathcal{J}_{jr}]|\mathcal{W}_{is},\mathcal{W}_{jr}].$$

Equation (1) implies $\mathbb{E}[y^{l}_{is}y^{l'}_{jr}|\mathcal{J}_{is},\mathcal{J}_{jr}]=y^{l}_{is}y^{l'}_{jr}$, and equation (4) implies $(\Delta\kappa^{ll'},\alpha)\subset(\mathcal{J}_{is}\cap\mathcal{J}_{jr})$. Consequently, if $z_s\subseteq\mathcal{W}_{is}$ and $z_r\subseteq\mathcal{W}_{jr}$, it is then the case that $z_s\subset\mathcal{J}_{is}$ and $z_r\subset\mathcal{J}_{jr}$, and we can thus further rewrite

$$\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))\times$$
$$(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha\mathbb{E}[\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}]+\nu^{ll'}_{is}|\mathcal{J}_{is},\mathcal{J}_{jr}])|\mathcal{W}_{is},\mathcal{W}_{jr}].$$

As $\Delta w^{ll'}_{is}=\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}]+\nu^{ll'}_{is}$, equation (5) further implies that

$$\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))\times$$
$$(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha\mathbb{E}[\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}]+\nu^{ll'}_{is}|\mathcal{W}_{is},\mathcal{W}_{jr}])|\mathcal{W}_{is},\mathcal{W}_{jr}],$$

and $\mathbb{E}[\nu^{ll'}_{is}|\mathcal{W}_{is},\mathcal{W}_{jr}]=0$ implies we can rewrite this moment as

$$\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w^{ll'}_{is}|\mathcal{W}_{is}])|\mathcal{W}_{is},\mathcal{W}_{jr}].$$

Given equation (5), we can further rewrite this moment as

$$\mathbb{E}[y^{l'}_{is}y^{l'}_{jr}+y^{l}_{is}y^{l'}_{jr}\exp(-g^{ll'}_{ijsr}(z_s,z_r,\alpha))(-(1+g^{ll'}_{ijsr}(z_s,z_r,\alpha))+\Delta\kappa^{ll'}+\alpha\mathbb{E}[\Delta w^{ll'}_{s}|\mathcal{W}_{is}])|\mathcal{W}_{is},\mathcal{W}_{jr}].$$

This moment is the same as the one entering the inequality in equation (A.10), which implies that the following inequality, written with the moment in (A.11), is equivalent to (A.10),

$$\mathbb{E}[y_{is}^{l'}y_{jr}^{l'} + y_{is}^{l}y_{jr}^{l'}\exp(-g_{ijsr}^{ll'}(z_s, z_r, \alpha))(-(1 + g_{ijsr}^{ll'}(z_s, z_r, \alpha)) + \Delta\kappa^{ll'} + \alpha\Delta w_{is}^{ll'})|\mathcal{W}_{is}, \mathcal{W}_{jr}] \geqslant 0. \tag{A.12}$$

This moment inequality is one of the two that we will combine to obtain the inequality that we use to bound the parameter α. To obtain the second moment inequality, we start from

$$y_{is}^{l}(y_{jr}^{l} + y_{jr}^{l'})(\mathbb{1}\{\Delta\kappa^{l'l} + \alpha\mathbb{E}[\Delta w_{jr}^{l'l}|\mathcal{W}_{jr}] + \Delta\varepsilon_{jr}^{l'l} \geqslant 0\} - y_{jr}^{l'}) = 0, \tag{A.13}$$

which is analogous to that in equation (A.9). Following the same steps described above to go from equation (A.9) to equation (A.12), we can derive the following inequality

$$\mathbb{E}[y_{is}^{l}y_{jr}^{l} + y_{is}^{l}y_{jr}^{l'}\exp(-g_{ijsr}^{ll'}(z_s, z_r, \alpha))(-(1 + g_{ijsr}^{ll'}(z_s, z_r, \alpha)) + \Delta\kappa^{l'l} + \alpha\Delta w_{jr}^{l'l})|\mathcal{W}_{is}, \mathcal{W}_{jr}] \geqslant 0. \tag{A.14}$$

As the moments in equations (A.12) and (A.14) have the same conditioning set, we can add them, obtaining the following moment inequality:

$$\begin{gathered}\mathbb{E}[y_{is}^{l}y_{jr}^{l} + y_{is}^{l'}y_{jr}^{l'} + y_{is}^{l}y_{jr}^{l'}\exp(-g_{ijsr}^{ll'}(z_s, z_r, \alpha))\times \\ (-(1 + g_{ijsr}^{ll'}(z_s, z_r, \alpha)) + \alpha(\Delta w_{is}^{ll'} + \Delta w_{is}^{l'l}))|\mathcal{W}_{is}, \mathcal{W}_{jr}] \geqslant 0.\end{gathered}$$

Finally, if $z_s \subset \mathcal{W}_{is}$ and $z_r \subset \mathcal{W}_{jr}$, we can use the LIE and conclude that

$$\begin{gathered}\mathbb{E}[y_{is}^{l}y_{jr}^{l} + y_{is}^{l'}y_{jr}^{l'} + y_{is}^{l}y_{jr}^{l'}\exp(-g_{ijsr}^{ll'}(z_s, z_r, \alpha))\times \\ (-(1 + g_{ijsr}^{ll'}(z_s, z_r, \alpha)) + \alpha(\Delta w_{is}^{ll'} + \Delta w_{is}^{l'l}))|z_s, z_r] \geqslant 0.\end{gathered} \tag{A.15}$$

The moment in this inequality is the moment in equation (16) when evaluated at $\theta_\alpha = \alpha$. Thus, equation (A.15) implies Theorem 3. ■

Online Appendix for "Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach"

Charly Porcher
Georgetown University

Eduardo Morales
Princeton University

Thomas Fujiwara
Princeton University

April 2024

B Other Proofs and Additional Derivations

In Appendix B.1, we show how we obtain several expressions appearing in Section 3.1.1. In Appendix B.2, we prove Corollary 1. In Appendix B.3, we show how we obtain several expressions appearing in Section 3.1.2. In Appendix B.4, we prove Corollary 2. In Appendix B.5, we show how we obtain several expressions appearing in Section 3.2. In Appendix B.6, we prove Corollary 3.

B.1 Second-Step Bounding Inequalities: Additional Derivations

Derivation of equation (9). Given $z_s \in \mathcal{Z}_s$, we compute the function in equation (9) by finding the value of $h^{ll'}_{is}(z_s, \Delta\theta_{ll})$ that minimizes the moment in equation (7) at each value of $\Delta\theta_{ll'}$. Specifically, given z_s and $\Delta\theta_{ll'}$ the first-order condition of the moment in equation (7) with respect to the scalar $h^{ll'}_{is}(z_s, \Delta\theta_{ll})$ is

$$\mathbb{E}[y^l_{is}(h^{ll'}_{is}(z_s, \Delta\theta_{ll'}) - (\Delta\theta_{ll'} + \alpha\Delta w^{ll'}_s))|z_s],$$

or, equivalently,

$$\mathbb{E}[h^{ll'}_{is}(z_s, \Delta\theta_{ll'}) - (\Delta\theta_{ll'} + \alpha\Delta w^{ll'}_s)|z_s, y^l_{is} = 1].$$

Setting this moment condition to zero and bearing in mind that, according to equation (4), $\alpha \subset \mathcal{J}_{is}$, we can solve for $h^{ll'}_{is}(z_s, \Delta\theta_{ll'})$ to obtain the solution in equation (9); i.e.,

$$h^{ll'}_{is}(z_s, \Delta\theta_{ll'}) = \Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w^{ll'}_s|z_s, y^l_{is} = 1]. \tag{B.1}$$

Derivation of equation (10). Equations (7) to (9) imply the following inequality

$$\mathbb{E}[y^{l'}_{is} - y^l_{is}\exp(-(\Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w^{ll'}_s|z_s, y^l_{is} = 1]))(1 - \alpha(\Delta w^{ll'}_s - \mathbb{E}[\Delta w^{ll'}_s|z_s, y^l_{is} = 1]))|z_s] \geqslant 0,$$

or, equivalently,

$$\mathbb{E}[y_{is}^{l'}|z_s] \geqslant \mathbb{E}[y_{is}^{l} \exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]))(1-\alpha(\Delta w_s^{ll'}-\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]))|z_s].$$

We can simplify this inequality as

$$\mathbb{E}[y_{is}^{l'}|z_s] \geqslant \exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]))\mathbb{E}[y_{is}^{l}(1-\alpha(\Delta w_s^{ll'}-\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]))|z_s],$$

or, equivalently,

$$\begin{aligned}\mathbb{E}[y_{is}^{l'}|z_s] \geqslant{}& \exp(-(\Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1])) \\ &\times (\mathbb{E}[y_{is}^l|z_s] - \alpha\mathbb{E}[y_{is}^l(\Delta w_s^{ll'} - \mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1])|z_s]).\end{aligned}$$

We can further rewrite this inequality as

$$\begin{aligned}\mathbb{E}[y_{is}^{l'}|z_s] \geqslant{}& \exp(-(\Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1])) \\ &\times (\mathbb{E}[y_{is}^l|z_s] - \alpha(\mathbb{E}[y_{is}^l\Delta w_s^{ll'}|z_s] - \mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]\mathbb{E}[y_{is}^l|z_s])),\end{aligned}$$

or equivalently,

$$\begin{aligned}\mathbb{E}[y_{is}^{l'}|z_s] \geqslant{}& \exp(-(\Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1])) \\ &\times (\mathbb{E}[y_{is}^l|z_s] - \alpha(\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]\mathbb{E}[y_{is}^l|z_s] - \mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]\mathbb{E}[y_{is}^l|z_s])),\end{aligned}$$

Eliminating terms that cancel each other, we obtain the inequality

$$\mathbb{E}[y_{is}^{l'}|z_s] \geqslant \exp(-(\Delta\theta_{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]))\mathbb{E}[y_{is}^l|z_s].$$

Rearranging terms, we obtain the expression in equation (10); i.e.,

$$\frac{\mathbb{E}[y_{is}^l|z_s]}{\mathbb{E}[y_{is}^{l'}|z_s]} \exp(-\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^l = 1]) \leqslant \exp(\Delta\theta_{ll'}).$$

Derivation of equation (11). Swapping the indices l and l' in equation (10) we obtain the following inequality

$$\frac{\mathbb{E}[y_{is}^{l'}|z_s]}{\mathbb{E}[y_{is}^{l}|z_s]} \exp(-\alpha\mathbb{E}[\Delta w_s^{l'l}|z_s, y_{is}^{l'} = 1]) \leqslant \exp(\Delta\theta_{l'l}).$$

Rearranging terms, we immediately obtain the inequality in equation (11); i.e.,

$$\frac{\mathbb{E}[y_{is}^{l}|z_s]}{\mathbb{E}[y_{is}^{l'}|z_s]}\exp(-\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_{is}^{l'}=1]) \geqslant \exp(\Delta\theta_{ll'}).$$

B.2 Second-Step Bounding Inequalities: Proof of Corollary 1

Equations (7) to (9) imply the following moment inequality:

$$\mathbb{E}[y_{is}^{l}\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_s^{l}=1]))-y_{is}^{l'}|z_s] \geqslant 0. \tag{B.2}$$

As Corollary 1 assumes $z_s \subset \mathcal{J}_{is}$, the definition $\mathcal{W}_{is} \equiv \mathcal{J}_{is}/\varepsilon_{is}$ implies $z_s \subseteq \mathcal{W}_{is}$. Using the LIE, we write

$$\mathbb{E}[\mathbb{E}[y_{is}^{l}\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_s^{l}=1]))-y_{is}^{l'}|\mathcal{W}_{is}]|z_s] \geqslant 0.$$

Given that $z_s \subseteq \mathcal{W}_{is}$, we can further rewrite

$$\mathbb{E}[\mathbb{E}[y_{is}^{l}|\mathcal{W}_{is}]\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s, y_s^{l}=1]))-\mathbb{E}[y_{is}^{l'}|\mathcal{W}_{is}]|z_s] \geqslant 0. \tag{B.3}$$

As y_s^{l} is a function of $(\mathcal{W}_{is}, \varepsilon_i)$, equation (5) implies $\mathbb{E}[\Delta w_s^{ll'}|z_s, \mathcal{W}_{is}, y_s^{l}] = \mathbb{E}[\Delta w_s^{ll'}|z_s, \mathcal{W}_{is}]$. Given that $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_{is}]$ according to Corollary 1, we rewrite equation (B.3) as

$$\mathbb{E}[\mathbb{E}[y_{is}^{l}|\mathcal{W}_{is}]\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s]))-\mathbb{E}[y_{is}^{l'}|\mathcal{W}_{is}]|z_s] \geqslant 0.$$

Given equation (6), we can further rewrite

$$\mathbb{E}\left[\frac{\exp(\Delta\kappa_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])}{\sum_{l''=1}^{L}\exp(\Delta\kappa_{l''l'}+\alpha\mathbb{E}[\Delta w_s^{l''l'}|\mathcal{W}_{is}])}\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|z_s]))-\mathbb{E}[y_{is}^{l'}|\mathcal{W}_{is}]|z_s\right] \geqslant 0.$$

Using a similar expression for the probability of choosing l' conditional on $\mathcal{W}_{is}$, we derive

$$\mathbb{E}\left[(\exp(\Delta\kappa_{ll'}-\Delta\theta_{ll'})-1)\left(\sum_{l''=1}^{L}\exp(\Delta\kappa_{l''l'}+\alpha\mathbb{E}[\Delta w_s^{l''l'}|\mathcal{W}_{is}])\right)^{-1}|z_s\right] \geqslant 0,$$

where we have used the assumption (imposed in Corollary 1) that $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]$. We can rewrite this inequality as

$$(\exp(\Delta\kappa_{ll'} - \Delta\theta_{ll'}) - 1)\mathbb{E}\left[\left(\sum_{l''=1}^{L} \exp(\Delta\kappa_{l''l'} + \alpha\mathbb{E}[\Delta w_s^{l''l'}|\mathcal{W}_{is}])\right)^{-1} |z_s\right] \geqslant 0.$$

The expectation in this inequality is always strictly positive. Thus, we rewrite it as

$$\exp(\Delta\kappa_{ll'} - \Delta\theta_{ll'}) - 1 \geqslant 0 \qquad \Leftrightarrow \qquad \Delta\kappa_{ll'} \geqslant \Delta\theta_{ll'}. \tag{B.4}$$

This inequality holds for any two locations l and l'. Thus, swapping the location indices l and l', we similarly obtain the following inequality:

$$\Delta\kappa_{l'l} \geqslant \Delta\theta_{l'l} \qquad \Leftrightarrow \qquad \Delta\kappa_{ll'} \leqslant \Delta\theta_{ll'}. \tag{B.5}$$

Combining the inequalities in equations (B.4) and (B.5), we obtain the following equality:

$$\Delta\kappa_{ll'} = \Delta\theta_{ll'}. \tag{B.6}$$

Thus, Corollary 1 holds. ■

B.3 Second-Step Odds-Based Inequalities: Additional Derivations

Derivation of equation (14). Equation (12) and (13) imply the following inequality

$$\mathbb{E}[y_{is}^{l} \exp(-(\Delta\theta_{ll'} + \alpha\Delta w_{is}^{ll'}))|z_s] \geqslant \mathbb{E}[y_{is}^{l'}|z_s].$$

We can rewrite this inequality as

$$\mathbb{E}[y_{is}^{l}|z_s]\mathbb{E}[\exp(-(\Delta\theta_{ll'} + \alpha\Delta w_{is}^{ll'}))|z_s, y_{is}^{l} = 1] \geqslant \mathbb{E}[y_{is}^{l'}|z_s],$$

or, equivalently,

$$\mathbb{E}[y_{is}^{l}|z_s] \exp(-\Delta\theta_{ll'})\mathbb{E}[\exp(-\alpha\Delta w_{is}^{ll'})|z_s, y_{is}^{l} = 1] \geqslant \mathbb{E}[y_{is}^{l'}|z_s].$$

Rearranging terms, we obtain the expression in equation (14); i.e.,

$$\frac{\mathbb{E}[y_{is}^{l}|z_s]}{\mathbb{E}[y_{is}^{l'}|z_s]}\mathbb{E}[\exp(-\alpha\Delta w_{is}^{ll'})|z_s, y_{is}^{l} = 1] \geqslant \exp(\Delta\theta_{ll'}).$$

Derivation of equation (15). Swapping the indices l and l' in equation (14), we obtain the following inequality

$$\frac{\mathbb{E}[y_{is}^{l'}|z_s]}{\mathbb{E}[y_{is}^{l}|z_s]}\mathbb{E}[\exp(-\alpha\Delta w_{is}^{l'l})|z_s, y_{is}^{l'}=1] \geqslant \exp(\Delta\theta_{l'l}).$$

Rearranging terms, we immediately obtain the inequality in equation (15); i.e.,

$$\frac{\mathbb{E}[y_{is}^{l}|z_s]}{\mathbb{E}[y_{is}^{l'}|z_s]}(\mathbb{E}[\exp(-\alpha\Delta w_{is}^{l'l})|z_s, y_{is}^{l'}=1])^{-1} \leqslant \exp(\Delta\theta_{ll'}). \tag{B.7}$$

B.4 Second-Step Odds-Based Inequalities: Proof of Corollary 2

Equations (12) and (13) imply the following moment inequality:

$$\mathbb{E}[y_{is}^{l}\exp(-(\Delta\theta_{ll'}+\alpha\Delta w_s^{ll'})) - y_{is}^{l'}|z_s] \geqslant 0. \tag{B.8}$$

The assumption that $z_s \subset \mathcal{J}_{is}$ in Corollary 2, and the definition of $\mathcal{W}_{is} \equiv \mathcal{J}_{is}/\varepsilon_{is}$, imply that $z_s \subseteq \mathcal{W}_{is}$. Using the LIE, we can then write

$$\mathbb{E}[\mathbb{E}[y_{is}^{l}\exp(-(\Delta\theta_{ll'}+\alpha\Delta w_s^{ll'})) - y_{is}^{l'}|\mathcal{W}_{is}]|z_s] \geqslant 0.$$

Given the assumption that $\Delta w_s^{ll'} = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]$ in Corollary 2, we can rewrite

$$\mathbb{E}[\mathbb{E}[y_{is}^{l}\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])) - y_{is}^{l'}|\mathcal{W}_{is}]|z_s] \geqslant 0,$$

or, equivalently,

$$\mathbb{E}[\mathbb{E}[y_{is}^{l}|\mathcal{W}_{is}]\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])) - \mathbb{E}[y_{is}^{l'}|\mathcal{W}_{is}]|z_s] \geqslant 0.$$

Given the expression for the probability of choosing l conditional on $\mathcal{W}_{is}$, we further rewrite

$$\mathbb{E}\left[\frac{\exp(\Delta\kappa_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])}{\sum_{l''=1}^{L}\exp(\Delta\kappa_{l''l'}+\alpha\mathbb{E}[\Delta w_s^{l''l'}|\mathcal{W}_{is}])}\exp(-(\Delta\theta_{ll'}+\alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])) - \mathbb{E}[y_{is}^{l'}|\mathcal{W}_{is}]|z_s\right] \geqslant 0.$$

Using a similar expression for the probability of choosing l' conditional on $\mathcal{W}_{is}$, we derive

$$\mathbb{E}\left[(\exp(\Delta\kappa_{ll'}-\Delta\theta_{ll'})-1)\left(\sum_{l''=1}^{L}\exp(\Delta\kappa_{l''l'}+\alpha\mathbb{E}[\Delta w_s^{l''l'}|\mathcal{W}_{is}])\right)^{-1}|z_s\right] \geqslant 0,$$

or, equivalently,

$$(\exp(\Delta\kappa_{ll'} - \Delta\theta_{ll'}) - 1)\mathbb{E}\left[\left(\sum_{l''=1}^{L}\exp(\Delta\kappa_{l''l'} + \alpha\mathbb{E}[\Delta w_s^{l''l'}|\mathcal{W}_{is}])\right)^{-1}|z_s\right] \geqslant 0. \tag{B.9}$$

The conditional expectation in this inequality will always be strictly positive. Thus, we can rewrite the inequality in equation (B.9) as

$$\exp(\Delta\kappa_{ll'} - \Delta\theta_{ll'}) - 1 \geqslant 0 \qquad \Leftrightarrow \qquad \Delta\kappa_{ll'} \geqslant \Delta\theta_{ll'}. \tag{B.10}$$

This inequality holds for any two locations l and l'. Thus, swapping the location indices l and l', we similarly obtain the following inequality:

$$\Delta\kappa_{l'l} \geqslant \Delta\theta_{l'l} \qquad \Leftrightarrow \qquad \Delta\kappa_{ll'} \leqslant \Delta\theta_{ll'}, \tag{B.11}$$

Combining the inequalities in equations (B.10) and (B.11), we obtain the following equality:

$$\Delta\kappa_{l'l} = \Delta\theta_{l'l}. \tag{B.12}$$

Thus, Corollary 2 holds. ■

B.5 First-Step Moment Inequalities: Additional Derivations

Derivation of equation (18)*.* Given values $z_s \in \mathcal{Z}_s$ and $z_r \in \mathcal{Z}_r$, we compute the function in equation (18) by finding the value of $g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha)$ that minimizes the moment in equation (16) at each value of θ_α. Specifically, given z_s, z_r, and θ_α, the first-order condition of the moment in equation (16) with respect to the scalar $g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha)$ is

$$\mathbb{E}[y_{is}^{l}y_{jr}^{l'}(2g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha) - \theta_\alpha(\Delta w_s^{ll'} + \Delta w_r^{l'l}))|z_s, z_r] = 0,$$

or, equivalently,

$$\mathbb{E}[2g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha) - \theta_\alpha(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^{l}y_{is'}^{l'} = 1].$$

Setting this moment condition to zero, we can solve for $g_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha)$ to obtain the solution in equation (18); i.e.,

$$g_{ijsr}^{ll'}(z_s, z_r, \Delta\theta_{ll'}) = \theta_\alpha\mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^{l}y_{jr}^{l'} = 1].$$

Derivation of equation (19). Equations (16) to (18) imply the following inequality

$$\mathbb{E}[y_{is}^l y_{jr}^l + y_{is}^{l'} y_{jr}^{l'} - y_{is}^l y_{jr}^{l'} \exp(-\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]) \times$$
$$(2 - \theta_\alpha((\Delta w_s^{ll'} + \Delta w_r^{l'l}) - \mathbb{E}[\Delta w_s^{ll'} + \Delta w_r^{l'l}|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]))|z_s, z_r] \geqslant 0,$$

or, equivalently,

$$\mathbb{E}[y_{is}^l y_{jr}^l + y_{is}^{l'} y_{jr}^{l'}|z_s, z_r] \geqslant \mathbb{E}[y_{is}^l y_{jr}^{l'} \exp(-\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]) \times$$
$$(2 - \theta_\alpha((\Delta w_s^{ll'} + \Delta w_r^{l'l}) - \mathbb{E}[\Delta w_s^{ll'} + \Delta w_r^{l'l}|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]))|z_s, z_r].$$

Using the LIE, we can rewrite this inequality as

$$\mathbb{E}[y_{is}^l y_{jr}^l + y_{is}^{l'} y_{jr}^{l'}|z_s, z_r] \geqslant \mathbb{E}[\exp(-\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]) \times$$
$$(2 - \theta_\alpha((\Delta w_s^{ll'} + \Delta w_r^{l'l}) - \mathbb{E}[\Delta w_s^{ll'} + \Delta w_r^{l'l}|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]))|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1] \times$$
$$\mathbb{E}[y_{is}^l y_{jr}^{l'}|z_s, z_r] \geqslant 0.$$

Simplifying this expression, we obtain

$$\mathbb{E}[y_{is}^l y_{jr}^l + y_{is}^{l'} y_{jr}^{l'}|z_s, z_r] \geqslant 2\mathbb{E}[\exp(-\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1])\mathbb{E}[y_{is}^l y_{jr}^{l'}|z_s, z_r].$$

Rearranging terms, we obtain the expression in equation (19); i.e.,

$$\frac{\mathbb{E}[y_{is}^l y_{jr}^{l'}|z_s, z_r]}{\mathbb{E}[0.5(y_{is}^l y_{jr}^l + y_{is}^{l'} y_{jr}^{l'})|z_s, z_r]} \leqslant \exp(\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]).$$

B.6 First-Step Moment Inequalities: Proof of Corollary 3

If $z_s \subset \mathcal{W}_{is}$ and $z_r \subset \mathcal{W}_{jr}$, we can use the LIE to rewrite equation (19) as

$$\frac{\mathbb{E}[\mathbb{E}[y_{is}^l y_{jr}^{l'}|\mathcal{W}_{is}, \mathcal{W}_{jr}]|z_s, z_r]}{\mathbb{E}[0.5(\mathbb{E}[y_{is}^l y_{jr}^l|\mathcal{W}_{is}, \mathcal{W}_{jr}] + \mathbb{E}[y_{is}^{l'} y_{jr}^{l'}|\mathcal{W}_{is}, \mathcal{W}_{jr}])|z_s, z_r]} \leqslant$$
$$\exp(\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]).$$

Equations (5) and (6) further imply that we can rewrite this inequality as

$$\frac{\mathbb{E}[\mathbb{E}[y_{is}^l|\mathcal{W}_{is}]\mathbb{E}[y_{jr}^{l'}|\mathcal{W}_{jr}]|z_s, z_r]}{\mathbb{E}[0.5(\mathbb{E}[y_{is}^l|\mathcal{W}_{is}]\mathbb{E}[y_{jr}^l|\mathcal{W}_{jr}] + \mathbb{E}[y_{is}^{l'}|\mathcal{W}_{is}]\mathbb{E}[y_{jr}^{l'}|\mathcal{W}_{jr}])|z_s, z_r]} \leqslant$$
$$\exp(\theta_\alpha \mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^l y_{jr}^{l'} = 1]). \tag{B.13}$$

Given equations (1) to (6), it holds that, for any $l_1 = 1, \dots, L$ and $l_2 = 1, \dots, L$, we can write

$$\mathbb{E}[y_{is}^{l_1}|\mathcal{W}_{is}] = \frac{\exp(\Delta\kappa^{l_1 l_2} + \alpha\mathbb{E}[\Delta w_s^{l_1 l_2}|\mathcal{W}_{is}])}{\sum_{l''=1}^{L}\exp(\Delta\kappa^{l'' l_2} + \alpha\mathbb{E}[\Delta w_s^{l'' l_2}|\mathcal{W}_{is}])},$$

and similarly for worker j of type r. We then rewrite the inequality in equation (B.13) as

$$\frac{\mathbb{E}[\exp(\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])\exp(\Delta\kappa^{l'l} + \alpha\mathbb{E}[\Delta w_r^{l'l}|\mathcal{W}_{jr}])|z_s, z_r]}{\mathbb{E}[0.5(\exp(\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]) + \exp(\Delta\kappa^{l'l} + \alpha\mathbb{E}[\Delta w_r^{l'l}|\mathcal{W}_{jr}]))|z_s, z_r]}$$
$$\leqslant \exp(\theta_\alpha\mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^{l}y_{jr}^{l'} = 1]).$$

Simplifying this expression, we obtain

$$\frac{\mathbb{E}[\exp(\alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}])\exp(\alpha\mathbb{E}[\Delta w_r^{l'l}|\mathcal{W}_{jr}])|z_s, z_r]}{\mathbb{E}[0.5(\exp(\Delta\kappa^{ll'} + \alpha\mathbb{E}[\Delta w_s^{ll'}|\mathcal{W}_{is}]) + \exp(\Delta\kappa^{l'l} + \alpha\mathbb{E}[\Delta w_r^{l'l}|\mathcal{W}_{jr}]))|z_s, z_r]}$$
$$\leqslant \exp(\theta_\alpha\mathbb{E}[0.5(\Delta w_s^{ll'} + \Delta w_r^{l'l})|z_s, z_r, y_{is}^{l}y_{jr}^{l'} = 1]).$$

If $\mathbb{E}[\Delta w_s^{ll'}|z_s] = \mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_{is}] = \Delta\bar{w}$, and $\mathbb{E}[\Delta w_r^{l'l}|z_r] = \mathbb{E}[\Delta w_r^{l'l}|\mathcal{J}_{ir}] = \Delta\bar{w}$, for a common constant $\Delta\bar{w} \in \mathbb{R}$ this inequality becomes

$$\frac{\exp(\alpha\Delta\bar{w})\exp(\alpha\Delta\bar{w})}{0.5(\exp(\Delta\kappa^{ll'} + \alpha\Delta\bar{w}) + \exp(\Delta\kappa^{l'l} + \alpha\Delta\bar{w}))} \leqslant \exp(\theta_\alpha\Delta\bar{w}).$$

If $\Delta\kappa^{ll'} = 0$, then it becomes

$$\frac{\exp(\alpha\Delta\bar{w})\exp(\alpha\Delta\bar{w})}{0.5(\exp(\alpha\Delta\bar{w}) + \exp(\alpha\Delta\bar{w}))} \leqslant \exp(\theta_\alpha\Delta\bar{w}), \quad \Leftrightarrow \quad \frac{\exp(\alpha\Delta\bar{w})\exp(\alpha\Delta\bar{w})}{\exp(\alpha\Delta\bar{w})} \leqslant \exp(\theta_\alpha\Delta\bar{w}),$$

and

$$\exp(\alpha\Delta\bar{w}) \leqslant \exp(\theta_\alpha\Delta\bar{w}).$$

Thus, two inequalities of this type, one with $\Delta\bar{w} > 0$ and the other one with $\Delta\bar{w} < 0$, will point identify α.

B.7 Using Inequalities for Estimation of the Wage Parameter

We describe here how we use the moment inequalities introduced in Section 3.2 to compute a confidence set for the wage parameter θ_α.

The moment in equation (16) depends on instrument vectors z_s and z_r. We construct z_s and z_r following equation (20). In our empirical application, we equate $\Delta z_s^{ll'}$ and $\Delta z_r^{ll'}$ to the one-year lag value of $\Delta w_s^{ll'}$ and $\Delta w_r^{ll'}$, respectively, and build vectors of instruments $\Delta x_s^{ll'}$ and $\Delta x_r^{ll'}$ for $q \in \{2, 4, 8, 16\}$.

The instruments $\Delta x_s^{ll'}$ and $\Delta x_r^{ll'}$ in equation (20) are weakly positive. Hence, Theorem 3 and the LIE imply that, for any locations l and l', worker i of type s, worker j of type r, and and deterministic function $g_{ijsr}^{ll'}\colon \mathcal{Z}_s \times \mathcal{Z}_r \times \Theta_\alpha \to \mathbb{R}$, the $q^2 \times 1$ vector of moment inequalities

$$\mathbb{E}[\mathbb{M}_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha, g_{ijsr}^{ll'}(\cdot))(\Delta x_s^{ll'} \otimes \Delta x_r^{ll'})] \geqslant 0, \tag{B.14}$$

holds for $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$ if $z_s^{ll'} \subset \mathcal{J}_{is}$ and $z_r^{ll'} \subset \mathcal{J}_{jr}$. The choice of the number of intervals q is consequential for the validity of the inequalities in equation (B.14). If $q = 2$, $\Delta x_s^{ll'}$ includes the following two elements:

$$\begin{aligned} \Delta x_{s,1}^{ll'} &= \mathbb{1}\{-\infty < \Delta z_s^{ll'} \leqslant \mathrm{med}(\Delta z_s^{ll'})\}|\Delta z_s^{ll'}|^d, \\ \Delta x_{s,2}^{ll'} &= \mathbb{1}\{\mathrm{med}(\Delta z_s^{ll'}) < \Delta z_s^{ll'} \leqslant \infty\}|\Delta z_s^{ll'}|^d; \end{aligned} \tag{B.15}$$

and $\Delta x_r^{ll'}$ includes the following two elements:

$$\begin{aligned} \Delta x_{r,1}^{ll'} &= \mathbb{1}\{-\infty < \Delta z_r^{ll'} \leqslant \mathrm{med}(\Delta z_r^{ll'})\}|\Delta z_r^{ll'}|^d, \\ \Delta x_{r,2}^{ll'} &= \mathbb{1}\{\mathrm{med}(\Delta z_r^{ll'}) < \Delta z_r^{ll'} \leqslant \infty\}|\Delta z_r^{ll'}|^d. \end{aligned} \tag{B.16}$$

Thus, if $q = 2$, the set of values of $\Delta\theta_{ll'}$ consistent with the inequalities in equation (B.14) includes $\Delta\kappa^{ll'}$ if, for locations l and l', worker i of type s knows whether the realized value of $\Delta z_s^{ll'}$ is above or below the median of the distribution of sector-wage differences across locations and sectors, and worker j of type r knows whether the realized value of $\Delta z_r^{ll'}$ is above or below the same median.

The setting described in Section 2 includes one observation per worker. The sample analogue of the moment inequality in equation (21) thus averages over only one observation. However, as this inequality is valid for every worker i of every type s, every worker j of every type r, and every pair of locations l and l', it holds that, for any function $g_{ijsr}^{ll'}\colon \mathcal{Z}_s \times \mathcal{Z}_r \times \Theta_\alpha \to \mathbb{R}$, the vector of moment inequalities

$$\sum_{s=1}^{S}\sum_{i=1}^{I_s}\sum_{r>s}\sum_{j=1}^{I_r} \mathbb{E}[\mathbb{M}_{ijsr}^{ll'}(z_s, z_r, \theta_\alpha, g_{ijsr}^{ll'}(\cdot))(\Delta x_s^{ll'} \otimes \Delta x_r^{ll'})] \geqslant 0, \tag{B.17}$$

is satisfied at $\Delta\theta_{ll'} = \Delta\kappa^{ll'}$ if $z_{s'}^{ll'} \subset \mathcal{J}_{is'}$ for every worker $i = 1, \ldots, I_{s'}$ of every type $s' = 1, \ldots, S$.

If the number of worker types S is small, it may be convenient to further aggregate the inequality in equation (B.17) across all location pairs (l, l'). However, it may be the case that the instrument vector $\Delta x_{s'}^{ll'}$ belongs to the information set of every worker only for a subset of location pairs (l, l'); e.g., only for urban locations. If this is the case, we may aggregate the inequality in equation (B.17) only over location pairs (l, l') that belong to some subset specified by the researcher.

Given any significance level, we compute a confidence interval for θ_α by applying the inference procedure in Andrews and Soares (2010) to the sample analogue of the moment inequalities in equation (B.17); see Appendix B.8 for details.

B.8 Inference Procedure: Andrews and Soares (2010)

We describe here our implementation of the asymptotic version of the Generalized Moment Selection (GMS) test described on page 135 of Andrews and Soares (2010). The content of this section follows closely that of Appendix A.7 in Dickstein and Morales (2018).

We base the construction of our confidence set for the true parameter on the modified method of moments (MMM) statistic. Denote by γ a generic parameter for which we want to compute a $1-\alpha$ confidence set. In our context, the parameter γ may equal either θ_α or θ_l for some location l. Assume we use K sample moment inequalities to compute a confidence set for γ and denote each of these inequalities by

$$\bar{m}_k(\gamma) \geqslant 0, \qquad k = 1, \ldots, K. \tag{B.18}$$

where, for each $k = 1, \ldots, K$,

$$\bar{m}_k(\gamma) \equiv \frac{1}{N} \sum_{c=1}^{C} \sum_{n=1}^{N_c} m_k(x_{nc}, \gamma), \tag{B.19}$$

and where observations are grouped into clusters $c = 1, \ldots, C$ and indexed by $n = 1, \ldots, N_c$ within each cluster c. The variable x_n is a generic vector of observed covariates. For example, in the context of the moment in equation (22), each of the clusters c in equation (B.19) corresponds to a sector s, and each observation n within a cluster c corresponds to a worker i within a sector s. In the context of the moment in equation (B.17), each cluster may correspond to a pair of sectors (s, r) and each observation n within a cluster may correspond to a tuple of individual indices and location indices (i, j, l, l'). Regardless of the definition of what an observation is, the variable N_c denotes the number of sample observations within a cluster c.

The MMM statistic is defined as

$$T(\gamma) = \sum_{k=1}^{K} \left(\min \left\{ \sqrt{N} \frac{\bar{m}_k(\gamma)}{\hat{\sigma}_k(\gamma)}, 0 \right\} \right)^2, \tag{B.20}$$

where $\hat{\sigma}_k(\gamma) = \sqrt{\hat{\sigma}_k^2(\gamma)}$ and

$$\hat{\sigma}_k^2(\gamma) = \frac{1}{N} \sum_{c=1}^{C} \left(\sum_{n=1}^{N_c} (m_k(x_{nc}, \gamma) - \bar{m}_k(\gamma)) \right)^2.$$

Given a set of K inequalities and a grid Γ_g covering the parameter space of γ, we implement the following steps to compute a confidence set for this parameter:

Step 1: choose a point $\gamma_p \in \Gamma_g$. Steps 2 to 8 test the null hypothesis that γ^* equals γ_p:

$$H_0 : \gamma^* = \gamma_p \qquad \text{vs.} \qquad H_0 : \gamma^* \neq \gamma_p.$$

Step 2: evaluate the MMM test statistic at γ_p**:**

$$T(\gamma_p) = \sum_{k=1}^{K} \left(\min \left\{ \sqrt{N} \frac{\bar{m}_k(\gamma_p)}{\hat{\sigma}_k(\gamma_p)}, 0 \right\} \right)^2. \tag{B.21}$$

Step 3: compute the correlation matrix of the moments evaluated at γ_p**:**

$$\hat{\Omega}(\gamma_p) = Diag^{-\frac{1}{2}}(\hat{\Sigma}(\gamma_p))\hat{\Sigma}(\gamma_p)Diag^{-\frac{1}{2}}(\hat{\Sigma}(\gamma_p)), \tag{B.22}$$

where $Diag(\hat{\Sigma}(\gamma_p))$ is the $L \times L$ diagonal matrix whose diagonal elements are equal to those of $\hat{\Sigma}(\gamma_p)$, $Diag^{-\frac{1}{2}}(\hat{\Sigma}(\gamma_p))$ is a matrix such that $Diag^{-\frac{1}{2}}(\hat{\Sigma}(\gamma_p))Diag^{-\frac{1}{2}}(\hat{\Sigma}(\gamma_p)) = Diag^{-1}(\hat{\Sigma}(\gamma_p))$ and

$$\hat{\Sigma}(\gamma_p) = \frac{1}{N} \sum_{c=1}^{C} \left(\sum_{n=1}^{N_c} (m(x_{nc}, \gamma_p) - \bar{m}(\gamma_p)) \right) \left(\sum_{i=1}^{N_c} (m(x_{ic}, \gamma_p) - \bar{m}(\gamma_p)) \right)', \tag{B.23}$$

where

$$m(x_{nc}, \gamma_p) = (m_1(x_{nc}, \gamma_p), \ldots, m_K(x_{nc}, \gamma_p))', \tag{B.24}$$

$$\bar{m}(\gamma_p) = (\bar{m}_1(\gamma_p), \ldots, \bar{m}_K(\gamma_p))'. \tag{B.25}$$

Step 4: simulate the asymptotic distribution of $T(\gamma_p)$**.** Take D draws from the multivariate normal distribution $\mathbb{N}(0_K, I_K)$ where 0_K is a vector of 0s of dimension K and

I_K is the identity matrix of dimension L. Denote each of these draws as ζ_d. Define the criterion function $T_d^{AA}(\gamma_p)$ as

$$T_d^{AA}(\gamma_p) = \sum_{k=1}^{K} \left\{ (\min\{[\hat{\Omega}^{\frac{1}{2}}(\gamma_p)\zeta_d]_k, 0\})^2 \times \mathbb{1}\{\sqrt{N}\frac{\bar{m}_k(\gamma_p)}{\hat{\sigma}_k(\gamma_p)} \leqslant \sqrt{\ln N}\} \right\}$$

where $[\hat{\Omega}^{\frac{1}{2}}(\gamma_p)\zeta_d]_k$ is the l-th element of the vector $\hat{\Omega}^{\frac{1}{2}}(\gamma_p)\zeta_r$.

Step 5: compute critical value. The critical value $\hat{c}^{AA}(\gamma_p, 1-\delta)$ is the $(1-\delta)$-quantile of the distribution of $T_d^{AA}(\gamma_p)$ across the D draws taken in the previous step.

Step 6: accept/reject γ_p. Include γ_p in the $(1-\delta)\%$ confidence set, $\hat{\Gamma}^{1-\delta}$, if $T(\gamma_p) \leqslant \hat{c}^{AA}(\gamma_p, 1-\delta)$.

Step 7: repeat steps 2 to 6 for every γ_p in the grid Γ_g.

Step 8: compare $\hat{\Gamma}^{1-\alpha}$ to Γ_g. If none of the points in $\hat{\Gamma}^{1-\alpha}$ are at the boundary of Γ_g, define $\hat{\Gamma}^{1-\alpha}$ as the 95% confidence set for γ^*. Otherwise, expand the limits of Θ_g and repeat steps 1 to 8.

C Additional Simulation Results

In Appendix C.1, we describe the moment inequalities we use to compute the confidence intervals in Table 1. In Appendix C.2, we explore the robustness of the results in Table 1 to the amenity terms κ^l differing across locations. In Appendix C.3, we explore alternative ways of building the moment inequalities. In Table C.3, we compare our two-step estimator to a one-step estimator that can be used in settings with small choice sets. In Table C.4, we study the performance of a different estimator from that described in Section 3 that may also be applied in settings with large choice sets. In Appendix C.6, we present estimates analogous to those in Table 1, but computed using a larger set of instruments. In Appendix C.4, we present figures that illustrate some of the results in Table 1.

C.1 Inequalities Used in Estimation in Table 1

First step: cases 1 to 4. To compute the confidence intervals for θ_α in cases 1 to 4 in Table 1, we use the following sample moment inequality for each pair of locations $l = \{1, 2, 3\}$ and $l' \neq l$:

$$\sum_{s=1}^{S} (y_s^l y_{r(s)}^l + y_s^{l'} y_{r(s)}^{l'} - y_s^l y_{r(s)}^{l'} \exp(-g_s^{ll'}(z_{2s}, z_{2r(s)}, \theta_\alpha))$$

$$(2 + 2g_s^{ll'}(z_{2s}, z_{2r(s)}, \theta_\alpha) - \theta_\alpha(\Delta w_s^{ll'} + \Delta w_{r(s)}^{l'l})))(\Delta x_s^{ll'} \otimes \Delta x_{r(s)}^{l'l}) \geqslant 0, \tag{C.1}$$

where $r(s)$ indexes the sector we match with sector s when computing the inequality in equation (C.1). For each $s = 1, \ldots, S$, we select $r(s)$ randomly among those that satisfy the restriction

$$\mathbb{1}\{|\hat{\mathbb{E}}[\Delta w_s^{ll'}|\Delta z_{2s}^{ll'}, y_s^l = 1] - \hat{\mathbb{E}}[\Delta w_{r(s)}^{l'l}|\Delta z_{2r(s)}^{l'l}, y_{r(s)}^{l'} = 1]| \leqslant \tau\} = 1, \tag{C.2}$$

with $\tau = 0.002$. For any s, $\hat{\mathbb{E}}[\Delta w_s^{ll'}|\Delta z_{2s}^{ll'}, y_s^l = 1]$ is the predicted value of $\Delta w_s^{ll'}$ computed using a linear regression of $\Delta w_s^{ll'}$ on $\Delta z_{2s}^{ll'}$ estimated on the subset of observations with $y_s^l = 1$. To understand the restriction in equation (C.2), one should notice that the moment inequality in equations (16) to (18) holds for any two sectors s and r. However, Corollary 3 indicates this inequality has desirable properties when combined with other conditions, sectors s and r satisfy

$$\mathbb{E}[\Delta w_s^{ll'}|\mathcal{J}_s] = \mathbb{E}[\Delta w_r^{l'l}|\mathcal{J}_r]. \tag{C.3}$$

We cannot directly impose this restriction as $\mathcal{J}_s$ and $\mathcal{J}_r$ are unobserved. However, as a feasible alternative, we impose the restriction in equation (C.2). In Appendix C.2, we show that the confidence interval for θ_α we obtain when using the moment inequality in equations (C.1) and (C.2) indeed increases as we increase the value of τ in equation (C.2).

To complete the description of the inequality in equation (C.1), we must determine the function $g_s^{ll'}(\cdot)$ and the instruments $\Delta x_s^{ll'}$ and $\Delta x_{r(s)}^{l'l}$. Building on the expression in equation (18), we impose

$$g_s^{ll'}(z_{2s}, z_{2r(s)}, \theta_\alpha) = \theta_\alpha 0.5(\hat{\mathbb{E}}[\Delta w_s^{ll'}|\Delta z_{2s}^{ll'}, y_s^l = 1] + \hat{\mathbb{E}}[\Delta w_{r(s)}^{l'l}|\Delta z_{2r(s)}^{l'l}, y_{r(s)}^{l'} = 1]). \tag{C.4}$$

The instrument vectors $\Delta x_s^{ll'}$ and $\Delta x_{r(s)}^{l'l}$ are defined as follows

$$\Delta x_s^{ll'}(\equiv \mathbb{1}\{-\infty < \Delta z_{2s}^{ll'} \leqslant 0\}, \mathbb{1}\{0 < \Delta z_{2s}^{ll'} \leqslant \infty\})', \tag{C.5a}$$

$$\Delta x_{r(s)}^{l'l} \equiv (\mathbb{1}\{-\infty < \Delta z_{2r(s)}^{l'l} \leqslant 0\}, \mathbb{1}\{0 < \Delta z_{2r(s)}^{l'l} \leqslant \infty\})'. \tag{C.5b}$$

First step: case 5. To compute the confidence interval for θ_α for case 5 in Table 1, we use the following sample moment inequality for each pair of locations $l = \{1, 2, 3\}$ and $l' \neq l$:

$$\sum_{s=1}^{S}(y_s^l y_{r(s)}^l + y_s^{l'} y_{r(s)}^{l'} - y_s^l y_{r(s)}^{l'} \exp(-g_s^{ll'}(w_s, w_{r(s)}, \theta_\alpha))$$

$$(2 + 2g_s^{ll'}(w_s, w_{r(s)}, \theta_\alpha) - \theta_\alpha(\Delta w_s^{ll'} + \Delta w_{r(s)}^{l'l})))(\Delta x_s^{ll'} \otimes \Delta x_{r(s)}^{l'l}) \geqslant 0, \tag{C.6}$$

where, for each $s = 1, \ldots, S$, the sector $r(s)$ is selected randomly among those that verify

$$\mathbb{1}\{|\Delta w_s^{ll'} - \Delta w_{r(s)}^{l'l}| \leqslant 0.002\} = 1; \tag{C.7}$$

the function $g_s^{ll'}(\cdot)$ is determined as

$$g_s^{ll'}(w_s, w_{r(s)}, \theta_\alpha) = \theta_\alpha 0.5(\Delta w_s^{ll'} + \Delta w_{r(s)}^{l'l}); \tag{C.8}$$

and the instrument vectors $\Delta x_s^{ll'}$ and $\Delta x_{r(s)}^{l'l}$ are defined as follows

$$\Delta x_s^{ll'} \equiv (\mathbb{1}\{-\infty < \Delta w_s^{ll'} \leqslant 0\}, \mathbb{1}\{0 < \Delta w_s^{ll'} \leqslant \infty\})', \tag{C.9a}$$
$$\Delta x_{r(s)}^{ll'} \equiv (\mathbb{1}\{-\infty < \Delta w_{r(s)}^{l'l} \leqslant 0\}, \mathbb{1}\{0 < \Delta w_{r(s)}^{l'l} \leqslant \infty\})'. \tag{C.9b}$$

Second step: bounding inequalities for cases 1 to 4. Given a value $\check{\theta}_\alpha$ of the parameter θ_α and a location $l = 2, \ldots, L$, we use the following bounding moment inequalities to compute the confidence interval for θ_l in cases 1 to 4 in Table 1:

$$\sum_{s=1}^{S}(y_s^1 - y_s^l \exp(-h_s^{l1}(z_{2s}, \theta_l))(1 + h_s^{l1}(z_{2s}, \theta_l) - (\theta_l + \check{\theta}_\alpha \Delta w_s^{l1})))\Delta x_s^{l1} \geqslant 0, \tag{C.10a}$$
$$\sum_{s=1}^{S}(y_s^l - y_s^1 \exp(-h_s^{1l}(z_{2s}, -\theta_l))(1 + h_s^{1l}(z_{2s}, -\theta_l) + (\theta_l + \check{\theta}_\alpha \Delta w_s^{l1})))\Delta x_s^{1l} \geqslant 0; \tag{C.10b}$$

with the instrument vectors Δx_s^{l1} and Δx_s^{1l} defined as

$$\Delta x_s^{l1} \equiv (\mathbb{1}\{-\infty < \Delta z_{2s}^{l1} \leqslant 0\}, \mathbb{1}\{0 < \Delta z_{2s}^{l1} \leqslant \infty\})', \tag{C.11a}$$
$$\Delta x_s^{1l} \equiv (\mathbb{1}\{-\infty < \Delta z_{2s}^{1l} \leqslant 0\}, \mathbb{1}\{0 < \Delta z_{2s}^{1l} \leqslant \infty\})'; \tag{C.11b}$$

and

$$h_s^{l1}(z_{2s}, \theta_l) = \theta_l + \check{\theta}_\alpha \hat{\mathbb{E}}[\Delta w_s^{l1} | \Delta z_{2s}^{l1}, y_{is}^l = 1], \tag{C.12a}$$
$$h_s^{1l}(z_{2s}, -\theta_l) = -\theta_l + \check{\theta}_\alpha \hat{\mathbb{E}}[\Delta w_s^{1l} | \Delta z_{2s}^{1l}, y_{is}^1 = 1]. \tag{C.12b}$$

For any s and locations l and l', $\hat{\mathbb{E}}[\Delta w_s^{ll'} | \Delta z_{2s}^{ll'}, y_s^l = 1]$ is the predicted value of $\Delta w_s^{ll'}$ computed using a linear regression of $\Delta w_{s'}^{ll'}$ on $\Delta z_{2s'}^{ll'}$ estimated on the subset of observations with $y_s^l = 1$.

Second step: bounding inequalities for case 5. Given a value $\check{\theta}_\alpha$ of θ_α and a location $l \neq 1$, we use the inequalities below to compute the confidence interval for θ_l in case 5 in Table 1:

$$\sum_{s=1}^{S}(y_s^1 - y_s^l \exp(-h_s^{l1}(w_s, \theta_l))(1 + h_s^{l1}(w_s, \theta_l) - (\theta_l + \check{\theta}_\alpha \Delta w_s^{l1})))\Delta x_s^{l1} \geqslant 0, \tag{C.13a}$$

$$\sum_{s=1}^{S}(y_s^l - y_s^1 \exp(-h_s^{1l}(w_s, \theta_l))(1 + h_s^{1l}(w_s, \theta_l) + (\theta_l + \check{\theta}_\alpha \Delta w_s^{l1})))\Delta x_s^{1l} \geqslant 0; \tag{C.13b}$$

with the instrument vectors Δx_s^{l1} and Δx_s^{1l} defined as

$$\Delta x_s^{l1} \equiv (\mathbb{1}\{-\infty < \Delta w_s^{l1} \leqslant 0\}, \mathbb{1}\{0 < \Delta w_s^{l1} \leqslant \infty\})', \tag{C.14a}$$

$$\Delta x_s^{1l} \equiv (\mathbb{1}\{-\infty < \Delta w_s^{1l} \leqslant 0\}, \mathbb{1}\{0 < \Delta w_s^{1l} \leqslant \infty\})'; \tag{C.14b}$$

and

$$h_s^{l1}(w_s, \theta_l) = \theta_l + \check{\theta}_\alpha \Delta w_s^{l1}, \tag{C.15}$$

with $h_s^{1l}(w_s, -\theta_l) = -h_s^{l1}(w_s, \theta_l)$.

Second step: odds-based inequalities for cases 1 to 4. Given a value $\check{\theta}_\alpha$ of the parameter θ_α and a location $l = 2, \ldots, L$, we use the following odds-based inequalities to compute the confidence interval for θ_l in cases 1 to 4 in Table 1:

$$\sum_{s=1}^{S}(y_s^l \exp(-(\theta_l + \check{\theta}_\alpha \Delta w_s^{l1})) - y_s^1)\Delta x_s^{l1} \geqslant 0, \tag{C.16a}$$

$$\sum_{s=1}^{S}(y_s^1 \exp(\theta_l + \check{\theta}_\alpha \Delta w_s^{l1}) - y_s^l)\Delta x_s^{1l} \geqslant 0; \tag{C.16b}$$

with the instrument vectors Δx_s^{l1} and Δx_s^{1l} defined as in equation (C.11).

Second step: odds-based inequalities for case 5. Given a value $\check{\theta}_\alpha$ of the parameter θ_α and a location $l \neq 1$, we compute the confidence interval for θ_l in case 5 in Table 1 using inequalities analogous to those in equation (C.16), but with Δx_s^{l1} and Δx_s^{1l} defined as in equation (C.14).

C.2 Amenity Differences Across All Locations

Table C.1: Simulation Results - Confidence Intervals With Amenity Differences

Case	σ_1	σ_3	z_s	$(\kappa_1, \kappa_2, \kappa_3)$	1st Step
					α
2	0	1	z_{2s}	$(0, 0, 1)$	[1 , 1.01]
2	0	1	z_{2s}	$(0, 0.5, 1)$	[0.92 , 1.09]
2	0	1	z_{2s}	$(0, 0, 2)$	[1 , 1.02]
2	0	1	z_{2s}	$(0, 1, 2)$	[0.73 , 1.29]
2	0	1	z_{2s}	$(0, 0, 3)$	[1 , 1.02]
2	0	1	z_{2s}	$(0, 1.5, 3)$	[0.55 , 1.49]

The column α contains a 95% confidence interval for θ_α based on the moment inequality estimator introduced in Section 3.2 and described in detail in Appendix C.1. The column $(\kappa_1, \kappa_2, \kappa_3)$ displays the true value of the location amenities in the simulated data. Confidence intervals are computed following the procedure in Andrews and Soares (2010) using a 1-dimensional grid with limits $[0.5, 1.5]$.

C.3 First-step Moment Inequalities with Loose Sectoral Matches

Table C.2: Simulation Results - Moment Inequality Confidence Intervals With Loose Matches

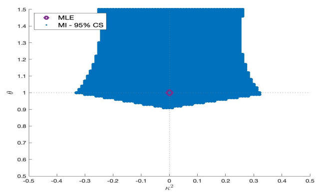
Case	σ_1	σ_3	z_s	τ	1st Step
					α
2	0	1	z_{2s}	8	[0.73 , 1.32]
2	0	1	z_{2s}	4	[0.79 , 1.25]
2	0	1	z_{2s}	2	[0.94 , 1.08]
2	0	1	z_{2s}	1	[0.98 , 1.03]
2	0	1	z_{2s}	0.8	[0.99 , 1.03]
2	0	1	z_{2s}	0.08	[1 , 1.02]
2	0	1	z_{2s}	0.008	[1 , 1.02]
2	0	1	z_{2s}	0.002	[1 , 1.01]

The column α contains a 95% confidence interval for θ_α based on the moment inequality estimator introduced in Section 3.2 and described in detail in Appendix C.1. The column τ displays the maximum feasible distance between $\mathbb{E}[\Delta w_s^{ll'}|z_s]$ and $\mathbb{E}[\Delta w_r^{l'l}|z_r]$ in each moment inequality. Confidence sets are computed using a 1-dimensional grid whose limits are $[0.5, 1.5]$. All confidence sets are computed following the procedure in Andrews and Soares (2010). The case with $\tau = 0.002$ corresponds to the baseline case reported in Table 1.

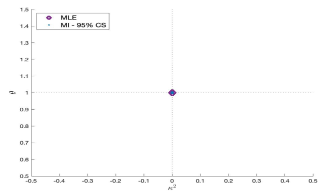
C.4 Combining Bounding and Odds-based Moment Inequalities

Figure C.1: Case 2 – $\sigma_1 = 0$ and $\sigma_3 = 1$

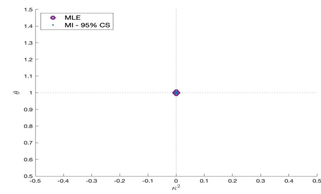
(a) Projected Confidence Set for (κ_2, θ)



Odds-based Ineq.

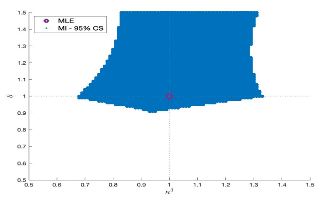


Bounding Ineq.

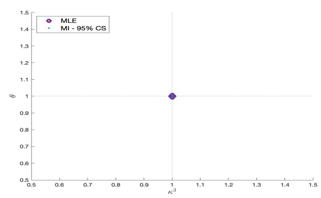


Both Types

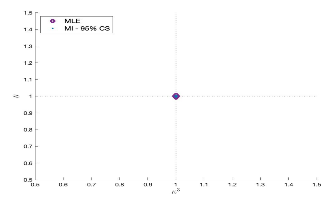
(b) Projected Confidence Set for (κ_3, θ)



Odds-based Ineq.



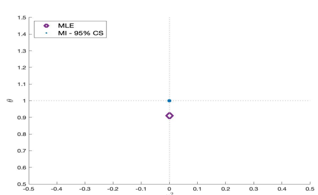
Bounding Ineq.



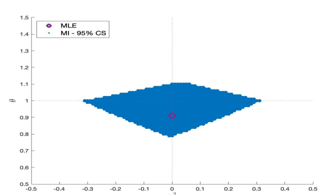
Both Types

Figure C.2: Case 3 – $\sigma_1 = 1$ and $\sigma_3 = 0$

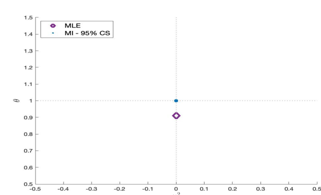
(a) Projected Confidence Set for (κ_2, θ)



Odds-based Ineq.

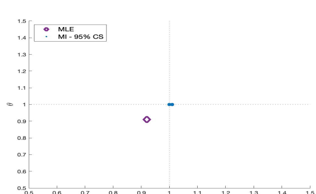


Bounding Ineq.

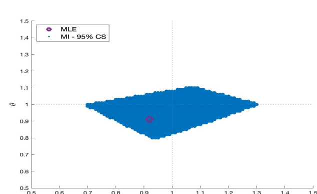


Both Types

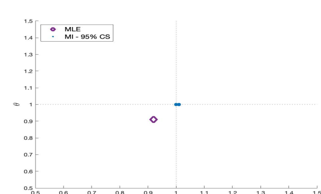
(b) Projected Confidence Set for (κ_3, θ)



Odds-based Ineq.



Bounding Ineq.



Both Types

C.5 Alternative Moment Inequality Estimators

One-step moment inequality estimator when the choice set is small. The key contribution of the two-step estimator described in Section 3 is to yield confidence intervals for all parameters in discrete-choice settings with many choice-specific fixed effects, as relevant in the migration context. However, when the choice set is small, as in our simulation setting, the inequalities in equations (8) and (13) in the second step of our estimator can be used alone to identify all parameters in one step. This is the approach in Dickstein et al. (2023). The wage coefficient is treated as an unknown parameter, and the bounding and the odds-based inequalities in Section 3.1 are used jointly to estimate a confidence set for the parameter vector $(\theta_\alpha, \theta_2, \theta_3)$. We report in Table C.3 the projection on each parameter of the three-dimensional confidence set computed following that procedure. Tables 1 and C.3 show that our two-step procedure yields similar confidence sets for the amenity parameters θ_2 and θ_3, but larger intervals for θ_α when agents' information sets are partly unobserved by the researcher (i.e., when $\sigma_1 > 0$). This procedure is, however, not applicable in our empirical setting, given a large number of choice-specific fixed effects.

Table C.3: Simulation Results - Confidence Intervals from One-Step Estimator

Case	σ_1	σ_3	z_i	Mom. Ineq.	α	κ_2	κ_3
1	0	0	z_{2s}	Bounding	[1 , 1]	[0 , 0]	[1 , 1]
				Odds-based	[1 , 1]	[0 , 0]	[1 , 1]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
2	0	1	z_{2s}	Bounding	[1 , 1]	[0 , 0]	[1 , 1]
				Odds-based*	[0.92 , 1.50]	[-0.33 , 0.33]	[0.67 , 1.33]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
3	1	0	z_{2s}	Bounding	[0.80 , 1.10]	[-0.30 , 0.30]	[0.70 , 1.30]
				Odds-based*	[1 , 1]	[0 , 0]	[1 , 1]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
4	1	1	z_{2s}	Bounding	[0.80 , 1.10]	[-0.30 , 0.30]	[0.70 , 1.30]
				Odds-based*	[0.92 , 1.50]	[-0.48 , 0.50]	[0.65 , 1.50]
				Both	[0.92 , 1.10]	[-0.33 , 0.30]	[0.70 , 1.30]
5	0	1	w_s	Bounding	[0.87 , 0.87]	[-0.05 , -0.03]	[0.85 , 0.88]
				Odds-based	$\varnothing$	$\varnothing$	$\varnothing$
				Both	$\varnothing$	$\varnothing$	$\varnothing$

This table contains projections of 95% confidence sets computed as in Andrews and Soares (2010) using a 3-dimensional grid with sides $[0.5, 1.5]$ (for α), $[-0.5, 0.5]$ (for κ_2) and $[0.5, 1.5]$ (for κ_3). We mark with an asterisk when the confidence set includes points outside the grid. Results in this table are taken from Dickstein et al. (2023).

Alternative moment inequality estimator when the choice set is large. In Table C.4, we apply an alternative inference procedure that only relies on the second-step inequalities described in Section 3.1. We compute confidence intervals for the vector $(\theta_\alpha, \theta_l)$ for each location $l = 2, \ldots, L$ using the inequalities in equations (8) and (13). We obtain $L-1$ two-dimensional confidence sets. After projecting them on each of their elements, we obtain a confidence interval for each amenity term $(\theta_2, \ldots, \theta_L)$ and $L-1$ intervals for θ_α. We can then report a randomly chosen confidence interval for θ_α among the $L-1$ available ones. This procedure yields results similar to those in Table 1, except in case 3, when adding information from the odds-based inequalities for the identification of θ_α results in a tighter confidence interval for this parameter. However, if information sets vary across destinations, the confidence interval for θ_α may be very sensitive to which of the $L-1$ available intervals is chosen.

Table C.4: Simulation Results - Alternative Confidence Intervals With Large Choice Set

Case	σ_1	σ_3	z_i	Mom. Ineq.	α	κ_2	κ_3
1	0	0	z_{2s}	Bounding	[1 , 1]	[0 , 0]	[1 , 1]
				Odds-based	[1 , 1]	[0 , 0]	[1 , 1]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
2	0	1	z_{2s}	Bounding	[1 , 1]	[0 , 0]	[1 , 1]
				Odds-based*	[0.91 , 1.50]	[-0.33 , 0.32]	[0.68 , 1.33]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
3	1	0	z_{2s}	Bounding	[0.80 , 1.10]	[-0.31 , 0.31]	[0.70 , 1.30]
				Odds-based*	[1 , 1]	[0 , 0]	[1 , 1.01]
				Both	[1 , 1]	[0 , 0]	[1 , 1.01]
4	1	1	z_{2s}	Bounding	[0.79 , 1.10]	[-0.31 , 0.31]	[0.69 , 1.30]
				Odds-based*	[0.92 , 1.50]	[-0.49 , 0.50]	[0.64 , 1.50]
				Both	[0.92 , 1.10]	[-0.31 , 0.31]	[0.69 , 1.31]
5	0	1	w_s	Bounding	[0.87 , 0.88]	[-0.05 , -0.01]	[0.86 , 0.88]
				Odds-based	$\varnothing$	$\varnothing$	$\varnothing$
				Both	$\varnothing$	$\varnothing$	$\varnothing$

This table contains projections of 95% confidence sets computed as in Andrews and Soares (2010) using a grids with sides $[0.5, 1.5]$ (for α), $[-0.5, 0.5]$ (for κ_2) and $[0.5, 1.5]$ (for κ_3). We mark with an asterisk when the confidence set includes points outside the grid.

C.6 Two-step Moment Inequalities with Additional Instruments

Table C.5: Simulation Results - Confidence Intervals With Additional Instruments

Case	σ_1	σ_3	z_s	First Step	Second Step		
				α	Mom. Ineq.	κ_2	κ_3
1	0	0	z_{2s}	[1 , 1.02]	Bounding	[0 , 0]	[1 , 1]
					Odds-based	[0 , 0]	[1 , 1]
					Both	[0 , 0]	[1 , 1]
2	0	1	z_{2s}	[1 , 1.01]	Bounding	[0 , 0]	[1 , 1]
					Odds-based	[-0.33 , 0.32]	[0.68 , 1.33]
					Both	[0 , 0]	[1 , 1]
3	1	0	z_{2s}	[0.91 , 1.15]	Bounding	[-0.31 , 0.31]	[0.70 , 1.30]
					Odds-based	[0 , 0]	[1 , 1.01]
					Both	[0 , 0]	[1 , 1.01]
4	1	1	z_{2s}	[0.91 , 1.19]	Bounding	[-0.31 , 0.31]	[0.70 , 1.30]
					Odds-based	[-0.32 , 0.32]	[0.68 , 1.33]
					Both	[-0.31 , 0.31]	[0.70 , 1.31]
5	0	1	w_s	$\varnothing$	Bounding	$\varnothing$	$\varnothing$
					Odds-based	$\varnothing$	$\varnothing$
					Both	$\varnothing$	$\varnothing$

The true parameter values are $\alpha = 1$, $\kappa^2 = 0$, and $\kappa^2 = 1$. The column α contains a 95% confidence interval for θ_α based on the estimator introduced in Section 3.2 and described in detail in Appendix C.1. The columns κ^2 and κ^3 contain 95% confidence intervals for θ_2 and θ_3 based on the estimators introduced in Section 3.1. The confidence intervals for θ_2 and θ_3 in the rows labeled *Bounding* use the inequalities introduced in Section 3.1.1; those in the row labeled *Odds-based* use the inequalities introduced in Section 3.1.2; and those in the row labeled *Both* combine both inequalities. Confidence sets are computed following the procedure in Andrews and Soares (2010) and using a 1-dimensional grid whose sides are $[0.5, 1.5]$ for α, $[-0.5, 0.5]$ for κ_2 and $[0.5, 1.5]$ for κ_3.

References

Andrews, Donald W.K. and Gustavo Soares, "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 2010, *78*, 119–157. [16, 19, 16, 19, 10, 16, 18, 19]

Dickstein, Michael J. and Eduardo Morales, "What do Exporters Know?," *The Quarterly Journal of Economics*, 2018, *133* (4), 1753–1801. [3, 6, 10, 21, 3, 6, 10, 21, 10]

_ , Jihye Jeon, and Eduardo Morales, "Patient Costs and Physicians' Information," *mimeo*, 2023. [3, 6, 10, 3, 6, 10, 18]

Supplementary Material for "Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach"

Charly Porcher
Georgetown University

Eduardo Morales
Princeton University

Thomas Fujiwara
Princeton University

April 2024

D Data and Summary Statistics

D.1 Data Sources and Sample Construction

The RAIS data. Our primary data source is the *Relação Anual de Informações Sociais* (RAIS), an administrative dataset maintained by Brazil's Ministry of Labor and Employment (local acronym MTE). It includes the universe of formal Brazilian employment spells in the private and public sectors. Individual workers are identified by their government-issued identification numbers (PIS/PASEP and CPF), allowing us to track them as they change employers (and employment location). For all employment spells observed between 1993 and 2011, we use information on their start and end dates, average monthly wage, number of work hours stipulated in the contract, 2-digit sector (industry) of production (according to the *Classificação Nacional de Atividades Econômicas*, CNAE), as well as information on the worker's gender, age, race, and level of education. All information on RAIS is reported by the employers.

RAIS only contains information on the formal employment of workers in Brazil. Thus, we have no information on the location of workers without formal jobs in a given year. These workers may be employed in the informal sector, self-employed, unemployed, or out of the labor force. Based on the 2010 Census, which directly asks respondents about their job status, 51% of the Brazilian labor force was in the formal sector. The implied total number of formal sector workers in the Census also closely matches the number of individual workers at RAIS.

Geography and wage definitions. To determine workers' location and migration decisions, we use the microregion of the *establishment* at which the worker is employed. This has the advantage of correctly locating workers in cases where a firm has several establishments in different locations. Microregions are groups of municipalities that span the entirety of the Brazilian territory. They are defined by the *Instituto Brasileiro de Geografia e Estatística* (IBGE). During our sample period, Brazil had 558 microregions. Microregions are also grouped into larger 136 mesoregions and are contained within 26 states and the federal district. While RAIS does not contain information on the residence of workers, Dix-Carneiro and Kovak (2017) use 2000 Census data to report that only 3.4% of individuals lived and worked in different microregions.

Previous research has used microregions as local labor markets (similar to commuting zones in the United States). For examples and further discussion, see Dix-Carneiro (2014), Dix-Carneiro and Kovak (2017), Dix-Carneiro and Kovak (2019), Felix (2022), and Szerman (2024).

Workers may hold multiple employment spells (jobs) in the same year. To obtain a dataset in which each unit of observation corresponds to a worker and a year, we assign to each worker-year pair the microregion and sector corresponding to the job that the worker held for the most extended period during that year. We compute the total labor income of a worker in a year by adding the labor income earned in every job this worker has been employed in the corresponding year. We calculate the total labor income of a worker in each of their jobs by transforming the average monthly wages reported in that job into a measure of average daily wages and multiplying it by the total number of days worked in the job reported in the data. If no start and end date is provided, we assume that these are January 1 and December 31, respectively.

Sample restrictions and sampling. We limit our data to workers between 25 and 64 years of age. This ensures we observe them after the vast majority of the population has completed their education and limits the age before a large share retires. However, we use the information from 1993-2001 to measure each worker's experience in each sector and microregion. To limit our data to workers with a sufficiently close labor relationship with the formal sector, we restrict our sample to workers observed at RAIS for at least seven years in the sample period.

Since we focus on studying heterogeneity in information sets due to workers' location, we restrict our sample to workers with similar demographic characteristics. Specifically, we focus on workers with at least a high school degree identified as male and white.

For computational reasons, our empirical application uses a sample of 10 million worker-year pairs. To ensure we observe a large enough number of individuals per market, we focus on 1,000 labor markets consisting of all combinations of the 50 microregions (out of 558) and 20 sectors (out of 51) with the largest total employment reported in RAIS. We then obtain our sample by randomly sampling 1 million individuals per year among those employed in the 1,000 labor markets of interest. Our sample period covers 10 years, from 2002-2011, and thus, our sample contains 10 million observations.

Additional data sources. Microregion population and demographic characteristics are from the 2000 and 2010 Censuses collected by the IBGE. We calculate distances between microregions by using the geodesic distance between their population centroids and data from the IBGE. Data on internet connections is from the *Agência Nacional de Telecomunicações* (ANATEL), which provides the number of broadband connections by municipality and year from 2007 onwards. We define internet access at the microregion level as the average share of households with broadband internet access in the 2007-2011 period.

D.2 Summary Statistics

D.2.1 Summary Statistics by Microregion and Sector

Table D.1 lists the 20 sectors in our sample and provides their average wage shifter (w_t^{sl} from equation E.1, described in Appendix E.1) for the sample period (2002-2011) and the share of the sample they represent. Wage shifters are measured in log units and are normalized so the sector with the lowest wage shifter has a value of zero. There is substantial variation in wage shifters across sectors.

Table D.2 lists the 50 microregions in our sample, and provides their average in-migration and out-migration rates (share of workers moving into and out of them between two consecutive years), as well as their average wage shifter in the sample period (2002-2011). Wage shifters are measured in log units and are normalized so the microregion with the lowest wage shifter has a value of zero. Wage shifters vary by microregion, sector, and year and are weighted by the number of workers in the sample in a microregion-sector-year cell. Thus wage shifter differences across microregions also reflect differences in their sectoral composition of employment. There is substantial variation in migration rates and wage shifters across microregions.

Table D.2 also provides the level of internet access for each microregion, their total population from the 2010 Census, and the share of the sample they represent. Note the population figure includes all persons residing in the microregion, while our sample only includes white male workers in the formal sector in a subset of sectors (industries). Thus, the correlation between a microregion population and its share of the sample is not equal to one.

D.2.2 Summary Statistics: Migration

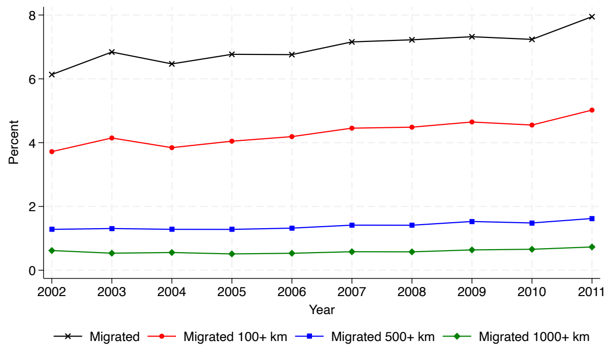
This subsection provides summary statistics describing key migration patterns. Given our sample definitions, it focuses on white male workers with at least a high school education in the 2002-2011 period. All figures are constructed using all workers with such characteristics (i.e., not only the sample focused on the 50 largest microregions and 20 largest sectors).

Figure D.1a provides migration rates (the share of workers that change microregion of employment between year t and $t-1$) by year. It shows an upward trend over the sample period, from close to 6% in 2002 to 8% in 2011. As a comparison, the overall migration rate in our sample in the same period is 6.2%. Figure D.1a also provides migration rates conditional on the distance between origin and destination microregions, indicating that about a third of moves are to microregions within 100 km from the origin, and less than a sixth of moves involve migration over a distance larger than 1,000 km. Figure D.1b provides a similar figure for changes in the employment sector, which are more common. It also provides the share of workers that change both microregion and sector of employment in a given year, indicating that most changes in the employment sector are not accompanied by migration.

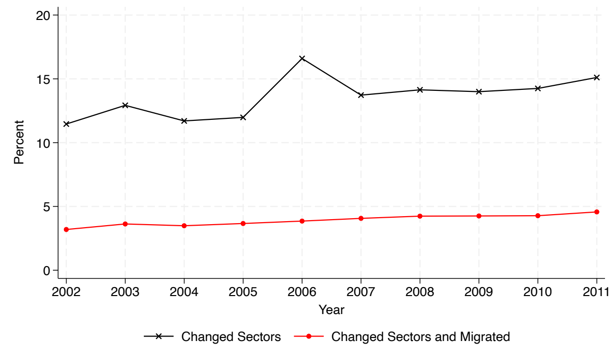
Figure D.2a depicts the distribution of distances between origin and destination microregions for those who migrate. Although more than half of moves occur between microregions within 200 km of distance, a sizable share of moves occur at larger distances, including substantial mass in distances over 500 km. Figure D.2b provides a scatter plot depicting the distribution of in-migration and out-migration rates for the 50 largest microregions which are the focus of our sample. Each marker represents one of these microregions, indicating its overall out- and in-migration rates (share of workers moving into and out of it between two consecutive years) in the sample period. In- and out-migration rates are strongly correlated across microregions, varying close to one-to-one. Note the figure also depicts the distribution of both variables. The bulk of microregions have migration rates in the 3 to 12% range with four others with migration rates of roughly 14% and two outliers that experience large migration flows. Figure D.3 provides similarly constructed figures exploring the relationship between migration rates and microregion size (measured as the number of workers observed in the data), indicating that their correlation is quite low.

Figure D.1: Migration and Sector Changes, by Year

(a) Migration Rates by Year



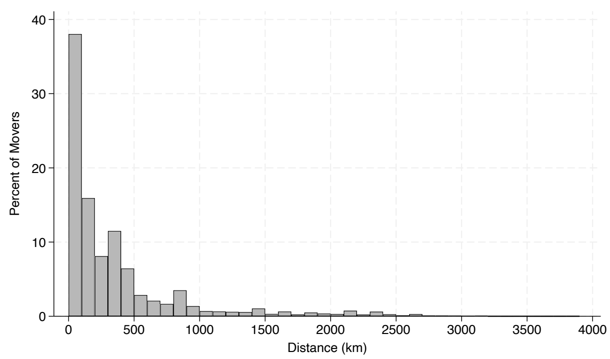
(b) Sector and Microregion Changes



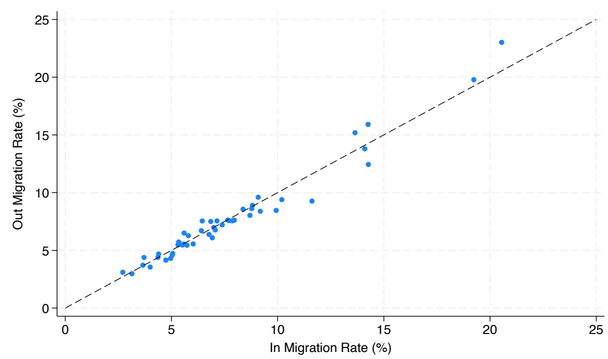
Panel (a) shows migration rates (share of individuals who changed microregion from the previous year) for each year. Migration rates are further refined by distance (only including workers that moved between microregions with a distance of 100km or more, 500km or more, and 1000km or more). Panel (b) shows the share of workers that changed sectors from the previous year (top line) and the share that changed both sector and migrated (bottom line). Data includes all white male workers with at least a high school education.

Figure D.2: Migration Patterns

(a) Histogram: Migration Distance



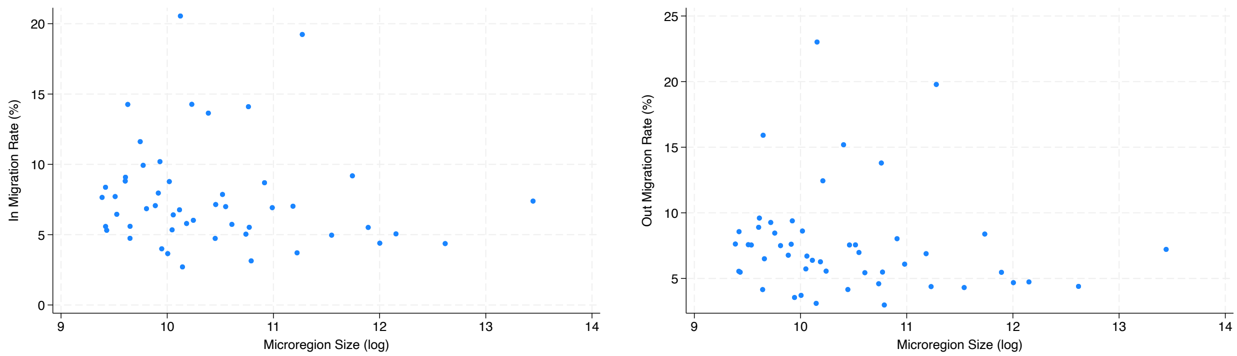
(b) Out- vs. In-Migration (sample)



Panel (a) provides a histogram of the distribution of distance between origin and destination of all observed migrations (changes in microregions between two consecutive years). In Panel (b), each marker represents one of the 50 microregions in our sample. The y-axis measures the out-migration rate (share of workers that move out of the microregion in a year), while the x-axis measures the in-migration (share of workers that move into the microregion). It considers migration with origins and destinations to all microregions (including outside the 50 largest ones). Dashed line represents the 45 degree line. In both panels, data includes all white male workers with at least a high school education in the 2002-2011 period.

Figure D.3: Migration and Microregion Size

(a) In-Migration Rates (b) Out-Migration Rates



Each marker in the figure represents one of the 50 microregions in our sample. The y-axis measures the in-migration and out-migration rates (share of workers that move into or out of a microregion between two consecutive years), respectively. The x-axis measures the log of average yearly number of workers. All variables are based on data including all white male workers with at least a high school degree in the 2002-2011 period.

Table D.1: Sector-level Summary Statistics

Sector Name	Sector Code	Wage Shifter (log)	Share of Sample (%)
Service Activities Mainly Provided to Businesses	74	0.140	20.52
Wholesale Trade in Goods and General Merchandise	51	0.055	8.28
Rail and Road Passengers Transportation	60	0.123	7.95
Construction Auxiliary Services and Installation Works	45	0.026	7.42
Motor Vehicle Parts and Accessories Manufacturing	34	0.207	6.28
Education and Teaching Activities	80	0.217	6.12
Household Appliance and Machinery Manufacturing	29	0.068	5.55
Food Processing and Manufacturing	15	0.120	4.74
Health Related Activities and Social Services	85	0.204	3.92
Metal Product Manufacturing	28	0.096	3.76
Software and Computer Development, Consulting	72	0.000	3.25
Plastics Product Manufacturing	25	0.110	3.08
Activities Related to the Organization of Freight Transport	63	0.027	2.99
Professional, Political Organizations and Trade Unions	91	0.258	2.91
Real Estate	70	0.007	2.77
Nonferrous Metal Foundries	27	0.268	2.76
Media Publishing, Printing, and Reproducing	22	0.354	2.14
Production and Distribution of Electricity, Energy	40	0.477	1.94
Performing, Arts, Sports and Leisure Activities	92	0.243	1.88
Electrical Machinery, and Supplies Manufacturing	31	0.072	1.75

This table presents the 20 sectors included in our sample, the share of the sample they represent, and their average wage shifter (w_t^{sl} from equation E.1) for the 2002-2011 period. Wage shifters are measured in log units and are normalized so the sector with the smallest wage shifter has a value equal to zero. Wage shifters vary by microregion, sector, and year and are weighted by the number of workers in the sample in a microregion-sector-year cell.

Table D.2: Microregion-level Summary Statistics

Microregion	Wage Shifter (log)	In-Migration Rate (%)	Out-Migration Rate (%)	Internet Access (%)	Population (1000s)	Share of Sample (%)
São Paulo - SP	0.365	6.63	6.43	17.81	13805	25.49
Rio de Janeiro - RJ	0.242	3.72	3.71	9.38	11601	9.47
Curitiba - PR	0.253	4.05	4.01	13.11	3060	6.26
Porto Alegre - RS	0.308	2.65	2.74	12.31	3628	5.71
Belo Horizonte - MG	0.258	3.69	3.82	10.00	4773	5.00
Campinas - SP	0.341	8.78	7.80	12.31	2631	4.23
Osasco - SP	0.235	17.71	18.99	10.01	1775	3.09
Brasília - DF	0.282	4.07	3.89	12.88	2570	2.24
São José dos Campos - SP	0.378	7.15	7.00	10.67	1415	2.23
Recife - PE	0.178	2.89	3.19	4.87	3259	2.21
Joinville - SC	0.085	3.20	3.01	11.26	843	2.05
Sorocaba - SP	0.323	7.38	6.70	8.51	1324	2.03
Guarulhos - SP	0.290	12.42	12.46	8.82	1347	1.85
Caxias do Sul - RS	0.213	2.57	2.20	8.59	770	1.47
Santos - SP	0.416	6.33	6.25	13.54	1471	1.46
Florianópolis - SC	0.136	6.16	5.31	16.70	878	1.37
Goiânia - GO	0.122	3.88	4.25	9.06	2117	1.36
Vitória - ES	0.203	4.32	4.07	8.70	1565	1.31
Fortaleza - CE	0.086	3.01	2.66	4.53	3351	1.21
Ribeirão Preto - SP	0.292	6.00	5.62	10.58	1033	1.19
Salvador - BA	0.240	7.69	7.70	6.04	3459	1.11
Moji das Cruzes - SP	0.228	14.89	17.42	6.55	1316	1.08
Jundiaí - SP	0.385	13.33	11.80	9.58	633	1.05
Itapecerica da Serra - SP	0.161	19.26	23.75	5.83	987	0.98
Londrina - PR	0.168	3.68	4.05	8.92	725	0.95
Piracicaba - SP	0.262	6.35	6.39	9.33	556	0.88
Blumenau - SC	0.061	4.94	4.59	10.21	677	0.80
Vale do Paraíba - RJ	0.338	5.85	6.19	4.99	680	0.79
Uberlândia - MG	0.228	4.35	4.52	9.18	820	0.74
São José do Rio Preto - SP	0.195	4.62	4.65	7.87	764	0.71
Limeira - SP	0.253	5.83	5.67	8.26	579	0.66
Bauru - SP	0.286	8.86	7.80	8.00	562	0.63
Maringá - PR	0.048	4.55	4.33	12.87	540	0.62
Bragança Paulista - SP	0.193	15.30	18.20	8.63	498	0.55
Itajaí - SC	0.084	7.36	6.37	10.23	571	0.53
Natal - RN	0.031	2.68	2.56	4.63	1031	0.52
Moji-Mirim - SP	0.255	7.68	7.51	7.98	383	0.52
Juiz de Fora - MG	0.253	4.52	5.52	6.35	729	0.52
Campo Grande - MS	0.148	4.03	4.98	9.55	874	0.52
Manaus - AM	0.284	4.27	3.43	2.66	2040	0.50
São Luís - MA	0.200	3.64	3.51	3.20	1309	0.48
Araraquara - SP	0.319	7.87	8.30	9.36	502	0.46
Ponta Grossa - PR	0.324	6.26	5.83	7.33	430	0.44
São Carlos - SP	0.379	5.05	5.51	9.68	309	0.44
Macaé - RJ	0.341	12.26	10.42	4.52	262	0.42
Passo Fundo - RS	0.270	2.72	3.14	14.85	328	0.41
João Pessoa - PB	0.023	3.12	3.09	5.55	1035	0.41
Cascavel - PR	0.109	4.72	4.89	7.76	433	0.39
Presidente Prudente - SP	0.327	4.52	5.78	6.27	573	0.36
Criciúma - SC	0.000	3.18	3.27	7.15	369	0.32

This table presents the 50 microregions included in our sample, the share of the sample they represent, their population in 2010, their share of households with broadband internet access, as well as their average in-migration rate, out-migration rate, and wage shifter (w_t^{sl} from equation E.1) for the 2002-2011 period. Wage shifters are measured in log units and are normalized so the microregion with the smallest wage shifter has a value equal to zero. Wage shifters vary by microregion, sector, and year and are weighted by the number of workers in the sample in a microregion-sector-year cell.

E Appendix to Empirical Analysis

E.1 Wage Proxies

In this section, we compute proxies for the log wages in all labor markets for all individuals in our sample. We consider a wage process that depends on an individual-by-sector-by-year component that workers consider as constant when deciding which location to move to. The relevant component of wages for their decision is then the location-by-sector-by-year component of log wages. The objective is to obtain a wage proxy that is as precise as possible after accounting for the individual and sector heterogeneity, in order to credibly interpret our proxies as local labor market demand shifters.

Workers are indexed by i. We assume that a worker's log wage w_{it}^{sl} can be expressed as the sum of a labor market-specific term w_t^{sl} (which we label the wage shifter), common to all agents, an individual skill for that sector that depends on a persistent match term and the number of years of age and experience in the sector, and an unexpected wage shock,

$$w_{it}^{sl} = w_t^{sl} + \underbrace{\alpha_i^s + \beta_e^s exp_{it}^s + \beta_{ee}^s (exp_{it}^s)^2 + \beta_a^s age_{it} + \beta_{aa}^s age_{it}^2}_{\text{sector-specific skill}} + \nu_{it}^{sl}, \tag{E.1}$$

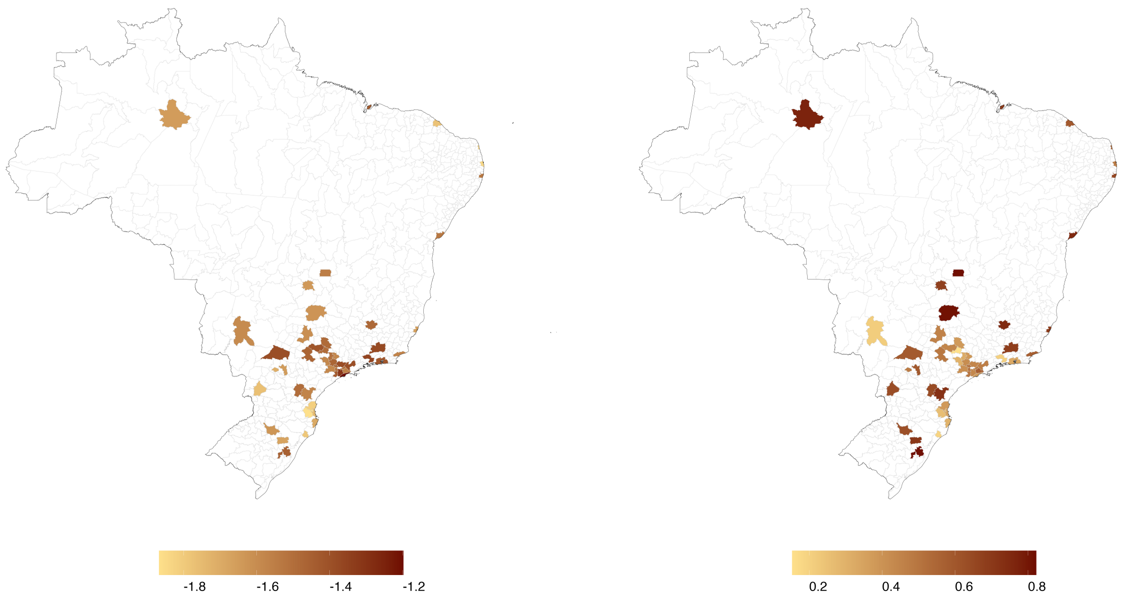
where we assume that the individual-sector skill is a function of age and age squared, and of the number of years of experience individual i has accumulated in sector s from 1993, the first year for which we collect individual employment information until the year they are observed. Hence, by construction, our measure of experience is capped at 8 years at the beginning of our analysis period in 2002 and at 17 years in 2011.

We estimate equation (E.1) on the universe of (white males with at least high school education) workers in the selected 1,000 labor markets over the 10 years of analysis that constitutes our sample (i.e., before sampling only 1 million individuals per year). This dataset consists of 15,313,848 observations. The mean value of experience across the 20 sectors and across the 10 years of our sample is 4.7 years, and the median is 3 years. The results from the estimation of equation (E.1) for each of the 20 sectors are reported in Table E.1. The median R^2 across the 20 regressions is 0.83, the median standard deviation of the individual-sector fixed effects is 1.20, and the median standard deviation of the labor market shifter is 0.48.

Figure E.1 provides maps depicting the geographic distribution of w_t^{sl} in the first and final year of our sample period. Since we observe different w_t^{sl} for each sector s in each microregion-year, the figure reports averages weighed by the number of workers in each microregion-sector-year cell.

Figure E.1: Geographic Distribution of Predicted Wages in Sample

(a) 2002 (b) 2011



The map presents the wage shifter (w_t^{sl} from equation E.1) for each microregion in the sample of the 50 largest microregions in the years 2002 and 2011. Wage shifters vary by microregion, sector, and year and are weighted by the number of workers in the sample in a microregion-sector-year cell. Microregions not included in the sample are shaded white.

E.2 Implementation of Moment Inequalities

Computing the bounding moments in equations (7) and (16) requires specifying the functions $h_{is}^{ll'}(\cdot)$ and $g_{ijsr}^{ll'}(\cdot)$, respectively. Equations (9) and (18) provide functional forms that yield the tightest identified sets. These expressions, however, depend on the expectation of specific wage differences conditional on the wage predictor used to build the corresponding inequality. Since we ignore the true value of those expectations, we use instead the following functions, for different guessed values α_m of α:

$$h_{ist}^{ll'}(z_s, \Delta\theta_{ll'}) = \Delta\theta_{ll'} + \alpha_m \Delta z_{st}^{ll'} \quad \text{and} \quad g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \theta_\alpha) = \theta_\alpha 0.5(\Delta z_{st}^{ll'} + \Delta z_{rt}^{l'l}). \tag{E.2}$$

Finally, the bounding moment inequality in equation (16) holds for any two locations l and l'. To avoid inference problems related to using many moment inequalities in our estimation procedure, we add these inequalities across location pairs. Our choice of which

Table E.1: Log Wage Proxies

Sector and Code		β_e^s	β_{ee}^s	β_{aa}^s	$cons$	R^2	N_{obs}	$\sigma_{\alpha_i^s}$	$\sigma_{w_t^{sl}}$
Business Services	74	0.365 (5.18e-4)	-1.16e-2 (2.27e-5)	-2.98e-5 (7.86e-6)	6.66 (0.011)	0.81	6,682,555	1.28	0.32
Wholesale Trade in Goods	51	0.420 (7.71e-4)	-1.18e-2 (3.17e-5)	-9.85e-5 (1.19e-5)	7.06 (0.017)	0.83	3,047,478	1.22	0.37
Rail and Road Transportation	60	0.414 (7.98e-4)	-1.06e-2 (2.68e-5)	7.28e-4 (1.12e-5)	5.34 (0.017)	0.78	2,952,869	1.36	0.48
Construction Services	45	0.319 (7.78e-4)	-8.94e-3 (3.27e-5)	-2.89e-5 (1.03e-5)	6.60 (0.016)	0.80	3,145,910	1.16	0.34
Motor Vehicle Manuf.	34	0.376 (1.04e-3)	-9.12e-2 (2.79e-5)	-1.49e-4 (1.35e-5)	7.68 (0.019)	0.84	1,827,610	0.94	0.65
Education	80	0.330 (9.24e-4)	-1.03e-2 (2.91e-5)	-8.64e-4 (1.03e-5)	9.34 (0.019)	0.86	2,092,113	1.58	0.94
Household Appliance Manuf.	29	0.417 (9.91e-4)	-1.01e-2 (3.08e-5)	-1.57e-4 (1.34e-5)	7.35 (0.020)	0.83	1,760,708	1.07	0.49
Food Processing and Manuf.	15	0.412 (9.02e-4)	-1.03e-2 (2.86e-5)	1.43e-4 (1.31e-5)	6.55 (0.017)	0.83	2,360,486	1.14	0.46
Health and Social Services	85	0.388 (1.24e-3)	-1.13e-2 (3.69e-5)	3.77e-4 (1.57e-5)	6.47 (0.024)	0.82	1,211,780	1.11	0.50
Metal Product Manuf.	28	0.434 (1.21e-3)	-1.08e-2 (4.23e-5)	1.07e-4 (1.72e-5)	6.69 (0.024)	0.82	1,184,199	1.15	0.46
Software and Consulting	72	0.426 (1.32e-3)	-1.19e-2 (5.11e-5)	-4.75e-6 (2.04e-5)	7.43 (0.028)	0.82	1,052,428	1.07	0.42
Plastics Product Manuf.	25	0.454 (1.41e-3)	-1.11e-2 (4.33e-5)	2.01e-4 (1.95e-5)	6.53 (0.027)	0.83	961,304	1.22	0.51
Freight Transport	63	0.420 (1.52e-3)	-9.86e-3 (5.01e-5)	1.13e-4 (2.02e-5)	6.76 (0.030)	0.83	937,390	1.23	0.45
Political Org. and Unions	91	0.414 (1.51e-3)	-1.26e-2 (4.60e-5)	-5.38e-4 (1.69e-5)	8.09 (0.028)	0.88	799,723	1.28	0.65
Real Estate	70	0.461 (1.48e-3)	-1.26e-2 (4.97e-5)	2.41e-4 (1.84e-5)	6.07 (0.029)	0.81	839,026	1.20	0.42
Nonferrous Metal Foundries	27	0.411 (1.53e-3)	-1.01e-2 (3.65e-5)	-5.46e-5 (1.71e-5)	7.31 (0.027)	0.87	903,982	0.98	0.74
Publishing and Printing	22	0.471 (1.90e-3)	-1.10e-2 (5.94e-5)	3.83e-4 (2.48e-5)	5.93 (0.037)	0.81	569,433	1.41	0.61
Production of Electricity	40	0.366 (2.17e-3)	-9.91e-3 (4.35e-5)	-3.92e-4 (1.80e-5)	8.69 (0.038)	0.90	553,837	1.08	1.01
Arts, Sports and Leisure	92	0.431 (1.88e-3)	-1.16e-2 (6.04e-5)	1.37e-4 (2.46e-5)	6.56 (0.038)	0.84	525,962	1.21	0.50
Machinery Manuf.	31	0.410 (1.92e-3)	-9.62e-3 (5.55e-5)	-3.81e-4 (2.46e-5)	7.68 (0.036)	0.85	528,438	1.11	0.56

This table summarizes the estimation of the wage proxies according to (E.1). We use all individuals in our selected demographic group in the 1,000 selected labor markets from our initial sample (i.e., before selecting 10 million observations). The linear coefficient on age is collinear with the year fixed effects. The last two columns report the standard deviation of the individual-sector fixed effects, α_i^s, and of the labor market shifters, w_t^{sl}.

location pairs to combine is guided by the results in Corollary 3. This corollary indicates that a requisite for the moment inequality in equation (16) to point identify the wage parameter α is that the locations l and l' entering the moment function offer the same amenity level in the population of reference; when $\kappa_{nt}^{l} - \kappa_{nt}^{l'} = 0$. Enforcing this condition is infeasible as these amenity levels are continuous variables that are only estimated later in our estimation procedure. However, as κ^l_{nt} accounts for migration costs in our setting, it is reasonable to expect it will vary with the distance between locations n and l. Consequently, locations l and l' that are at a similar distance to the origin n are more likely to have similar amenity levels and, when combined in the context of the moment inequality in equation (16), yield smaller identified sets. Thus, we only form the moment function (16) for location l and l' for which the difference between the distance from n to l and the distance from n to l' is in the lower tercile of all pairwise differences in distance to n. Given such location pairs, we form the moment function in equation (16) by aggregating across all sector pairs s and r, and across worker pairs within those sectors.

E.3 PPML-IV Estimator

We describe here the estimator in Artuç and McLaren (2015). To derive it in our setting, we must assume all workers employed in the same sector s have the same information set in any given period t, regardless of their location of residence. Thus, $\mathcal{J}_{ist} = \mathcal{J}_{i'st}$ for any sector s, period t, and any two workers i and i' employed in s at t. Given the assumption that all workers within a sector s and period t have identical information sets, we can write the model-implied number of sector s workers that migrate between locations n and l at t as:

$$\begin{aligned} M^l_{nst} &= \frac{\exp\left(\alpha\mathbb{E}[w^l_{st}|\mathcal{J}_{st}] - \kappa^l_{nt}\right)}{\sum_k \exp\left(\alpha\mathbb{E}[w^k_{st}|\mathcal{J}_{st}] - \kappa^k_{nt}\right)} L_{nst-1} \\ &= \exp\left(\alpha\mathbb{E}[w^l_{st}|\mathcal{J}_{st}] - \kappa^l_{nt} + \Gamma_{nst}\right) \\ &= \exp\left(\alpha w^l_{st} - \alpha\xi^l_{st} - \kappa^l_{nt} + \Gamma_{nst}\right) \\ &= \exp\left(\Lambda^l_{st} + \Gamma_{nst} + \Psi^l_{nt}\right), \end{aligned} \tag{E.3}$$

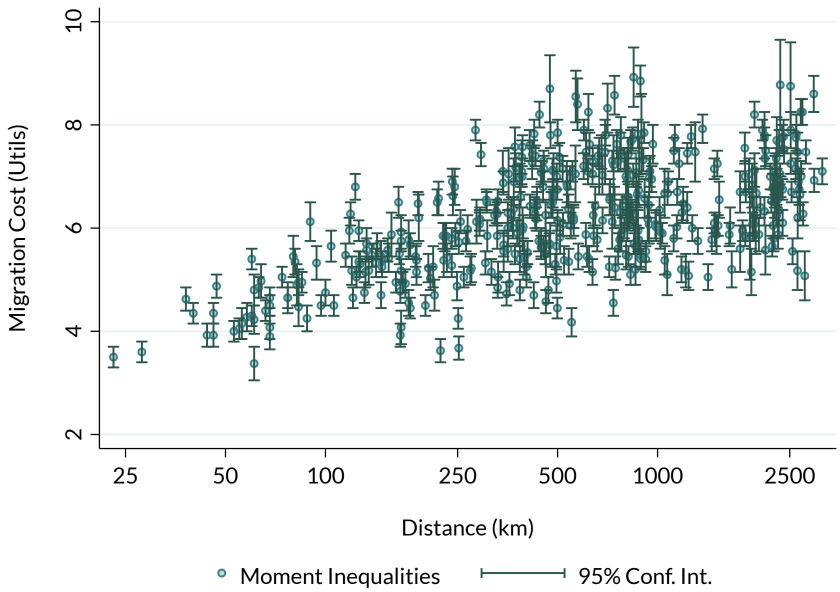
where L_{nst-1} is the total number of workers in location n and sector s at period $t-1$; $\xi^l_{st} \equiv w^l_{st} - \mathbb{E}[w^l_{st}|\mathcal{J}_{st}]$ is these workers' expectational error when predicting wages in sector s, period t, and location l; and

$$\Lambda^l_{st} \equiv \alpha w^l_{st} - \alpha\xi^l_{st}, \quad \Gamma_{nst} \equiv -\ln\left(\sum_k \exp\left(\alpha\mathbb{E}[w^k_{st}|\mathcal{J}_{st}] - \kappa^k_{nt}\right)\right) + \ln L_{nst}, \quad \Psi^l_{nt} \equiv -\kappa^l_{nt}.$$

Using information on $\{M^l_{nst}\}^{L,L}_{n=1,l=1}$ for a period t, sector s, and L origin and destination locations, the procedure in Artuç and McLaren (2015) recovers estimates of α and $\{\kappa^l_{nt}\}^{L,L}_{n=1,l=1}$ in three steps. First, compute PPML estimates of $\{\Lambda^l_{st}\}^L_{l=1}$, $\{\Gamma_{nst}\}^L_{n=1}$ and $\{\Psi^l_{nt}\}^{L,L}_{n=1,l=1}$ using the expression in the last line in equation (E.3). Second, under the assumption that a variable z^l_{st} correlated with w^l_{st} belongs to the information set $\mathcal{J}_{st}$, compute an IV estimate of α by regressing $\hat{\Lambda}^l_{st}$ on w^l_{st} using z^l_{st} as an instrument, with $\hat{\Lambda}^l_{st}$ the first-step estimate of Λ^l_{st}. Third, recover $\hat{\kappa}^l_{nt} = -\hat{\Psi}^l_{nt}$ for every origin n and destination l.

E.4 Additional Results

Figure E.2: Migration Costs from Moment Inequalities with Confidence Intervals



This figure displays the main estimates of the preference parameters from our working sample. Each point shows the midpoint of the 95-percent confidence interval for a given bilateral migration cost $\{\kappa^l_{nt}\}$ in the year 2011, expressed in utils, and the associated 95-percent confidence interval.

F Dynamic Model: Proofs and Additional Derivations

F.1 Details on Two-Step Estimation Procedure

We denote by $(\theta_\alpha, \theta_\beta)$ the parameter vector with true value (α, β), and by $\Theta_{(\alpha,\beta)}$ the set of possible values of $(\theta_\alpha, \theta_\beta)$. Similarly, for each $l = 2, \ldots, L$ and sample period t, we denote by θ_t^l the parameter with true value λ_t^l and by Θ_t^l the set of possible values of θ_t^l. In Appendix F.1.1, we describe the estimation of the parameter vector $(\theta_\alpha, \theta_\beta)$. In Appendix F.1.2, we discuss the estimation of the parameter θ_t^l for all $l = 2, \ldots, L$ and sample period t.

F.1.1 First-Step: Estimating Migration Costs and Wage Coefficient

For any period t, any locations l and l', any worker i of type s and any worker j of type r, any vectors $z_{st} \in \mathcal{Z}_{st}$ and $z_{rt} \in \mathcal{Z}_{rt}$, and any deterministic function $g_{ijsrt}^{ll'} \colon \mathcal{Z}_s \times \mathcal{Z}_r \times \Theta_{(\alpha,\beta)} \to \mathbb{R}$, we define the moment

$$\begin{gathered}\tilde{\mathbb{M}}_{ijsrt}^{ll'}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta, g_{ijsrt}^{ll'}(\cdot)) \equiv \\ \mathbb{E}[y_{ist}^l y_{jrt}^l + y_{ist}^{l'} y_{jrt}^{l'} - y_{ist}^l y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) \times \\ (2 + 2g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta) - (\Delta\tilde{v}_{ist}^{ll'} + \Delta\tilde{v}_{jrt}^{l'l}))|z_s, z_r].\end{gathered} \tag{F.1}$$

Theorem 4 establishes a property of this moment when evaluated at $\theta_\alpha = \alpha$ and $\theta_\beta = \beta$.

Theorem 4 *Assume equations (26) to (33) hold, $z_{st} \subset \mathcal{J}_{ist}$ for worker i of type s at period t, and $z_{rt} \subset \mathcal{J}_{jrt}$ for worker j of type r at period t. Then, $\tilde{\mathbb{M}}_{ijsrt}^{ll'}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta, g_{ijsrt}^{ll'}(\cdot)) \geqslant 0$ for any period t, locations l and l', worker i of type s, worker j of type r, $z_{st} \in \mathcal{Z}_{st}$, $z_{rt} \in \mathcal{Z}_{rt}$, and deterministic function $g_{ijsrt}^{ll'} \colon \mathcal{Z}_{st} \times \mathcal{Z}_{rt} \times \Theta_{(\alpha,\beta)} \to \mathbb{R}$.*

We prove Theorem 4 in Appendix F.4. Theorem 4 states that, given equations (26) to (33), the assumption that z_{st} belongs to the information set of worker i of type s at period t, and the assumption that z_{rt} belongs to the information set of worker j of type r at period t, the moment in equation (F.1) is positive when evaluated at $(\theta_\alpha, \theta_\beta) = (\alpha, \beta)$. Furthermore, this is true regardless of the period t, regardless of the two locations l and l' we compare, regardless of the workers is and jr we consider, regardless of the vectors z_{st} and z_{rt} on which we condition, and regardless of the deterministic function $g_{ijsrt}^{ll'}(\cdot)$ we use to form the moment. We thus may compute the set of values of $(\theta_\alpha, \theta_\beta)$ that satisfy

$$\tilde{\mathbb{M}}_{ijsrt}^{ll'}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta, g_{ijsrt}^{ll'}(\cdot)) \geqslant 0, \tag{F.2}$$

and, if equations (26) to (33) hold, $z_{st} \subset \mathcal{J}_{ist}$, and $z_{rt} \subset \mathcal{J}_{jrt}$, (α, β) will belong to this set.

F.1.2 Second-Step: Estimating Location-Specific Amenities

Denote by $\Delta\theta_t^{ll'} \equiv \theta_t^l - \theta_t^{l'}$ the parameter with true value $\Delta\lambda_t^{ll'} \equiv \lambda_t^l - \lambda_t^{l'}$, and by $\Theta_t^{ll'}$ the set of possible values of $\Delta\theta_t^{ll'}$. For any type s and period t, denote by $\mathcal{Z}_{st}$ the support of z_{st}. Then, for any $z_{st} \in \mathcal{Z}_{st}$ and deterministic function $h_{ist}^{ll'} \colon \mathcal{Z}_{st} \times \Theta_t^{ll'} \rightarrow \mathbb{R}$, we define the moment

$$\begin{gathered}\tilde{\mathbb{m}}_{ist}^{ll'}(z_{st}, \Delta\theta_t^{ll'}, h_{ist}^{ll'}(\cdot)) \equiv \\ \mathbb{E}[y_{ist}^{l'} - y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\theta_t^{ll'}))(1 + h_{ist}^{ll'}(z_{st}, \Delta\theta_t^{ll'}) - \Delta\tilde{v}_{ist}^{ll'}(\Delta\theta_t^{ll'}))|z_{st}],\end{gathered} \tag{F.3}$$

with

$$\Delta\tilde{v}_{ist}^{ll'}(\Delta\theta_t^{ll'}) = \beta\Delta x_{nt}^{ll'} + \Delta\theta_t^{ll'} + \alpha\Delta w_{ist}^{ll'} + \delta\beta \sum_{l''=1}^{L} y_{ist+1}^{(lt)l''}(x_{lt+1}^{l''} - x_{l't+1}^{l''}). \tag{F.4}$$

Theorem 5 establishes a property of this moment when evaluated at $\Delta\theta_t^{ll'} = \Delta\lambda_t^{ll'}$.

Theorem 5 *Assume equations (26) to (33) hold and $z_{st} \subset \mathcal{J}_{ist}$. Then, for any period t, locations l and l', worker i of type s, $z_{st} \in \mathcal{Z}_{st}$, and deterministic function $h_{ist}^{ll'} \colon \mathcal{Z}_{st} \times \Theta_t^{ll'} \rightarrow \mathbb{R}$, $\tilde{\mathbb{m}}_{ist}^{ll'}(z_{st}, \Delta\theta_t^{ll'}, h_{ist}^{ll'}(\cdot)) \geqslant 0$.*

We prove this theorem in Appendix F.2, and provide an expression for the optimal function $h_{ist}^{ll'}(\cdot)$ in Appendix F.3. Theorem 5 shows that, given knowledge of (α, β), one may bound the amenity difference $\Delta\lambda_t^{ll'}$ for any sample period t and locations l and l'. Appendix F.1.1 shows how to derive moment inequalities that are informative about (α, β).

F.2 Second-Step Bounding Inequalities: Proof of Theorem 5

Equation (26) implies that, for any worker i of type s, period t, and locations l and l',

$$(y_{ist}^l + y_{ist}^{l'})(\mathbb{1}\{\mathbb{E}[\mathcal{V}_{is}^l - \mathcal{V}_{is}^{l'}|\mathcal{J}_{ist}] \geqslant 0\} - y_{ist}^l) = 0.$$

Equations (27a) and (29) imply we can rewrite this inequality as

$$(y_{ist}^l + y_{ist}^{l'})(\mathbb{1}\{\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{J}_{ist}] + \Delta\varepsilon_{ist}^{ll'} \geqslant 0\} - y_{ist}^l) = 0,$$

where $\Delta v_{ist}^{ll'} = v_{ist}^l - v_{ist}^{l'}$ and $\Delta\varepsilon_{ist}^{ll'} = \varepsilon_{ist}^l - \varepsilon_{ist}^{l'}$. Equation (31) implies we can further write this equality as

$$(y_{ist}^l + y_{ist}^{l'})(\mathbb{1}\{\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \Delta\varepsilon_{ist}^{ll'} \geqslant 0\} - y_{ist}^l) = 0. \tag{F.5}$$

This equality holds for any worker i of any type s, any period t, and any two locations l and l'. As the next step, we take the expectation of both sides of this equality conditional on a value of $\mathcal{W}_{ist}$ and a dummy variable that equals one if worker i of type s chooses either location l or location l' at period t; i.e.,

$$\mathbb{E}[\mathbb{1}\{\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \Delta\varepsilon_{ist}^{ll'} \geqslant 0\} - y_{ist}^{l}|\mathcal{W}_{ist}, y_{ist}^{l} + y_{ist}^{l'} = 1] = 0.$$

Equation (33) implies we can rewrite this moment equality as

$$\mathbb{E}\left[\frac{\exp(\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}])}{1 + \exp(\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}])} - y_{ist}^{l}\middle|\mathcal{W}_{ist}, y_{ist}^{l} + y_{ist}^{l'} = 1\right] = 0,$$

or, equivalently,

$$\mathbb{E}[(1 + \exp(-\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}]))^{-1} - y_{ist}^{l}|\mathcal{W}_{ist}, y_{ist}^{l} + y_{ist}^{l'} = 1] = 0.$$

Multiplying by $1 + \exp(-\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}])$, we obtain

$$\mathbb{E}[1 - y_{is}^{l}(1 + \exp(-\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}]))|\mathcal{W}_{ist}, y_{ist}^{l} + y_{ist}^{l'} = 1] = 0,$$

or, equivalently,

$$\mathbb{E}[1 - y_{ist}^{l} - y_{ist}^{l}\exp(-\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}])|\mathcal{W}_{ist}, y_{ist}^{l} + y_{ist}^{l'} = 1] = 0.$$

Given the conditioning on the event $y_{ist}^{l} + y_{ist}^{l'} = 1$, we can further rewrite

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l}(-\exp(-\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}]))|\mathcal{W}_{ist}] = 0. \tag{F.6}$$

As $-\exp(-x)$ is concave in $x \in \mathbb{R}$, a linear approximation to this function at any value $a \in \mathbb{R}$ will bound it from above. The formula for this linear approximation is

$$-\exp(-a) + \exp(-a)(x - a) = \exp(-a)(-(1 + a) + x).$$

Thus, given any deterministic function $h_{ist}^{ll'} \colon \mathcal{Z}_{st} \times \Theta_{t}^{ll'} \to \mathbb{R}$ and equation (F.6), we can derive

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l}\exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_{t}^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_{t}^{ll'})) + \mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}])|\mathcal{W}_{ist}] \geqslant 0. \tag{F.7}$$

Defining as $\nu_{ist}^{ll'}$ the expectational error worker i of type s makes when forecasting at period t the variable Δv_{ist}^{ll}, the assumption that workers expectations are rational (see equation (2)) implies

$$\nu_{ist}^{ll'} \equiv \Delta v_{ist}^{ll'} - \mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}], \qquad \Longrightarrow \qquad \mathbb{E}[\nu_{ist}^{ll'}|\mathcal{W}_{ist}] = 0. \tag{F.8}$$

Let's consider the moment

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + \Delta v_{ist}^{ll'})|\mathcal{W}_{ist}]. \tag{F.9}$$

or, equivalently,

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + (\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \nu_{ist}^{ll'}))|\mathcal{W}_{ist}].$$

Given that $\mathcal{W}_{ist} \subset \mathcal{J}_{ist}$, we can use the LIE to rewrite this moment as

$$\mathbb{E}[\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + (\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \nu_{ist}^{ll'}))|\mathcal{J}_{ist}]|\mathcal{W}_{ist}].$$

Equation (26) implies $\mathbb{E}[y_{is}^{l}|\mathcal{J}_{is}] = y_{is}^{l}$. Consequently, if $z_s \subseteq \mathcal{W}_{is}$, it is then the case that $z_s \subset \mathcal{J}_{is}$ and we can further rewrite

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + \mathbb{E}[\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \nu_{ist}^{ll'}|\mathcal{J}_{ist}])|\mathcal{W}_{ist}].$$

As $\Delta v_{ist}^{ll'} = \mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \nu_{ist}^{ll'}$, equation (31) further implies that

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + \mathbb{E}[\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}] + \nu_{ist}^{ll'}|\mathcal{W}_{ist}])|\mathcal{W}_{ist}],$$

and equation (F.8) implies we can rewrite this moment as

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + \mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}])|\mathcal{W}_{ist}].$$

However, this moment is exactly the same entering the moment inequality in equation (F.7), which implies that the following inequality involving the moment in equation (F.9) is equivalent to that in equation (F.7):

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + \Delta v_{ist}^{ll'})|\mathcal{W}_{ist}] \geqslant 0.$$

Finally, if $z_{st} \subset \mathcal{W}_{ist}$, we can use the LIE and conclude that

$$\mathbb{E}[y_{ist}^{l'} + y_{ist}^{l} \exp(-h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'}))(-(1 + h_{ist}^{ll'}(z_{st}, \Delta\lambda_t^{ll'})) + \Delta v_{ist}^{ll'})|z_{st}] \geqslant 0. \tag{F.10}$$

This moment is for the dynamic model described in Section 6 analogous to that in equation (7) for the case of the static model described in Section 2. The only difference between both moments is that the difference in the static utility between choosing alternatives l and l' entering the moment in equation (7) (i.e., $\Delta\theta_{ll'} + \alpha\Delta w_s^{ll'}$) is substituted by the difference in the corresponding choice-specific value functions (i.e., $\Delta v_{ist}^{ll'}$).

However, the moment inequality in equation (F.10) is not immediately useful for the identification of the parameters of the dynamic model described in Section 6. To make this point clear, rewrite $\Delta v_{ist}^{ll'} = v_{ist}^{l} - v_{ist}^{l'}$. Thus, the difference in the choice-specific value functions $\Delta v_{ist}^{ll'}$ depends on the optimal choices of the ith worker of type s in every period $t' > t$ both conditional on choosing l at period t (which matters for the value of v_{ist}^{l}) and conditional on choosing l' at period t (which matters for the value of $v_{ist}^{l'}$). To derive from the inequality in equation (F.10) a moment inequality that may be used for the (partial) identification of the parameters of the model described in Section 6, we follow the approach in Morales et al. (2019). Specifically, we substitute the difference in choice-specific value functions $\Delta v_{ist}^{ll'}$ by the variable

$$\Delta\tilde{v}_{ist}^{ll'} \equiv v_{ist}^{l} - \tilde{v}_{ist}^{l'}, \tag{F.11}$$

where $\tilde{v}_{ist}^{l'}$ is the discounted sum of static utilities from period t onwards (that is, in every period $t' \geqslant t$) if worker i of types s chooses location l' at period t but follows in every subsequent every period $t' > t$ the path of choices that would be optimal if they had instead chosen location l at t. To define v_{ist}^{l}, $v_{ist}^{l'}$, and $\tilde{v}_{ist}^{l'}$ formally, denote by

$$y_{ist'}^{(lt)} = (y_{ist'}^{(lt)1}, \ldots, y_{ist'}^{(lt)L}) \tag{F.12}$$

the vector that codes the optimal choice of worker i of type s at period t' if they were to choose alternative l at period t, i.e., $y_{ist'}^{(lt)l'} = 1$ if i would choose l', zero otherwise. Similarly, $y_{ist'}^{(l't)} = (y_{ist'}^{(l't)1}, \ldots, y_{ist'}^{(l't)L})$ codes the optimal choice of worker i in period t' if they were to choose alternative l' at period t. Then, we can write

$$\begin{aligned} v_{ist}^{l} &= \kappa_{nt}^{l} + \alpha w_{st}^{l} + \delta \sum_{l''=1}^{L} y_{ist+1}^{(lt)l''}(\kappa_{lt+1}^{l''} + \alpha w_{st+1}^{l''} + \varepsilon_{ist+1}^{l''}) \\ &\quad + \sum_{t'>t+1} \delta^{t'-t} \sum_{n'=1}^{L}\sum_{l''=1}^{L} y_{ist'}^{(lt)n'} y_{ist'+1}^{(lt)l''}(\kappa_{n't'+1}^{l''} + \alpha w_{st'+1}^{l''} + \varepsilon_{ist'+1}^{l''}), \end{aligned} \tag{F.13a}$$

$$v_{ist}^{l'} = \kappa_{nt}^{l'} + \alpha w_{st}^{l'} + \delta \sum_{l''=1}^{L} y_{ist+1}^{(l't)l''}(\kappa_{l't+1}^{l''} + \alpha w_{st+1}^{l''} + \varepsilon_{ist+1}^{l''})$$

$$+\sum_{t'>t+1}\delta^{t'-t}\sum_{n'=1}^{L}\sum_{l''=1}^{L}y_{ist'}^{(l't)n'}y_{ist'+1}^{(l't)l''}(\kappa_{n't'+1}^{l''}+\alpha w_{st'+1}^{l''}+\varepsilon_{ist'+1}^{l''}), \tag{F.13b}$$

$$\tilde{v}_{ist}^{l'}=\kappa_{nt}^{l'}+\alpha w_{st}^{l'}+\delta\sum_{l''=1}^{L}y_{ist+1}^{(lt)l''}(\kappa_{l't+1}^{l''}+\alpha w_{st+1}^{l''}+\varepsilon_{ist+1}^{l''})$$

$$+\sum_{t'>t+1}\delta^{t'-t}\sum_{n'=1}^{L}\sum_{l''=1}^{L}y_{ist'}^{(lt)n'}y_{ist'+1}^{(lt)l''}(\kappa_{n't'+1}^{l''}+\alpha w_{st'+1}^{l''}+\varepsilon_{ist'+1}^{l''}). \tag{F.13c}$$

Equations (26) and (30) imply that

$$\mathbb{E}[v_{ist}^{l'}|\mathcal{W}_{ist}]\geqslant\mathbb{E}[\tilde{v}_{ist}^{l'}|\mathcal{W}_{ist}], \tag{F.14}$$

and, consequently,

$$\mathbb{E}[\Delta v_{ist}^{ll'}|\mathcal{W}_{ist}]\geqslant\mathbb{E}[\Delta\tilde{v}_{ist}^{ll'}|\mathcal{W}_{ist}]. \tag{F.15}$$

Equations (F.10) and (F.15) imply the following moment inequality

$$\mathbb{E}[y_{ist}^{l'}+y_{ist}^{l}\exp(-h_{ist}^{ll'}(z_{st},\Delta\lambda_t^{ll'}))(-(1+h_{ist}^{ll'}(z_{st},\Delta\lambda_t^{ll'}))+\Delta\tilde{v}_{ist}^{ll'})|z_{st}]\geqslant 0. \tag{F.16}$$

Comparing the expressions for v_{ist}^{l} and $\tilde{v}_{ist}^{l'}$ in equations (F.13a) and (F.13c), we can write

$$\begin{aligned}\Delta\tilde{v}_{ist}^{ll'}=v_{ist}^{l}-\tilde{v}_{ist}^{l'}&=(\kappa_{nt}^{l}-\kappa_{nt}^{l'})+\alpha(w_{st}^{l}-w_{st}^{l'})+\delta\sum_{l''=1}^{L}y_{ist+1}^{(lt)l''}(\kappa_{lt+1}^{l''}-\kappa_{l't+1}^{l''})\\&=\beta(x_{nt}^{l}-x_{nt}^{l'})+(\lambda_t^{l}-\lambda_t^{l'})+\alpha(w_{st}^{l}-w_{st}^{l'})+\delta\beta\sum_{l''=1}^{L}y_{ist+1}^{(lt)l''}(x_{lt+1}^{l''}-x_{l't+1}^{l''}).\end{aligned} \tag{F.17}$$

By plugging equation (F.17) into equation (F.16), we obtain a moment inequality whose moment equals that in equation (F.3) when evaluated at $\Delta\theta_{ll'}=\Delta\lambda^{ll'}$. Equations (F.16) and (F.17) thus imply Theorem 5. ■

F.3 Second-Step Bounding Inequalities: Additional Derivations

Derivation of optimal function $h_{ist}^{ll'}(z_{st},\Delta\theta_{ll})$. Given $z_{st}\in\mathcal{Z}_{st}$, we establish the value of $h_{ist}^{ll'}(z_{st},\Delta\theta_{ll})$ that minimizes the moment in equation (F.3) at each value of $\Delta\theta_{ll'}$. Specifically, given z_{st} and $\Delta\theta_{ll'}$ the first-order condition of the moment in equation (F.3) with respect to the scalar $h^{ll'}_{ist}(z_{st}, \Delta\theta_{ll})$ is

$$\mathbb{E}[y^l_{ist}(h^{ll'}_{ist}(z_{st}, \Delta\theta_{ll'}) - \Delta\tilde{v}^{ll'}_{ist})|z_{st}],$$

or, equivalently,

$$\mathbb{E}[h^{ll'}_{ist}(z_{st}, \Delta\theta_{ll'}) - \Delta\tilde{v}^{ll'}_{ist}|z_{st}, y^l_{ist} = 1].$$

Setting this moment condition to zero, we solve for $h^{ll'}_{ist}(z_{st}, \Delta\theta_{ll'})$ to obtain the following solution:

$$h^{ll'}_{ist}(z_{st}, \Delta\theta_{ll'}) = \mathbb{E}[\Delta\tilde{v}^{ll'}_{ist}|z_{st}, y^l_{ist} = 1], \tag{F.18}$$

with $\Delta\tilde{v}^{ll'}_{ist}$ defined as in equation (F.17).

F.4 First-Step Moment Inequalities: Proof of Theorem 4

For any locations l and l', any period t, and any worker i of type s, equation (F.5) indicates the following equality holds

$$(y^l_{ist} + y^{l'}_{ist})(\mathbb{1}\{\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}] + \Delta\varepsilon^{ll'}_{ist} \geqslant 0\} - y^l_{ist}) = 0.$$

For any locations l and l', any period t, any worker i of type s, and any worker j of type r, we can thus derive the following equality:

$$y^{l'}_{jrt}(y^l_{ist} + y^{l'}_{ist})(\mathbb{1}\{\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}] + \Delta\varepsilon^{ll'}_{ist} \geqslant 0\} - y^l_{ist}) = 0. \tag{F.19}$$

Taking the expectation of both sides of this equality conditional on $\mathcal{W}_{ist}$, $\mathcal{W}_{jrt}$, and on dummy variable that equals one if worker i of type s chooses either location l or location l' at period t, we obtain

$$\mathbb{E}[y^{l'}_{jrt}(y^l_{ist} + y^{l'}_{ist})(\mathbb{1}\{\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}] + \Delta\varepsilon^{ll'}_{ist} \geqslant 0\} - y^l_{ist})|\mathcal{W}_{ist}, \mathcal{W}_{jrt}, y^l_{ist} + y^{l'}_{ist} = 1] = 0.$$

Given equations (26) and (33), we can rewrite this moment equality as

$$\mathbb{E}\left[y^{l'}_{jrt}\left(\frac{\exp(\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}])}{1 + \exp(\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}])} - y^l_{ist}\right)\middle|\mathcal{W}_{ist}, \mathcal{W}_{jrt}, y^l_{ist} + y^{l'}_{ist} = 1\right] = 0,$$

or, equivalently,

$$\mathbb{E}[y^{l'}_{jrt}((1+\exp(-\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}]))^{-1} - y^{l}_{ist})|\mathcal{W}_{ist}, \mathcal{W}_{jrt}, y^{l}_{ist} + y^{l'}_{ist} = 1] = 0.$$

Multiplying by $1 + \exp(-\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}])$, we obtain

$$\mathbb{E}[y^{l'}_{jrt}(1 - y^{l}_{ist}(1+\exp(-\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}])))|\mathcal{W}_{ist}, \mathcal{W}_{jrt}, y^{l}_{ist} + y^{l'}_{ist} = 1] = 0,$$

or, equivalently,

$$\mathbb{E}[y^{l'}_{jrt}(1 - y^{l}_{ist} - y^{l}_{ist}\exp(-\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}]))|\mathcal{W}_{ist}, \mathcal{W}_{jrt}, y^{l}_{ist} + y^{l'}_{ist} = 1] = 0.$$

Given that this expectation conditions on the event $y^{l}_{ist} + y^{l'}_{ist} = 1$, we can further rewrite

$$\mathbb{E}[y^{l'}_{jrt}(y^{l'}_{ist} + y^{l}_{ist}(-\exp(-\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}])))|\mathcal{W}_{ist}, \mathcal{W}_{jrt}] = 0,$$

or, equivalently,

$$\mathbb{E}[y^{l'}_{ist}y^{l'}_{jrt} + y^{l}_{ist}y^{l'}_{jrt}(-\exp(-\mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}]))|\mathcal{W}_{ist}, \mathcal{W}_{jrt}] = 0.$$

Since the function $-\exp(-x)$ is concave in $x \in \mathbb{R}$, we can derive the following inequality, given any deterministic function $g^{ll'}_{ijsrt}\colon \mathcal{Z}_{st} \times \mathcal{Z}_{rt} \times \Theta_{(\alpha,\beta)} \to \mathbb{R}$,

$$\begin{gathered}\mathbb{E}[y^{l'}_{ist}y^{l'}_{jrt} + y^{l}_{ist}y^{l'}_{jrt}\exp(-g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \alpha, \beta))\times \\ (-(1+g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \alpha, \beta)) + \mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}])|\mathcal{W}_{ist}, \mathcal{W}_{jrt}] \geqslant 0.\end{gathered} \tag{F.20}$$

Let's consider the moment

$$\mathbb{E}[y^{l'}_{ist}y^{l'}_{jrt} + y^{l}_{ist}y^{l'}_{jrt}\exp(-g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \alpha, \beta))(-(1+g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \alpha, \beta)) + \Delta v^{ll'}_{ist})|\mathcal{W}_{ist}, \mathcal{W}_{jrt}], \tag{F.21}$$

or, equivalently,

$$\begin{gathered}\mathbb{E}[y^{l'}_{ist}y^{l'}_{jrt} + y^{l}_{ist}y^{l'}_{jrt}\exp(-g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \alpha, \beta))\times \\ (-(1+g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \alpha, \beta)) + \mathbb{E}[\Delta v^{ll'}_{ist}|\mathcal{W}_{ist}] + \nu^{ll'}_{ist})|\mathcal{W}_{ist}, \mathcal{W}_{jrt}],\end{gathered}$$

where $\nu^{ll'}_{ist}$ is defined as in equation (F.8).

Equation (31) implies $\nu_{ist}^{ll'} = \Delta w_{ist}^{ll'} - \mathbb{E}[\Delta w_{st}^{ll'} | \mathcal{W}_{ist}, \mathcal{W}_{jrt}]$ and, thus, we can conclude that

$$\mathbb{E}[\nu_{ist}^{ll'} | \mathcal{W}_{ist}, \mathcal{W}_{jrt}] = 0. \tag{F.22}$$

As $\mathcal{W}_{ist} \subset \mathcal{J}_{ist}$ and $\mathcal{W}_{jrt} \subset \mathcal{J}_{jrt}$, we use the LIE to rewrite the moment in equation (F.21) as

$$\mathbb{E}[\mathbb{E}[y_{ist}^{l'} y_{jrt}^{l'} + y_{ist}^{l} y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) \times \\ (-(1 + g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) + \mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{W}_{ist}] + \nu_{ist}^{ll'})) | \mathcal{J}_{ist}, \mathcal{J}_{jrt}] | \mathcal{W}_{ist}, \mathcal{W}_{jrt}].$$

Equation (26) implies $\mathbb{E}[y_{ist}^{l} y_{jrt}^{l'} | \mathcal{J}_{ist}, \mathcal{J}_{jrt}] = y_{ist}^{l} y_{jrt}^{l'}$. Consequently, if $z_{st} \subseteq \mathcal{W}_{ist}$ and $z_{rt} \subseteq \mathcal{W}_{jrt}$, then $z_{st} \subset \mathcal{J}_{ist}$ and $z_r \subset \mathcal{J}_{jrt}$, and we can rewrite the moment in equation (F.21) as

$$\mathbb{E}[y_{ist}^{l'} y_{jrt}^{l'} + y_{ist}^{l} y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) \times \\ (-(1 + g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) + \mathbb{E}[\mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{W}_{ist}] + \nu_{ist}^{ll'} | \mathcal{J}_{ist}, \mathcal{J}_{jrt}]) | \mathcal{W}_{ist}, \mathcal{W}_{jrt}].$$

Equation (31) further implies that

$$\mathbb{E}[y_{ist}^{l'} y_{jrt}^{l'} + y_{ist}^{l} y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) \times \\ (-(1 + g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) + \mathbb{E}[\mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{W}_{ist}] + \nu_{ist}^{ll'} | \mathcal{W}_{ist}, \mathcal{W}_{jrt}]) | \mathcal{W}_{ist}, \mathcal{W}_{jrt}],$$

and equation (F.22) implies we can rewrite the moment in equation (F.21) as

$$\mathbb{E}[y_{ist}^{l'} y_{jrt}^{l'} + y_{ist}^{l} y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) \times \\ (-(1 + g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) + \mathbb{E}[\Delta v_{ist}^{ll'} | \mathcal{W}_{ist}]) | \mathcal{W}_{ist}, \mathcal{W}_{jrt}].$$

However, this moment is the same as that in equation (F.20), which implies that the following inequality involving the moment in equation (F.21) is equivalent to that in equation (F.20):

$$\mathbb{E}[y_{ist}^{l'} y_{jrt}^{l'} + y_{ist}^{l} y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta))(-(1 + g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) + \Delta v_{ist}^{ll'}) | \mathcal{W}_{ist}, \mathcal{W}_{jrt}] \geqslant 0. \tag{F.23}$$

Following steps analogous to those we follow to derive the inequality in equations (F.16) from that in equation (F.10) (see Appendix Section F.2), we derive the following inequality from that in equation (F.23):

$$\mathbb{E}[y_{ist}^{l'} y_{jrt}^{l'} + y_{ist}^{l} y_{jrt}^{l'} \exp(-g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta))(-(1 + g_{ijsrt}^{ll'}(z_{st}, z_{rt}, \alpha, \beta)) + \Delta \tilde{v}_{ist}^{ll'}) | \mathcal{W}_{ist}, \mathcal{W}_{jrt}] \geqslant 0. \tag{F.24}$$

where $\Delta\tilde{v}_{ist}^{ll'}$ is defined as in equation (F.17).

The inequality in equation (F.24) is one of the two moment inequalities we combine to obtain the inequality that we use to bound the parameters α and β. To obtain the second moment inequality, we start with the following expression

$$y_{ist}^{l}(y_{jrt}^{l}+y_{jrt}^{l'})(\mathbb{1}\{\mathbb{E}[\Delta v_{jrt}^{l'l}|\mathcal{W}_{jrt}]+\Delta\varepsilon_{jrt}^{l'l}\geqslant 0\}-y_{jrt}^{l'})=0, \tag{F.25}$$

which is analogous to that in equation (F.19). Following the same steps we implement to go from equation (F.19) to equation (F.24), we can derive from equation (F.25) the following inequality

$$\mathbb{E}[y_{ist}^{l'}y_{jrt}^{l'}+y_{ist}^{l}y_{jrt}^{l'}\exp(-g_{ijsrt}^{ll'}(z_{st},z_{rt},\alpha,\beta))(-(1+g_{ijsrt}^{ll'}(z_{st},z_{rt},\alpha,\beta))+\Delta\tilde{v}_{jrt}^{l'l})|\mathcal{W}_{ist},\mathcal{W}_{jrt}]\geqslant 0. \tag{F.26}$$

As the moments in equations (F.23) and (F.26) share the same function $g_{ijsrt}^{ll'}\colon \mathcal{Z}_{st}\times\mathcal{Z}_{rt}\times\Theta_{(\alpha,\beta)}\to\mathbb{R}$ and have the same conditioning set and, we obtain the following moment inequality when we add them:

$$\begin{gathered}\mathbb{E}[y_{ist}^{l}y_{jrt}^{l}+y_{ist}^{l'}y_{jrt}^{l'}-y_{ist}^{l}y_{jrt}^{l'}\exp(-g_{ijsrt}^{ll'}(z_{st},z_{rt},\alpha,\beta))\times\\ (2+2g_{ijsrt}^{ll'}(z_{st},z_{rt},\alpha,\beta)-(\Delta\tilde{v}_{ist}^{ll'}+\Delta\tilde{v}_{jrt}^{l'l}))|\mathcal{W}_{ist},\mathcal{W}_{jrt}]\geqslant 0,\end{gathered}$$

with

$$\Delta\tilde{v}_{ist}^{ll'}=\beta(x_{nt}^{l}-x_{nt}^{l'})+(\lambda_t^{l}-\lambda_t^{l'})+\alpha(w_{st}^{l}-w_{st}^{l'})+\delta\beta\sum_{l''=1}^{L}y_{ist+1}^{(lt)l''}(x_{lt+1}^{l''}-x_{l't+1}^{l''}), \tag{F.27a}$$

$$\Delta\tilde{v}_{jrt}^{l'l}=\beta(x_{nt}^{l'}-x_{nt}^{l})+(\lambda_t^{l'}-\lambda_t^{l})+\alpha(w_{rt}^{l'}-w_{rt}^{l})+\delta\beta\sum_{l''=1}^{L}y_{ist+1}^{(lt)l''}(x_{l't+1}^{l''}-x_{lt+1}^{l''}). \tag{F.27b}$$

Finally, if $z_{st}\subset\mathcal{W}_{ist}$ and $z_{rt}\subset\mathcal{W}_{jrt}$, we can use the LIE and conclude that

$$\begin{gathered}\mathbb{E}[y_{ist}^{l}y_{jrt}^{l}+y_{ist}^{l'}y_{jrt}^{l'}-y_{ist}^{l}y_{jrt}^{l'}\exp(-g_{ijsrt}^{ll'}(z_{st},z_{rt},\alpha,\beta))\times\\ (2+2g_{ijsrt}^{ll'}(z_{st},z_{rt},\alpha,\beta)-(\Delta\tilde{v}_{ist}^{ll'}+\Delta\tilde{v}_{jrt}^{l'l}))|z_{st},z_{rt}]\geqslant 0,\end{gathered} \tag{F.28}$$

with

$$\Delta\tilde{v}_{ist}^{ll'}+\Delta\tilde{v}_{jrt}^{l'l}=\alpha(w_{ist}^{l}-w_{ist}^{l'}+w_{jrt}^{l'}-w_{jrt}^{l})+\delta\beta\sum_{l''=1}^{L}y_{ist+1}^{(lt)l''}(x_{lt+1}^{l''}-x_{l't+1}^{l''}+x_{l't+1}^{l''}-x_{lt+1}^{l''}) \tag{F.29}$$

By plugging equation (F.29) into equation (F.28), we obtain a moment inequality whose moment equals that in equation (F.1) when evaluated at $(\theta_\alpha, \theta_\beta) = (\alpha, \beta)$. Equations (F.28) and (F.29) thus imply Theorem 4. ■

F.5 First-Step Moment Inequalities: Additional Derivations

Derivation of optimal function $g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta)$. We find the value of $g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta)$, given $z_{st} \in \mathcal{Z}_{st}$ and $z_{rt} \in \mathcal{Z}_{rt}$, minimizes the moment in equation (F.28) at each value of $(\theta_\alpha, \theta_\beta)$. Specifically, given z_{st}, z_{rt}, θ_α, and θ_β, the first-order condition of the moment in equation (F.28) with respect to the scalar $g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta)$ is

$$\mathbb{E}[y^l_{ist} y^{l'}_{jrt}(2g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta) - (\Delta\tilde{v}^{ll'}_{ist} + \Delta\tilde{v}^{l'l}_{jrt}))|z_s, z_r] = 0,$$

or, equivalently,

$$\mathbb{E}[2g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta) - (\Delta\tilde{v}^{ll'}_{ist} + \Delta\tilde{v}^{l'l}_{jrt})|z_{st}, z_{rt}, y^l_{ist} y^{l'}_{jrt} = 1].$$

Setting this moment condition to zero, we solve for $g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta)$ to obtain:

$$g^{ll'}_{ijsrt}(z_{st}, z_{rt}, \theta_\alpha, \theta_\beta) = \mathbb{E}[\Delta\tilde{v}^{ll'}_{ist} + \Delta\tilde{v}^{l'l}_{jrt}|z_{st}, y^l_{ist} = 1], \tag{F.30}$$

with $\Delta\tilde{v}^{ll'}_{ist} + \Delta\tilde{v}^{l'l}_{jrt}$ defined as in equation (F.29).

G Model with Endogenous Worker Types

We model workers' choice of market to supply labor. Each labor market is defined by a sector $s = 1, \ldots, S$ and a location $l = 1, \ldots, L$, and we index each market by the combination of indices sl. We focus on the choice of workers in a population of interest defined by the worker's demographic characteristics and prior location and sector. Thus, we omit for simplicity any index identifying the worker's demographic group or prior location or sector of employment.

Defining a variable y^{sl}_i that equals one if worker i chooses market sl (and zero otherwise), we assume

$$y^{sl}_i \equiv \mathbb{1}\{l = \operatorname*{argmax}_{\substack{l'=1,\ldots,L \\ s'=1,\ldots,S}} \mathbb{E}[\mathcal{U}^{s'l'}_i|\mathcal{J}_i]\} \qquad \text{for any } l = 1, \ldots, L \text{ and } s = 1, \ldots, S.$$

We assume worker expectations are rational; that is, determined as in equation (2).

Instead of equation (3), we assume the utility of choosing market sl for a worker i is:

$$\mathcal{U}_i^{sl} = \kappa^l + \tau^s + \alpha w_i^{sl} + \varepsilon_i^{sl},$$

where the new term τ^s is a sector-specific unobserved term that accounts for sector-specific amenities as well as for sector-specific switching costs.

The assumption in equation (4) extends directly to the model with endogenous worker types. More specifically, defining $\varepsilon_i = \{\varepsilon_i^{sl}\}_{s=1,l=1}^{S,L}$, $\kappa = \{\kappa^l\}_{l=1}^{L}$ and $\tau = \{\tau^s\}_{s=1}^{S}$, we assume

$$(\varepsilon_i, \alpha, \kappa, \tau) \subseteq \mathcal{J}_i.$$

The assumption in equation (5) also extends naturally to the model considered here. Specifically, for any sectors s and r, locations l and l', and worker indices i and j, it holds:

$$\mathbb{E}[\Delta w_i^{sll'} | \mathcal{J}_i, \mathcal{J}_j] = \mathbb{E}[\Delta w_i^{sll'} | \mathcal{J}_i] = \mathbb{E}[\Delta w_i^{sll'} | \mathcal{W}_i] = \mathbb{E}[\Delta w^{sll'} | \mathcal{W}_i],$$

where $\Delta w_i^{sll'} = w_i^{sl} - w_i^{sl'}$ and $\Delta w^{sll'} = w^{sl} - w^{sl'}$, with w^{sl} a market-level wage shifter.

Finally, instead of equation (6), we assume that, for any workers i and j, it holds that

$$F_\varepsilon(\varepsilon_i, \varepsilon_j | \mathcal{W}_i, \mathcal{W}_j) = F_\varepsilon(\varepsilon_i) F_\varepsilon(\varepsilon_j) = \exp\left(- \sum_{s=1}^{S} \left(\sum_{l=1}^{L} \exp(-\varepsilon_i^{sl}) \right)^{\psi} - \sum_{s=1}^{S} \left(\sum_{l=1}^{L} \exp(-\varepsilon_j^{sl}) \right)^{\psi} \right),$$

where ψ measures the extent to which the type I extreme value idiosyncratic shocks are correlated across sectors within a location.

References

Artuç, Erhan and John McLaren, "Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility," *Journal of International Economics*, 2015, *97* (2), 278–294. [4, 5, 25, 4, 5, 25, 11]

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885. [4, 8, 4, 8, 2]

_ **and Brian K. Kovak**, "Trade Liberalization and Regional Dynamics," *American Economic Review*, 2017, *107* (10), 2908–2946. [2]

_ **and** _ , "Margins of Labor Market Adjustment to Trade," *Journal of International Economics*, 2019, *117*, 125–142. [2]

Felix, Mayara, “Trade, Labor Market Concentration, and Wages,” *mimeo*, 2022. [2]

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, “Extended Gravity,” *The Review of Economic Studies*, 2019, *86* (6), 2668–2712. [3, 34, 3, 34, 17]

Szerman, Christiane, “The Labor Market Effects of Disability Hiring Quotas,” *mimeo*, 2024. [2]