# Contracted Labor Mobility and Self-selection on Job Match Quality

Juho Alasalmi*
University of Konstanz / Pellervo Economic Research

July 2021

**Abstract**

Contracted migrants observe their wages in the source and destination locations before their migration choice. With labor markets characterized by wage dispersion, the migration choices are then not based on the mean or the variance of source and destination location wage distributions but on specific realizations from these distributions. The deviations of the wage realizations from their expectations, the job match qualities in the source and destination, become factors of selection. The Roy-Borjas model extended by wage dispersion in source and destination location labor markets predicts negative selection on source and positive selection on destination job match quality. Mobility costs amplify the selection on job match quality. I discuss how the comparison of migrants and stayers is not helpful in identifying selection on of job match quality and instead, using Finnish administrative data, compare the contracted migrants to workers who similarly contract a job outside their region of residence but choose to commute. Comparison of the contracted migrants and commuters, groups facing different costs of mobility, but otherwise arguably comparable, reveals selection on job match quality.

*JEL classification:* J61; R23; C24; C25
*Keywords:* Geographical labor mobility, internal migration, migrant selection, job match quality

# 1 Introduction

The economic literature on labor related migration has mainly interpreted migration as speculative. That is, that migrants relocate in the hopes of finding employment in their destination inducing uncertainty in destination location payoffs. However, migration is often contracted, occuring only after a job in destination has been found and the payoff in the destination has realized. The processes that select speculative and contracted

*juho.alasalmi@uni-konstanz.de, Comments most welcome.

migrants are different. Hence, models of speculative migration applied to explain selection of contracted migrants are misspecified leading to biases in the interpretation of results. Thus, the migrant selection literature would advance by taking the possibility of contracted migration more seriously. This paper takes a step to this direction.

The selection mechanisms when opportunities can be searched and secured in other markets before entering them are complicated, and not properly acknowlegded in the economic literature of market self-selection following Roy (1951). The first selective behavior is to search: only those who see it potentially increasing their welfare net search and mobility costs search for opportunities in other markets. The second hurdle is not up to the searcher: the market provides opportunities to those it judges to fit its goals the best. In the labor market, employers select their employees and the hiring of profit-maximizing employers is not random. Finally, if a searcher receives an offer, she judges whether, net relocation costs the offer is worth accepting.

The approach to reduce this complexity taken in this paper is to take the first two of these hurdles as exogenous and focus on a subset of source location workers who have passed these hurdles and are at risk of mobility and to focus on the contribution of the last hurdle in selecting those who successfully change the market of their supply. This allows the use of tractable techniques for studying selection but, at the same time, letting those at risk of mobility be nonrandomly selected from the source location workers respects the selection that search and employers' hiring choice create. As a result, I provide a very tractable extension of Roy-Borjas model to contracted migration. This modeling approach is accompanied by and empirical setting is that is not confounded by the first two hurdles as I compare mobile workers who all have passed them.

I capture essential aspects of contracted migration by extending the Roy-Borjas (Borjas, 1987) model of migrant selection by allowing within skill wage dispersion and observability of both the current wage in the source and the potential wage in the destination prior to the migration choice. Choice of contracted mobility is thus not only based on the mean or the variance of source and destination region wage distributions but on a specific realizations from this distributions. Given current wage, low wage offers may not be enough to compensate for relocation costs whereas high offers may be. Thus, the deviation of the offered wage from the expected wage, job match quality in the destination, becomes a factor of relocation choice and there is positive selection on destination location job match quality. Given the offered wage, the lower the current wage is, the larger are the gains from relocation. Thus, job match quality in the source becomes a factor of relocation choice and there is negative selection on source location job match quality.

Empirically, the relationship between contracted migration and an empirical measure of a wage's deviation from its expectation, a residual from a wage regression, is confounded by the unobservable skills and abilities which have been documented to correlate with migration (Borjas, Kauppinen, & Poutvaara, 2018; Bartolucci, Villosio, & Wagner, 2018; Gould & Moav, 2016). This complicates the identification of job match quality in at least three ways. First, in the source, contracted migrants may not be negatively selected relative to stayers on pre-migration residuals if they are positively selected relative to stayers on their unobservable skills. The positive selection of the sample that receives job offers may mask any effect of negative job match quality when this group is compared to stayers. Second, in the destination, contracted migrants may be positively selected relative to destination region workers on post-migration residuals, not only due their better job match quality but also due to their more valuable unobservable skills. Third, contracted migrants may increase their residuals when relocating their labor supply not only due to good new job match in comparison to current job match but due to higher compensation for unobservable skills in their destination than in their source.

It is thus difficult to identify effects of job match quality by comparing migrants to stayers. An alternative comparison group for the contracted migrants are those who contract a job outside their location of residence, but who choose to (tele)commute.[1] Commuting is often, and increasingly, a substitute to migration in supplying labor outside the current region of residence. These two groups using different technology to supply labor outside their current region of residence make a relevant comparison for three reasons. First, the commuters, as contracted migrants have received and accepted a job offer, and, thus have self-selected to search interregionally and have been selected by employers. Second, when comparing commuters and migrants within source-destination pairs, commuters and migrants experience the same change in the compensation paid for unobservable skills due to the change of labor market. Third, those who choose to commute and those who choose to migrate are, however, different in the relocation costs they incur. Relocation costs, on the other hand, magnify selection effects. Comparison of commuters and migrants thus helps us study how the selection effects change in costs and helps us discern what sort of selection, selection on job match quality or selection on unobservable skills is magnified.

This paper contributes to the migrant selection literatue in four ways. First, I extend the Roy-Borjas migrant selection model to labor markets with within skill wage dispersions and to contracted migration. Second, I provide the first explicit evidence on selection on job match quality both in the source and destination locations. Third, I corroborate the selection on unobservable skills predicted by the Roy-Borjas model in the context of internal mobility. Fourth, I provide evidence on how mobility costs magnify the selection effects both on job match quality and unobservable skills. The findings imply that interpreting results on selection on residuals without taking job match quality into account underestimates positive selection on unobservable skills in the source and overestimates

[1] In the following, I use the term commuting but commuting thoughout the paper may contain telecommuting as well. Also, by commuters I refer to those who newly start commuting or change their commuting destination.

positive selection on unobservable skill in the destination.

My exploration of these ideas proceeds as follows. The next section positions the work into the existing literature. Section 3 presents the model selection of contracted migrants. Section 4 outlies the empirical approach. Section 5 introduces the data and the chosen empirical counterparts of theoretical concepts. Section 6 provides evidence of the selection of contracted migrants on job match quality relative to the commuters. Section 7 concludes. All proofs of lemmas and propositions are in the Appendix.

# 2 Literature

Since Hicks (1932), Schultz (1961) and Sjaastad (1962) the economic literature on migration has studied the role economic incentives play in migration. Borjas (1987) made explicit the heterogeneity in incentives stemming from differences in individual productivity and modelled the consequences of this heterogeneity on migrant selection. As a large fraction of variation in wages is not explainable by observable determinants of productivity, empirical studies of migrant selection have, in addition to studying migrant selection on observable determinants, studied selection on unobservable determinats of wages, or on their supposed approximate analogy, on residuals (Moraga, 2011; Abramitzky, 2009; Kaestner & Malamud, 2014; Borjas et al., 2018).

Gould and Moav (2016) take the analysis of selection on residuals forward by decomposing unobsevable skills into a location-invariant and location specific components. This is important as these two types of unobservable skill generate different selection patterns as Gould and Moav (2016) show. But unobservable skills, be it time-invariant or location speciific skills, are not the sole factors of wages unobservable to the researcher (Mortensen, 2003). I follow Gould and Moav (2016) in decomposing the unobservable variation i wagegs but into a location-invariant component and a job match component. This acknowledges that not all variation in residuals is due to unobservable skills and generates novel selection results in the context of contracted migration where the job match is observed both in the source and destination locations.

The dominant method of an econometric study of migration is to compare emigrants (Moraga, 2011; Kaestner & Malamud, 2014; Borjas et al., 2018) or immigrants to stayers. However, in controlling for the heterogeneity of migrants and stayers, the literature studying migrant selection has also compared incoming migrants from different locations (Borjas, 1987; Abramitzky, 2009), outgoing migrants to different locations (Hunt & Mueller, 2004; Dostie & Léger, 2009; Parey, Ruhose, Waldinger, & Netz, 2017) and also migrants working in different industries (Gould & Moav, 2016). On the other hand, the literature aiming to estimate the labor market returns to migration has distinguished wage changes due to job changes and wage changes due to location changes by comparing migrants to job movers (Bartel, 1979; Yankow, 2003; Ham, Li, & Reagan, 2011; Emmler & Fitzenberger, 2020). The comparison of migrants and commuters in this paper adds to this selection of settings by simultaneously controlling for the effect of job change but still comparing two groups that are both interregionally mobile.

The distinction between contracted and speculative migration was introduced by Silvers (1977). The evidence on the respective roles of speculative and contracted migration is, however, very scarce. Saben (1964) reports that 62 percent of high-skilled intercounty migrants had a job at hand when migrating whereas 38 percent of other migrants moved having accepted a job in the destination in the US in year 1962. Detang-Dessendre and Molho (1999), using a small survey of young first-time migrants from rural regions, report contracted migration to be more common than speculative migration. The share of contracted migration has likely increased and will be increasing. As job search more and more often occurs online job opportunities can more easily be searched and secured in distant labor markets. Policies regarding international immigration have been gearing toward favoring high-skilled migration and such policies often contain requirements of job contract at arrival (Kerr, Kerr, Özden, & Parsons, 2017). It is thus likely that contracted migration, rather than speculative migration is the dominant form of labor related migration, at least in the developed countries.

While not made explicit, the literature on labor related migration has often very likely studied contracted migration. For instance, by not allowing a long gap between job spells (Ham et al., 2011) or by defining migration as a change in job location (Emmler & Fitzenberger, 2020). Moreover, even without restrictions that increase the likelihood of contracted migration in the data, contracted migration is likely common in data given the likely role of contracted migration as the dominant form of labor related migration. Thus, for a more transparent interpretations of empirical results, the literature would benefit in making the distinction between contracted and speculative migration explicit.

# 3 Selection of Contracted Migrants

Consider three locations or labor markets indexed by $h = j, k, l$. The set of workers in location $h$ is defined by a skill distribution $\nu_i \sim \mathcal{N}(\mu_h^{\nu}, 1)$. The skills are time-invariant and perfectly transferable across locations.[2] Let worker $i$ reside in location $l$, work in location $j$, and potentially search for a job in location $k$. If worker $i$ searches for a job in $k$, she may receive a job offer. I call the subset $I$ of location $j$ workers that search for

[2] This is plausible within a country and this region-invariance of unobservables is also an assumption maintained, for instance, by Borjas, Bronars, and Trejo (1992) who study internal migration in the US. Note however, that while I assume that the absolute level of skill is location-invariant, I do not assume that the ranking of skills is location-invariant. A migrant with above mean skill on source location may have below mean skill in the destination if the destination location mean skill is higher than the source location mean skill. This may well occur, for instance, in case of rural to urban migration.

location $k$ jobs and receive a job offer the *population at risk of contracted mobility.* Only the workers in the population at risk of contracted mobility can relocate.[3] The annualized cost for a worker residing in $l$ of supplying labor in location $h$ is $\pi_{lh}$. The worker $i \in I$ maximizes a discounted stream of per period earnings $e_{ilh} = e(w_{ih}, \pi_{lh})$, $h = j, k$. Job offers $w_{ik}$ from location $k$ are sampled from distribution $F_{ik}$. The asset value of search for $i \in I$ is

$$rV_i(e_{ilj}) = e_{ilj} + \int \max\{0, V_i(e_{ilk}) - V_i(e_{ilj})\} dF_{ik}(w_{ik}) \tag{1}$$

where the second term on the right-hand side is the option value of optimal job acceptance behavior. This optimal behavior accepts a job offer if and only if $V_i(e_{ilk}) > V_i(e_{ilj})$ which, since $V_i$ is strictly increasing, requires $e_{ilk} > e_{ilj}$. Now letting $e(w, \pi) = w - \pi$, we have the job acceptance and migration condition

$$w_{ik} > w_{ij} + \pi_{ljk} \tag{2}$$

where $\pi_{ljk} = \pi_{lk} - \pi_{lj}$ is the cost of relocating labor supply from $j$ to $k$.[4]

Decompose worker $i$'s wage in location $h$ as $w_{ih} = \bar{\mu}_h + \rho_h(\nu_i - \mu_h^\nu) + q_{ih}$ where $\bar{\mu}_h$ is the compensation for mean skill, $\int w dF_{ih}(w) = \bar{\mu} + \rho_h(\nu_i - \mu_h^\nu)$ is the expected compensation for skill $\nu_i$ and the deviation $q_{ih}$ from $i$'s expected wage allows within skill wage dispersion. This wage dispersion may have many sources (see e.g. Mortensen (2003)) but here it is taken as exogenous. From the perspective of a worker, the source of such variation is probably of little importance and I interpret $q_{ih}$ generally as job match quality: the quality of $w_{ih}$ relative to $i$'s market expectation. Plugging this wage decomposition into (2), we have for the mobility condition

$$w_{ij} < w_{ik} - \pi \iff \mu_j - \mu_k + \pi < (\rho_k - \rho_j)\nu_i + q_{ik} - q_{ij}. \tag{MC}$$

where $\mu_h := \bar{\mu}_h - \rho_h \mu_h^\nu$ is the compensation paid for zero skill level in location $h$. Even if worker $i$ when deciding whether to accept the job offer $w_{ik}$ does not care how her current and offered wage can be decomposed, the decompositions determine the likelihood of worker $i$ receiving an acceptable job offer. Thus, to study the selection that the migration condition (MC) generates, specify heterogeneity in I on the different components of the

[3] Allowing speculative migration, a parameter restriction ensuring speculative migration does not occur requires specifying how speculative migrants land jobs in the destination. For instance, if they search in continuous time with unemployment income $b$ and discount rate $r$ and accept the first job offer that arrives at rate $\varphi$, a condition is $\frac{\int w dF_{ik}(w) - w_{ij}}{r} < \frac{w_{ij} - b + \pi_{ijk}}{\varphi}$ (See Lemma 1 in Section A.1). Adding risk aversion makes speculative migration less attractive.

[4] Interpreting $w$ as the logarithm of wage, a formally equivalent model follows from a time-equivalent labor supply cost $e(w, \pi) = w(1 - \pi)$ as $ln[w(1 - \pi)] \approx lnw - \pi$.

wage and see what types of agents satisfy the migration condition. Let

$$\nu_i|i \in I \sim \mathcal{N}(\mu_\nu, \sigma_\nu^2). \tag{3}$$

As those who are in a position to choose whether to relocate or not may be nonrandomly selected, allow $\mu_\nu \neq \mu_j^\nu$, $\sigma_\nu^2 \neq 1$. The values of individual effects in the source and destination regions in I are then distributed as[5]

$$\begin{bmatrix} \nu_{ij} \\ \nu_{ik} \end{bmatrix} |i \in I \sim \mathcal{N}\left( \begin{bmatrix} \rho_j \\ \rho_k \end{bmatrix} \mu_\nu, \begin{bmatrix} \rho_j^2 & \rho_j\rho_k \\ \rho_j\rho_k & \rho_k^2 \end{bmatrix} \sigma_\nu^2 \right). \tag{4}$$

For within skill wage dispersion,

$$\begin{bmatrix} q_{ij} \\ q_{ik} \end{bmatrix} |i \in I \sim \mathcal{N}\left( 0, \begin{bmatrix} \sigma_j^2 & 0 \\ 0 & \sigma_k^2 \end{bmatrix} \right). \tag{5}$$

The population at risk of contracted mobility may have been already selected on job match quality. Low job match quality increases the relative payoff from job search and if only job-seeker-employer meetings with high job match quality lead to job offers, then the population at risk of contracted mobility is negatively selected on source location job match quality and positively selected on destination region job match quality. As we will see, such selection is qualitatively equivalent to the selection that (MC) generates and, thus, abstracting from the selection generated by job search and hiring choices simplifies without affecting the theoretical results.[6]

**Proposition 1.** *With heterogeneity specified by (3), (4), and (5), (i) the expected pre-mobility wages of migrants are*

$$\begin{aligned} E[w_{ij}|(\text{MC})] &= \bar{\mu}_j - \rho_j\mu_j^\nu + E[\rho_j\nu_i|(\text{MC})] + E[q_{ij}|(\text{MC})] \\ &= \bar{\mu}_j + \rho_j(\mu_\nu - \mu_j^\nu) + \frac{1}{\sigma_\Delta}\sigma_\nu^2(\rho_k - \rho_j)\rho_j\lambda(z) - \frac{1}{\sigma_\Delta}\sigma_j^2\lambda(z) \\ &= \bar{\mu}_j + \rho_j(\mu_\nu - \mu_j^\nu) + \frac{1}{\sigma_\Delta}\big(\sigma_\nu^2(\rho_k - \rho_j)\rho_j - \sigma_j^2\big)\lambda(z), \end{aligned} \tag{6}$$

[5] $Var[\nu_{ik}] = Var[\rho_k\nu_i] = \rho_k^2 Var[\nu_i] = (\rho_k\sigma_\nu)^2$, $Cov(\nu_{ik}, \nu_{ij}) = E[\nu_{ik}\nu_{ij}] - E[\nu_{ik}]E[\nu_{ij}] = \rho_k\rho_j(E[\nu_i^2] - E[\nu_i]^2) = \rho_k\rho_j Var[\nu_i] = \rho_k\rho_j\sigma_\nu^2$.

[6] The selection into the population at risk of contracted mobility may also generate correlation between job match quality and skills. For clarity, I ignore these potential correlations as well here. For formalisation and more discussion on the selection into the population at risk of contracted mobility, see companion paper available from the author.

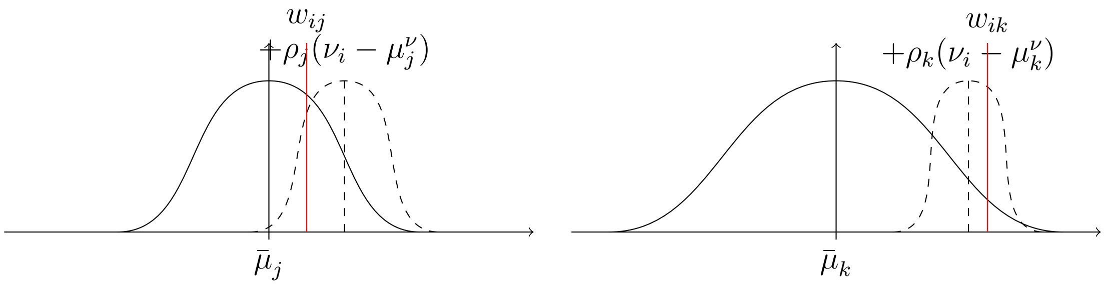


Figure 1: Source and destination wages of speculative and contracted migrants.

*and (ii) the expected post-mobility wages of migrants are*

$$\begin{aligned} E[w_{ik}|(\text{MC})] &= \bar{\mu}_k - \rho_k\mu_k^\nu + E[\rho_k\nu_i|(\text{MC})] + E[q_{ik}|(\text{MC})] \\ &= \bar{\mu}_k + \rho_k(\mu_\nu - \mu_k^\nu) + \frac{1}{\sigma_\Delta}\sigma_\nu^2(\rho_k - \rho_j)\rho_k\lambda(z) + \frac{1}{\sigma_\Delta}\sigma_k^2\lambda(z) \\ &= \bar{\mu}_k + \rho_k(\mu_\nu - \mu_k^\nu) + \frac{1}{\sigma_\Delta}\big(\sigma_\nu^2(\rho_k - \rho_j)\rho_k + \sigma_k^2\big)\lambda(z), \end{aligned} \tag{7}$$

*where* $\sigma_\Delta^2 \coloneqq \sigma_k^2 + \sigma_j^2 + (\rho_k - \rho_j)^2\sigma_\nu^2$, $\lambda(z) \coloneqq \phi(z)/(1 - \Phi(z))$ *is the inverse Mill's ratio, where* $\phi$, *and* $\Phi$ *denote the density and distribution functions of the standard normal, respectively, and*

$$z \coloneqq \frac{1}{\sigma_\Delta}\left(\mu_j - \mu_k + \pi - (\rho_k - \rho_j)\mu_\nu\right). \tag{8}$$

Job match quality enters the expected pre-mobility disturbance with a negative sign indicating negative selection on job match quality in the source. Given a wage offer, the lower the current job match quality is, the larger is the gain of acceptance and relocation and, thus, the more likely relocation costs are covered. Hence, those with low current job match quality are more likely to relocate than those with high job match quality.

Job match quality enters the expected post-mobility disturbance with a positive sign indicating positive selection on job match quality in the destination. Low wage offers may not be enough to compensate for the costs of relocation whereas high job offers may be. Hence, those realizing job offers with high job match quality are more likely to migrate. Figure 1 illustrates the effect of wage dispersion. Since the difference between destination and source wages have to compensate for the migration cost, a typical contracted migrant is negatively selected in her skill specific wage distribution in the source and positively selected in her skill specific wage distribution in the destination.

The expected pre- and post mobility wages nest the migrant mean wages computed by Borjas (1987) and the selection results with respect to skills are as reported therein.[7] Whether selection on job match quality or unobservable skill dominate selection depends on the relative magnitudes of the spreads of within skill wage disperions and and distributions of expected wages. If there is not much variation in the distributions of expected wages across locations, the locational variation in $\mu_h$ and $\rho_h$ may not be enough to cover the relocation costs. With enough wage dispersion in both location, however, a change in job match quality may still incentivize relocation and selection on job match quality dominates. This is likely to be especially the case when studying internal migration. On the other hand, if there are large locational differences in wage distributions relative to within skill wage dispersion, then selection on unobservable skills dominates.

We can interpret the model's relation to the Roy-Borjas migrant selection model in two ways. First, the extension can be seen as allowing within skill wage dispersion. With degenerate within skill wage dispersion each worker is always compensated exactly the value of their skills and there is no selection on job match quality. Second, if we interpret the wages in Roy-Borjas model as means of the within skill wage dispersions, the extension is that the migrants observe their exact wage realizations in the source and destination before their relocation choice. Not observing an exact wages the choices are made on expected wages which with risk neutrality is again equivalent to the wages in the model of degenerate within skill dispersion.

The expected wages of Roy-Borjas model can, however, also be interpreted more in line of the human capital approach to migration of Sjaastad (1962) as the expectations of discounted income streams. Such an interpretation allows then wage expectation to contain the option value of further job-to-job mobility. In contrast, interpreting the offered destination wage $w_{ih}$ as an intertemporal utility stream may seem to restrict the model of contracted mobility to no further job-to-job mobility. However, the mobility condition (MC) was derived from on-the-job search model that clearly allows on-the-job mobility in the destination. The crux is that while the Sjaastad's human capital approach

[7] Setting $\mu_\nu = q_{ik} = q_{ij} = 0$, (6) can be written as

$$E[w_{ij}|(\mathrm{MC})] = \mu_j + \frac{sd(\nu_{ik})sd(\nu_{ij})}{sd(\nu_{ik}-\nu_{ij})}\left(\frac{sd(\nu_{ik})}{sd(\nu_{ij})} - \frac{Cov[\nu_{ik},\nu_{ij}]}{sd(\nu_{ik})sd(\nu_{ij})}\right)\lambda(z), \tag{9}$$

and (7) can be written as

$$E[w_{ik}|(\mathrm{MC})] = \mu_k + \frac{sd(\nu_{ik})sd(\nu_{ij})}{sd(\nu_{ik}-\nu_{ij})}\left(\frac{Cov[\nu_{ik},\nu_{ij}]}{sd(\nu_{ik})sd(\nu_{ij})} - \frac{sd(\nu_{ik})}{sd(\nu_{ij})}\right)\lambda(z), \tag{10}$$

where $sd(\cdot) := \sqrt{Var[\cdot]}$ and

$$z = \frac{1}{sd(\nu_{ik}-\nu_{ij})}\left(\mu_j - \mu_k + \pi\right). \tag{11}$$

giving us the truncated expectations as formulated by Borjas (1987).

and the Roy-Borjas model typically assume irreversibility of migration choice, containing the choice of staying, here staying does not preclude mobility in later time. Later job opportunities in the potential destination remain available even after declining a job offer and, thus, in comparing staying and mobility, the on-the-job mobility prospects in the destination cancel out.

Borjas (1987) categorizes the possible selection patterns his model is able to produce into positive selection, negative selection and refugee sorting. While positive and negative selection have found use in further theoretical and empirical work, the pattern of refugee sorting, in its somewhat narrow interpretation of high-skill but low wage emigrants suppressed in communist countries immigrating to non-communist countires and earning above average wages in a market economy has been left as a mere mathematical possibility. Selection on job match quality here as it sets the correlation of source and destination location match qualities to zero, corresponds to the case of Borjas' refugee sorting. Here, workers are leaving jobs that do not pay them what they would expect to earn in the markets into jobs that do, as if refugees fleeing job paying jobs into high paying jobs. This interpretation makes the refugee sorting pattern relevant also in studies focusing on market economies.

# 4 Empirical Strategy

I now reformulate the model of selection of constracted migrants as a model of disturbances where disturbances are the sum of the value of unobservable skills and job match quality. This is the empirically relevant model as the value of unobservable skills and job match quality are not separately observed, a problem already raised in the context of migrant selection literature, for instance, by Borjas et al. (2018) I then discuss how this makes the dominant method of selection analysis, the comparison of migrants to stayers obsolete in identifying the selection on job match quality and propose an alternative strategy based on variation in mobility costs otherwise arguably similar groups of mobile workers.

## 4.1 A Model of Disturbances

Job match quality is defined as the wage's deviation from the worker's expected wage. The corresponding empirical measure is the residuals from wage regressions. Unobservable determinants of productivity cannot be included as regressors and, thus, the resulting residuals reflect both the value job match quality and the price of unobservables. The residuals of the wage regressions are assumed to be a sum of a term that captures the effect of unobservable abilities and a term that captures the effect of job match quality. This corresponds to an error decomposition devised by Flinn (1986) and Garen (1989). Such a factor model has also been used in migration literature (Borjas et al., 1992; Gould

& Moav, 2016; Bartolucci et al., 2018), but there the component $q_{ik}$ has been interpreted as a transitory location specific productivity shock. Here locations are tied with jobs such that a location specific productivity shock is an employer specific productivity shock, i.e. job match quality.

For a model of residuals, decompose the individual skills into observable and unobservable skills as $w_{ih} = \bar{\mu}_{ih} + u_{ih}$ where $\bar{\mu}_{ih}$ is the component of wage that can be predicted by observable variables including location fixed effect and controlled for in the empirical analysis and $u_{ih}$, the sum of the value of unobservables and job match quality, is the disturbance. With an abuse of notation, reinterpret $\nu_i$ as the unobservable skills of $i$ and $\mathcal{N}(\mu_\nu, \sigma_\nu^2)$ as the distribution of unobservable skill in the population at risk of contracted mobility such that $u_{ih} = \rho_h(\nu_i - \mu_h^\nu) + q_{ih}$.

**Proposition 2** (A model of disturbances)**.** *The theoretical conditional expectation of migrants' pre-mobility disturbances is*

$$E[u_{ij}|(\mathrm{MC})] = \rho_j(\mu_\nu - \mu_j^\nu) + \frac{1}{\sigma_\Delta}\big(\sigma_\nu^2(\rho_k - \rho_j)\rho_j - \sigma_j^2\big)\lambda(z_{ijk}), \tag{12}$$

*and post-mobility disturbances is*

$$E[u_{ik}|(\mathrm{MC})] = \rho_k(\mu_\nu - \mu_k^\nu) + \frac{1}{\sigma_\Delta}\big(\sigma_\nu^2(\rho_k - \rho_j)\rho_k + \sigma_k^2\big)\lambda(z_{ijk}), \tag{13}$$

*and the change in disturbances is*

$$\begin{aligned} E[u_{ik} - u_{ij}|(\mathrm{MC})] = {} & (\rho_k - \rho_j)\mu_\nu - \rho_k\mu_k^\nu + \rho_j\mu_j^\nu \\ & + \big(\sigma_\nu^2(\rho_k - \rho_j)^2 + \sigma_k^2 + \sigma_j^2\big)^{\frac{1}{2}}\lambda(z_{ijk}). \end{aligned} \tag{14}$$

*where $\sigma_\Delta^2 \coloneqq \sigma_k^2 + \sigma_j^2 + (\rho_k - \rho_j)^2\sigma_\nu^2$, $\lambda(z) \coloneqq \phi(z)/(1 - \Phi(z))$ is the inverse Mill's ratio, where $\phi$, and $\Phi$ denote the density and distribution functions of the standard normal, respectively, and*

$$z_{ijk} \coloneqq \frac{1}{\sigma_\Delta}\left(\mu_{ij} - \mu_{ik} + \pi_{jk} - (\rho_k - \rho_j)\mu_\nu\right). \tag{15}$$

The presence of unobservables confounds the relationship between job match quality and disturbances in two ways. First, as disturbance is the sum of job match quality and unobservables, those with higher unobservables have higher disturbances. The first terms on the right-hand sides of (12), (13), and (15) capture this. Second, if the destination region compensates the unobservable factors well relative to the source region, it generates acceptable job offers disproportionally for those whose skill composition has a high weight on unobservables selecting them into mobility toward this region. The first terms in the parenthesis of the second terms of (12), (13), and (15) capture this.

## 4.2 Commuters and Migrants

The population at risk of contracted mobility is likely not a random sample of workers in location $j$. Prior to being in a position of choosing between contracted mobility and staying, these workers have chosen to search for a job interregionally and have received a job offer. Both of these hurdles are likely to select positively on unobservable skills. Looking at (12) it is, thus, clear how even if there is negative selection on job match quality in the source, the positive selection of the population at risk of mobility, $\mu_\nu > \mu_j^\nu$, may mask this.[8] Thus, the positive selection of those who are in a position to choose whether to migrate or not confounds the effect of job match quality when comparing the mobile to the stayers. Looking at (13), on the other hand, positive selection on disturbances in the destination may be due to both job match quality and unobservable skills and, thus, observing positive selection on residuals in the destination is not evidence for selection on job match quality. Thus, the dominant method of the migrant selection literature, comparison of migrants and stayers is not helpful in identifying selection on job match quality.

Hence, for a more relevant comparison, I study two groups of workers who relocate their labor supply but employ different mobility technologies: those that migrate and those that start commuting. I call these two groups of mobile workers *contracted migrants* and *commuters*. A contracted migrant and a commuter make a relevant comparison since they have both chosen to search for jobs outside their region of residence and have both accepted a job offer. Hence, this comparison controls for the selection into the group of population at risk of contracted mobility. Furthermore, when restricting comparisons to source-destination pairs, the contracted migrants and commuters experience the same change in the compensation for unobservable skills.

The contracted migrants and commuters, however, incur very different costs of mobility. The choice between migration and commuting is based on the relative costs of these two mobility modes. These costs depend on a variety of factors such as access to public transport or a car, housing status and family. The employed mobility mode is thus seen here as a statistic summarizing these different cost structures. Presumably, migrants, who relocate both their residence and work locations incur on average larger costs than the commuters who relocate only their work locations.

On a more fundamental level, the identification of the effect of job match quality thus is based on variation in relocation costs. Figure 2 illustrates how mobility costs magnify

[8] For formalisation and more discussion on the selection into the population at risk of contracted mobility, see companion paper available from the author.

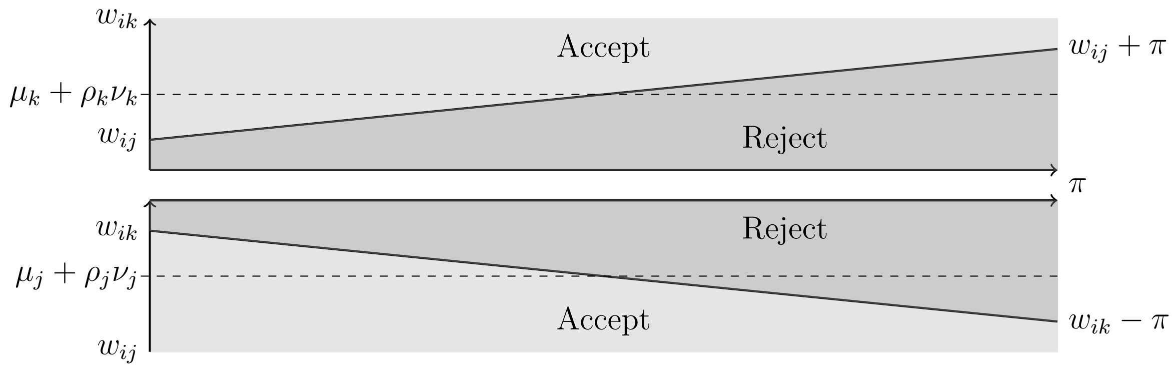


Figure 2: Reservation and inverse reservation wages as a function of relocation cost.

selection on job match quality. Looking at the upper half, the vertical axis tracks the support of $i$'s within skill wage dispersion in location $k$. The worker $i$'s reservation wage $w_{ij}+\pi$ is an element on this support and increasing in mobility cost. Given current wage, $w_{ij}$ the larger is the mobility cost, the smaller subset of possible location $k$ wages are acceptable and the larger is job match quality required for $i$ to accept a job offer and migrate. Looking at the lower half, the vertical axis tracks the support of $i$'s within skill wage dispersion in location $j$. Ther worker $i$'s inverse reservation wage $w_{ik}-\pi$ is an element on this support and decreasing in mobility cost. Given job offer, the larger is the mobility cost, the lower is current job match quality required for this offer to be acceptable. Summing up, the gap in source and destination location job match qualities required for relocation increases in relocation costs. Relocation costs create what Borjas (1987) called the scale effect. When costs are high, also a large changes in wages is required. With higher costs then, only those with largest changes in wages are mobile.

**Proposition 3.** *(i) The expected difference of the pre-mobility residuals of the contracted migrants and commuters is*

$$E[u_{ij}|(\mathrm{MC}),\mathrm{migrant}]-E[u_{ij}|(\mathrm{MC}),\mathrm{commuter}]$$
$$=\rho_j(\mu_\nu^m-\mu_\nu^c)+\frac{1}{\sigma_\Delta}\big(\sigma_\nu^2(\rho_k-\rho_j)\rho_j-\sigma_j^2\big)[\lambda(z_{ijk}^m)-\lambda(z_{ijk}^c)] \tag{16}$$

*(ii) The expected difference of the post-mobility residuals of the contracted migrants and commuters is*

$$E[u_{ik}|(\mathrm{MC}),\mathrm{migrant}]-E[u_{ik}|(\mathrm{MC}),\mathrm{commuter}]$$
$$=\rho_k(\mu_\nu^m-\mu_\nu^c)+\frac{1}{\sigma_\Delta}\big(\sigma_\nu^2(\rho_k-\rho_j)\rho_k+\sigma_k^2\big)[\lambda(z_{ijk}^m)-\lambda(z_{ijk}^c)] \tag{17}$$

*(ii) The expected difference of the changes in residuals of the contracted migrants and*

*commuters is*

$$E[u_{ik} - u_{ij}|(\text{MC}), \text{migrant}] - E[u_{ik} - u_{ij}|(\text{MC}), \text{commuter}]$$
$$= (\rho_k - \rho_j)(\mu_\nu^m - \mu_\nu^c) + \left(\sigma_\nu^2(\rho_k - \rho_j)^2 + \sigma_k^2 + \sigma_j^2\right)^{\frac{1}{2}} [\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)]. \quad (18)$$

Note that in the comparison of contracted migrants and commuters, the unobservable location mean skill level effects $\rho_h \mu_h^\nu$ and their changes cancel out. Now, if the commuters and migrants in the population at risk of mobility are approximately equal in their mean unobservable skills, the first term in expressions (16), (17) and (18) can be ignored and the comparision of contracted migrants and commuters reveals the sign of the multiplier of $\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)$ and so reveals whether variation in mobility costs magnify selection on job-match quality or on unobservable skills. Section 6.2.3 develops this argument formally.

# 5 Data and Empirical Definitions

In this section, I describe the data I use, define the two mobility groups of comparison, contracted migrants and commuters, and the outcome variables of interest, wages, predicted wages and residuals.

## 5.1 Data

I use total population annual individual level data compiled from various administrative registers provided by Statistics Finland.[9] For each observation, I assemble data from three periods: the year before the potential mobility event $t-1$, the year of the potential mobility event, $t$, and one year after the potential mobility event $t+1$ (Figure 4 uses data all the way to the period $t-5$). As the theory in Section 3 models the behavior of salaried workers well attached to the labor market, I use a sample of such workers. The sample for a year $t$ consists of those aged weakly between 30 and 60 in year $t$ and who were salaried employees, alive and in Finland in the last week of years $t-1$, $t$ and $t+1$, who had positive earnings and zero registered unemployment days in these years and to whom all variables used in the analysis are observed. All residents and workers in and all migrants and commuters to and from the Åland Islands and students and retirees using information on the longest principal activity during the year are excluded. The age restriction is to reduce mobility of students, the young, the first-time movers and mobility that may occur with retirement in the sample. I also, for reasons explained later, require

[9]The data are available in data sets called FOLK modules. To build the data used I combine information from FOLK Basic data, FOLK Employment and FOLK Cohabitation modules. See https://taika.stat.fi/en/.

that employer and establishment does not change between years $t$ and $t + 1$. I pool data such that $t \in \{2011, 2012, 2013, 2014\}$.

## 5.2 Empirical Mobility

The contracted mobile are defined as those who change the location of their employment. This definition of mobility aims to capture labor related migration by requiring a change in job location but also captures changes in commuting destinations. The stayers are those who are not mobile and do not change the location of their residence. To remove some speculative and residential migration from the sample, all observations that are not classified as stayers or mobile are discarded. The mobile are further categorized as contracted migrants and commuters according to their post-mobility residential locations. As I study contracted migration, I do not aim to separate speculative and contracted migration, which is an infeasible task with annual administrative data, but to restrict the sampel such that the mobility events observed are very likely contracted.

First, define job movers as those whose postal code area or municipality of work place changes between years $t-1$ and $t$ and, to capture employment-to-employment transitions, who have zero days in registered unemployment in year $t$. Then, for all job-movers, I compute the distance from the location of their residence in year $t - 1$ to the location of their work place in year $t$. The distance to the location of the new job is defined differently for those who eventually commute and those who eventually migrate: For the commuters, the data contains information of the commuting distance in year $t$ as an Euclidean distance between job and residence location with the accuracy of 250m x 250m squares computed in Statistic Finland. As commuters, by definition, do not change their residence, their commuting distance equals the distance between the location of their residence and the location of their new job.

For the contracted migrants, the commuting distance in year $t$ does not equal the distance to new job location. For them, the distances to new job location are computed as the Euclidean distances between the centroids of the postal code area they resided in year $t - 1$ and the postal code area of new job in year $t$. The different accuracies used in measuring the distance to new job location are unlikely to be an issue as the distances to new job location are typically longer for the migrants than to commuters. Hence, relatively, distances measured using postal code areas are probably not subject to larger measurement errors than distances computed using the 250m x 250m squares. On the other hand, Euclidean distances between postal code area centroids are subject to most severe measurement errors for small distances. Using information on commuting distances for these distances avoids this problem. All location information is from the last week of the year similarly to employment information.

Given the distance to the location of new job for all job-movers, the extent of dislocation that qualifies as mobility is defined by a distance threshold. Using distance threshold in place of changes in administrative regions has the advantage of removing the issue of mobility propensities depending on the distance to border, effectively leading to an omitted variable bias if the distance to border is not controlled for. The threshold distance for mobility is set to 50 kilometers in the main analysis. The threshold of 50 kilometers is somewhat arbitrary, chosen to ensure that there are both commuters and migrants in the sample. I study the robustness of the results to different thresholds.

Categorization of the mobile into contracted migrants and commuters is based on the post-mobility residential locations relative the pre-mobility residential locations. If the residential location does not change, a mobile individual is defined as commuter. Otherwise, she is a migrant. I leave out from the mobile those who migrate but such that the distance between the location of their residence and their job increases as these moves are likely motivated by factors unrelated to labor market outside the scope of the theory of Section 3.

A concern in dividing the mobile into commuters and migrants is that commuting often is a temporary solution: commuting may simply be migration with a lag. For instance, rigidities in the housing market may force the individual to commute for a while even if she later migrates. Thus, defining commuting and migration by the residential locations on the year of mobility may overestimate the number of commuters due to the lags in residential adjustment. Ideally, migrants could be distinguished from commuters by studying their residential locations in all years after the mobility event. However, requiring commuting, say, until after $n$ years after the mobility event would drop all mobile observations whose employment lasted fewer than $n$ years from the analysis. The lost observations would clearly be selected nonrandomly potentially leading to sample selection problems. To avoid dropping observations, the required years of commuting after the mobility event could be defined individually by the length of the employment. Then, however, the condition for commuting would be less stringend for those with shorter employments potentially leading to the durability of employment to increase the likelihood of commuting mechanically by definition. As a compromise, commuters are defined to be those that have still not changed their original residence location in year $t + 1$.[10]

The choice between commuting and migration is, of course, not truly binary. The migrants have a strictly positive post-mobility commuting distance as well. Figure 3, however, shows, that the fraction of approximate corner solutions is large enough to render the binary classification of migrants and commuters a reasonable approximation.

[10] Note that the post-mobility earnings are computed using the earnings information in year $t + 1$ and by requiring that the employer does not change between years $t$ and $t + 1$. See Section 5.3. Given this constraint, the definition of mobility does not drop observations from the analysis.

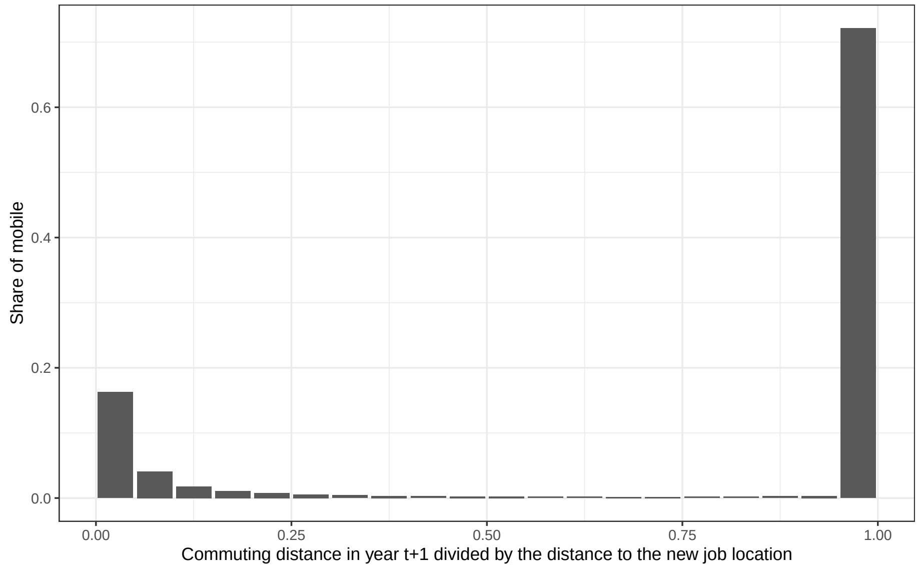


Figure 3: **New commuting distance as a fraction of distance to new job location.**
*Notes: Distribution of commuting distances in year $t+1$ relative to the distance between the location of residence in year $t-1$ and location of work in year $t$ among the mobile. See definitions of the mobile and the distances in Section 5.2. For data, see Section 5.1*

## 5.3 Empirical Job Match Quality

Residuals from wage regressions estimate disturbances. First following models are estimated in complete sample for each year $t-1$, $t$ and $t+1$:

$$\hat{w}_{ih} = \hat{E}_{ih}[w] + u_{ih} = x_i'\gamma + \mu_h + u_{ih}. \tag{19}$$

where $x_i'$ contains observable determinants of productivity and $\mu_h$ is a municipality effect. Residuals are then computed as $\hat{u}_{ih} \coloneqq \hat{w}_{ih} - \hat{\mu}_h - x_i'\hat{\gamma}$, where variables with hats indicate estimated quantities.

### 5.3.1 Wages

Earnings are observed annually and wages are computed as the ratio of earnings to employment days. As only annual earnings and employment days are observed, earnings cannot be allocated to jobs in source and destination locations that the mobile hold in the year of mobility $t$. Thus, year $t$ information cannot be used to construct the pre- or post-mobility wages. Pre-mobility wages are measured using year $t-1$ information. For the contracted post-mobility wage I use year $t+1$ information. This does not come without problems. Whereas the contracted wage is determined as mobility occurs, the wages in year $t+1$ may partly be a consequence of a certain mobility mode choice. This may happen, for instance, if source and destination location labor markets differ in their on-the-job search possibilities putting migrants and commuters in different positions with respect to their on the job search options. I remove this problem by restricting the analysis to those mobile workers who do not change their employer or establishment before the end of year $t + 1$. Within firm career advancement and wage growth likely does not depend on the residence location and is, thus, not suspect to this concern. Further, convex combinations of year $t+1$ and $t$ wages are used as alternative measures of the contracted wage for robustness.

The interpretation of year $t + 1$ wage as a determinant of mobility follows from the underlying assumption that all studied migration is contracted, that is, that the wage in the destination is observed prior to migration choice. As described in Section 5.2 the mobile are defined so as to highly likely be contracted migrants. Especially, the requirement of zero days in registered unemployment in the year of mobility is likely to exclude speculative migrants in the sample of workers with solid labor market histories to whom claiming unemployment benefits in case of unemployment is well incentivized. Nevertheless, it is possible that the sample contains speculative migrants who gained employment in the destination quickly and without drawing unemployment benefits. For these workers, the destination wage is determined after migration. However, speculative migrants transition from unemployment to employment and are, thus, in a weaker position to realized high job match qualities than the commuters who transition from employment to employment. Thus, if there were speculative migrants in the sample, they would likely bias the estimated positive selection on job match quality in the destination toward zero. Also note that the whole sample, including the stayers, are restricted to those who do not change their employer between years $t$ and $t + 1$ to avoid any conditioning of mobility classification on specific employment paths.

### 5.3.2 Wage Predictors

In computing the residuals $x_i'$ contains gender, whether born in Finland, age, age squared, indicators for level of education, indicators field of education, indicators for occupation and indicators for industry.

If commuting time and labor supply compete from the same finite endowment of time, then keeping the wage fixed, the daily income should decrease in commuting time. As the empirical measure of a wage is strictly speaking daily income, the concern is that wages are underestimated for those who commute. If commuting costs are then left in the error term when computing the residuals, the commuters would have lower post-mobility residuals than the contracted migrants simply because they have higher commuting costs. However, if wages compensate for commuting costs, then removing commuting costs from the residuals would rather remove variation in job match quality that this paper aims to explain.

That is, if the empirical measure of wage has a negative association with commuting

| | Mobile | | Stayers |
|---|---|---|---|
| | Commuters | Migrants | |
| Females | 0.32 | 0.47 | 0.52 |
| Age, year $t$ | 44.7 | 39.7 | 45.5 |
| Born abroad | 0.03 | 0.06 | 0.04 |
| **Education, year $t$** | | | |
| Basic education | 0.08 | 0.06 | 0.10 |
| Secondary education | 0.34 | 0.28 | 0.41 |
| Tertiary education | 0.55 | 0.62 | 0.48 |
| Doctoral or equivalent | 0.03 | 0.03 | 0.02 |
| **Work, year $t-1$** | | | |
| Tenure in current job, days | 2,209.6 | 1,601.6 | 2,878.1 |
| Employment days | 360.83 | 359.05 | 363.22 |
| Unemployment days | 0 | 0 | 0 |
| Log wage | 4.75 | 4.60 | 4.59 |
| **Mobility experience, year $t-1$** | | | |
| Migration experience | 0.18 | 0.38 | 0.10 |
| Commuting experience | 0.79 | 0.59 | 0.45 |
| **Family, year $t-1$** | | | |
| Spouse working | 0.68 | 0.45 | 0.65 |
| Living alone | 0.05 | 0.08 | 0.06 |
| Living with spouse | 0.80 | 0.57 | 0.76 |
| Children | 0.60 | 0.42 | 0.58 |
| **Housing, year $t-1$** | | | |
| Right of occupancy dwelling | 0.01 | 0.02 | 0.02 |
| Rents the dwelling | 0.10 | 0.36 | 0.15 |
| Owns the dwelling | 0.88 | 0.60 | 0.82 |
| **Distance to (new) job, km** | | | |
| Mean | 147.77 | 216.68 | 14.86 |
| Median | 104.38 | 159.60 | 6.86 |
| **Number** | 30,896 | 12,172 | 3,131,309 |

Table 1: **Descriptive statistics by mobility group.** *Notes: Means for continuous variables unless otherwise mentioned and shares for categorical variables. For the data, see Section 5.1. See Section 5.2 for the definitions of commuters, contracted migrants and stayers. See Section 5.1 for the sample restrictions.*

costs, we have the first case and we should include commuting costs into the wage regression but if the the association is zero or positive, we have the second case and we should not include commuting costs into the wage regression. The most straightforward measure of commuting costs in the data is the commuting distance. The estimated coefficient for the commuting distance in the wage regression in the whole sample of the mobile and stayers is significantly positive (not reported). Thus, when computing the residuals no measures of commuting costs are included in the wage regressions.

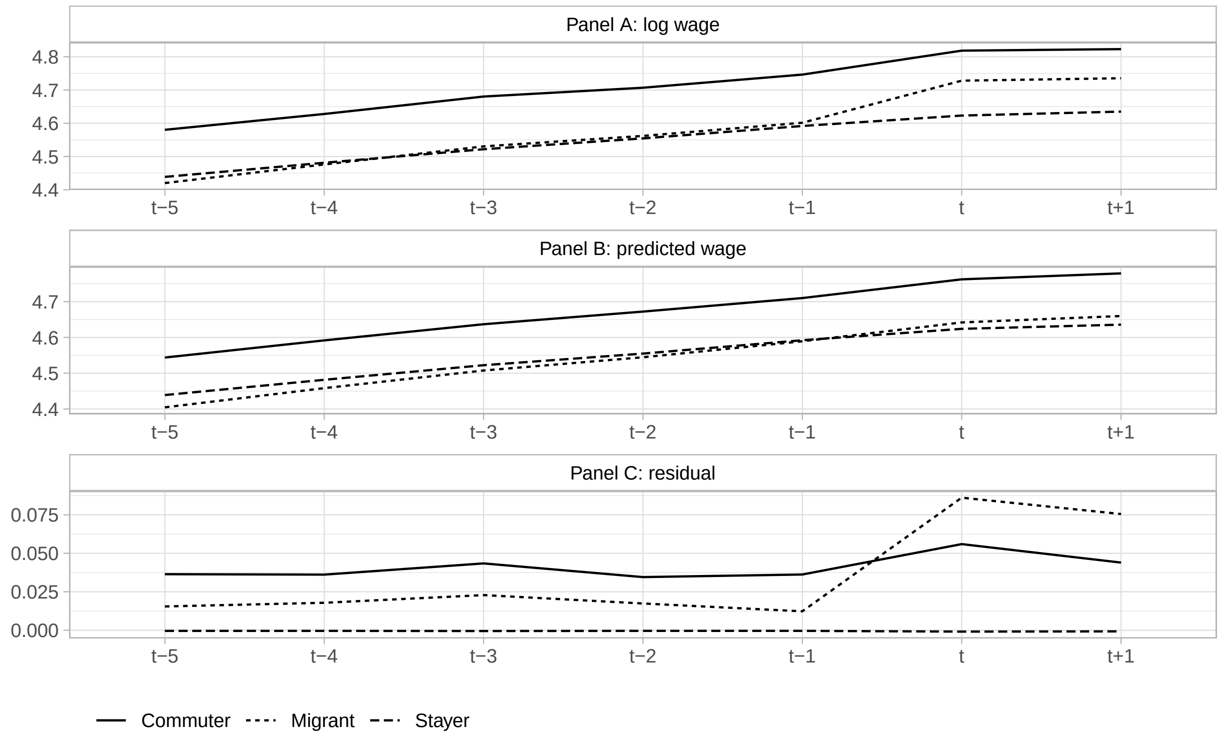


Figure 4: **Wage decompositions of commuters, contracted migrants and stayers.** *Notes: For the data, see Section 5.1. See definitions in Section 5.2. For the wage models, see Section 5.3. Mobility potentially occurs between time points $t-1$ and $t$.*

## 5.4 Descriptive Statistics

Table 1 presents the descriptive statistics of the sample by subsample. The commuters are more often males and older than the contracted migrants. The contracted migrants are clearly less constrained by family: they less often have a working spouse and children and they more often live alone. The contracted migrants are also less constrained by housing: they more often rent and less often own their dwelling. Contracted migrants tend to have higher education that the commuters. Choices between commuting and migrating seem to autocorrelate as well: previous migration experience is higher among the contracted migrants whereas previous commuting experience is more prevalent among the commuters.

# 6 Empirical results

## 6.1 Selection on Residuals

Figure 4 presents the wage paths of commuters, contracted migrants and stayers in Panel A. While we observe the wages of stayers growing along a stable path, for the mobile the growth of the wages exceed this stable growth path during the period of potential mobility. Moreover, the contracted migrants increase their wages more than the commuters. The decomposition of wages into their predicted and residual components, depicted in Panels B and C, respectively, shows that the residuals drive the pattern observed in the Panel A. While the contracted migrants have lower residuals than the commuters prior to mobility in the source, they have higher residuals than the commuters after mobility in the destination. The first column in Table 2 shows that the unconditional differences in residuals pre- and post mobility are statistically different from zero with signs as in Figure 4.

The unconditional residual means of contracted migrants and commuters in Figure 4 Panel C form a pattern consistent with negative selection on job match quality in the source and positive selection on job match quality in the destination. Define an indicator on the set of the mobile, who satisfy (MC), $D_i = \mathbb{1}(\text{migrant})$, such that $D_i = 1$ if $i$ is a contracted migrant and $D_i = 0$ if $i$ is a commuter. A model of residuals, for $h = j, k$ is

$$\hat{u}_{ih} = \mu_{jkl} + \lambda_t + \alpha_1 \hat{E}_{ij}[w] + \alpha_2 \hat{E}_{ik}[w] + \tau D_i + \alpha_3 \Delta d_i + \gamma w_i + \varepsilon_i \tag{20}$$

where $\tau := E[u_{ih}|D = 1] - E[u_{ih}|D = 0]$. The triadic residence-source-destination municipality (LAU 2 region) fixed effect $\mu_{jkl}$ restricts the comparison to commuters and migrants within residence-source-destination triplets.[11] Due to selection on unobservable skills the pre-mobility residuals of the outgoing mobile workers depend on the destination location and post-mobility residuals of the incoming mobile workers on the source location. Thus, to control for the effect of destination region in the model of pre-mobility residuals and the effect of source location in the model of post-mobility residuals, the triadic fixed that interacts the source and destination locations is more appropriate than monadic source and destination location fixed effects. The triadic fixed effect also controls for the differences in residence, source and destination region characteristics, e.g. prices and availability of housing, and for transport infrastructures connecting the residence and destination locations. The time fixed effect $\lambda_t$ controls for common year effects. The change in distance to job location $\Delta d_i$ is the difference between the distance to the distance to job in period $t - 1$. To control for potential losses in firm, industry or occupation specific human captial, the vector $w_i$ contains all the main and interaction effects of indicators of employer, industry of occupation changes and $\gamma$ is a conformable vector of coefficients.

The indicator $D_i$ can be thought of as being such a mobility cost structure that makes $i$ prefer migration to commuting between locations $j$ and $k$. There are many factors that contribute to this cost structure such as family, housing, ties to source location, networks, access to public transport and access to private transport. Here, a sufficient statistic for these costs is the preference over mobility mode. The indicator $D_i$ captures this preference. Hence, controlling for any mobility costs or proxies of mobility costs is

[11] Residence location refers to year $t - 1$ location of residence.

| Dependent variable: | $\hat{u}_{ij}$ | | | |
|---|---|---|---|---|
| Migrant (ref: Commuter) | −0.0239 (0.0055)*** | −0.0214 (0.0091)* | −0.0228 (0.0092)* | −0.0203 (0.0092)* |
| **Dependent variable:** | $\hat{u}_{ik}$ | | | |
| Migrant (ref: Commuter) | 0.0316 (0.0051)*** | 0.0249 (0.0082)** | 0.0239 (0.0083)** | 0.0244 (0.0083)** |
| **Dependent variable:** | $\hat{u}_{ik} - \hat{u}_{ij}$ | | | |
| Migrant (ref: Commuter) | 0.0554 (0.0030)*** | 0.0462 (0.0090)*** | 0.0468 (0.0091)*** | 0.0447 (0.0088)*** |
| Cost controls | | | | |
| Distance to new job | No | No | Yes | Yes |
| Commuting distance | No | No | Yes | Yes |
| $j-k-l$ triad FE | No | Yes | Yes | Yes |
| $\hat{E}_{ik}[w]$, $\hat{E}_{ij}[w]$ | No | No | No | Yes |
| Year FE | No | Yes | Yes | Yes |
| Constant term | Yes | No | No | No |
| Observations | 43,068 | 43,068 | 43,068 | 43,068 |
| Migrants | 12,172 | 12,172 | 12,172 | 12,172 |
| Commuters | 30,896 | 30,896 | 30,896 | 30,896 |
| $R^2$, $\hat{u}_{ij}$ | 0.0005 | 0.0010 | 0.0013 | 0.0215 |
| $R^2$, $\hat{u}_{ik}$ | 0.0008 | 0.0027 | 0.0029 | 0.0267 |
| $R^2$, $\hat{u}_{ik} - \hat{u}_{ij}$ | 0.0024 | 0.0020 | 0.0022 | 0.0695 |

Table 2: **Selection on residuals.** *Notes: Data pooled from years 2011-2014. See Section 5.2 for the definitions of commuters, contracted migrants and stayers. See Section 5.1 for sample construction. For the computation of residuals, see Section 5.3. White heteroskedasticity robust standard errors clustered at residence-source-destination municipality level in parenthesis.*

would amount to controlling for the causal pathways of the effect of interest. On the other hand, if there is unobservable variation in costs, controlling for many variables that can be seen as proxies for costs, such as family and housing, commuting and migration experience may make the assumption of balanced unobservables across the two groups of mobile more plausible.

Table 2 presents the results of least squares regression estimation of (20) for pre- and post-mobility residuals. The contracted migrants are negatively selected relative to the commuters on pre-mobility residuals and positively selected relative to the commuters on post-mobility residuals. Thus, the pattern of residuals in Panel C is observed conditionally as well.

## 6.2 Selection on Job Match Quality

Selection on residuals may be due to either selection job match quality or selection unobservable skills. I now aim to distinguish these two types of selection. I first show how balance on unobservable skills among the observed commuters and migrants identify the effect of job match quality and discuss its plausibility. I then discuss the conditions under which difference-in-difference estimation identifies the effect of job match quality and study the differences in residual changes. Lastly, I pit selection on unobservables and job match quality against each other by deriving corresponding coefficient sign restrictions from the theory and show that a selection on job match quality is more consistent with the data than seleciton on unobservable skills.

### 6.2.1 Balanced Unobservable Skills

Identification of the effect of job match quality on pre- and post-mobility residuals requires that we compare commuters and migrants who are equal in their mean unobservable skills. To see this, consider the conditional difference in the residuals of migrants and commuters:

$$\begin{aligned} \hat{u}_{ih} &= controls + \tau D_i + \varepsilon_i \\ &= controls + \left[E[q_{ih}|D=1] - E[q_{ih}|D=0]\right] D_i + \epsilon_i \end{aligned} \tag{21}$$

with $\epsilon_i = \rho_h \left[E[\nu_i|D=1] - E[\nu_i|D=0]\right] D_i + \varepsilon_i$. If $E[D_i\varepsilon_i] = 0$, that is if *controls* contain all factors with a simultaneous effect on $\hat{u}_{ih}$ and $D_i$ excluding unobservable skills, unbiased estimation of $E[q_{ih}|D=1] - E[q_{ih}|D=0]$ requires $E[D_i\epsilon_i] = 0$ which requires

$$\rho_h \left(E[\nu_i|D=1] - E[\nu_i|D=0]\right) = 0 \iff E[\nu_i|D=1] = E[\nu_i|D=0]. \tag{22}$$

There is, of course, no direct way of knowing whether (22) is anywhere close to being satisfied in the estimation of (20). The theory suggests that everything else hold constant, migrants are selected more strongly on unobservables if they incur larger relocation costs than commuters speaking against (22). However, if the selection on unobservables plays a relatively minor role in selection, (22) might be plausible owing to the sample choice ensuring high homogeneity with respect to labor market outcomes and the similar hurdles of search and job finding that the mobile pass.

### 6.2.2 Difference-in-difference

It is clear that if the pre-mobility job match quality of migrants is lower and post-mobility job match quality is higher than that of the commuters, then the change in job match quality among the migrants should be greater than among the commuters as well. Modeling the change in job match quality is, however, interesting in its own right since if the unobservables balance in the two mobility groups or if the compensations for unobservables in the source and destination region equal, the individual unobservable effect has a common trend among the commuters and migrants and, hence, within source-destination pair, a difference-in-difference estimator identifies the change in job match quality.

To see this, consider model (20) but let now the residual change $u_{ik} - u_{ij}$ be the dependent variable. The coefficient $\tau$ then identifies the difference in the residual differences

$$\begin{aligned} &E[u_{ik} - u_{ij}|D = 1] - E[u_{ik} - u_{ij}|D = 0] \\ =&(\rho_k - \rho_j)\left(E[\nu_i|D = 1] - E[\nu_i|D = 0]\right) + E[q_{ik} - q_{ij}|D = 1] - E[q_{ik} - q_{ij}|D = 0], \quad (23) \end{aligned}$$

where the first term is the difference in unobservable trends. Note that even if, as visible in Figure 4, the residuals of the commuters and migrants evolve very similarly prior to mobility, these common trends cannot be extrapolated to the year of mobility as in that year the compensation for the unobservables changes so that the group specific unobservable effects do not cancel out in the before-after comparison.[12] This means that if the commuters and migrants are on average different in their unobservables, then the compensations for their unobservables evolve differently in the year of mobility violating common trend. As is clear from (23) common trend is satisfied when $\rho_k = \rho_j$.[13]

It makes thus an interesting exercise to study sample restricted by the values the difference $\rho_k - \rho_j$ can take. There is no direct measure of the location's compensation for unobservable skills. Previous research has proxied $\rho_h$ with varius inequality measures of location $h$ wages such the 90-20 income share ratio (Borjas, 1987), ratio of 75th to 25 the percentile of the earnings distribution (Parey et al., 2017), Gini coefficient (Liebig & Sousa-Poza, 2004) standard deviation of log wage (Borjas et al., 1992) and standard deviation of residuals from wage regressions (Borjas et al., 1992; Gould & Moav, 2016).[14] I proxy $\rho_h$ with three measures if wage variation: the standard deviation of wages, standard deviation of predicted wages and standard deviation of residuals among location $h$ workers. All these measures are computed in the analysis sample to capture wage variation in the labor markets that the workers in the sample face. As there is substantial year-to-year variation, for each municipality an average over the sample years is computed.

Figure 5 presents estimated coefficients in these samples. The upper panel uses as a measure of $\rho_h$ the standard deviation of residuals of workers in location $h$ whereas the lower panel uses the standard deviation of wages of location $h$ workers. The most interesting sample is the one where $\rho_k = \rho_j$ is restricted to be close to zero. In this sample, the evolutions of the value of unobservables among the contracted migrants and commuters have close to parallel trend. It seems that the migrants have larger changes

[12]Note that since the group unobservable effect $E[\nu_i|D = d]$ and the effect of changing location $\rho_k - \rho_j$ are multiplicative, common trend is violated for both levels and logarithmic transformation of wages. See e.g. Lechner (2011) page 186.

[13]Several authors have studied wage growth of migrants in order to cancel out the individual (or group) fixed effect (Bartel, 1979; Yankow, 2003; Ham et al., 2011). However, as is clear above, the individual fixed effect in a wage regression should be interpreted as a the price paid for the individual fixed effect. If this price changes as might be if the location of labor supply changes, then the individual fixed effect is not time-invariant and does not cancel out in a before-after comparison.

[14]Also various measures of returns to education have been used (Gould & Moav, 2016; Moraga, 2013).

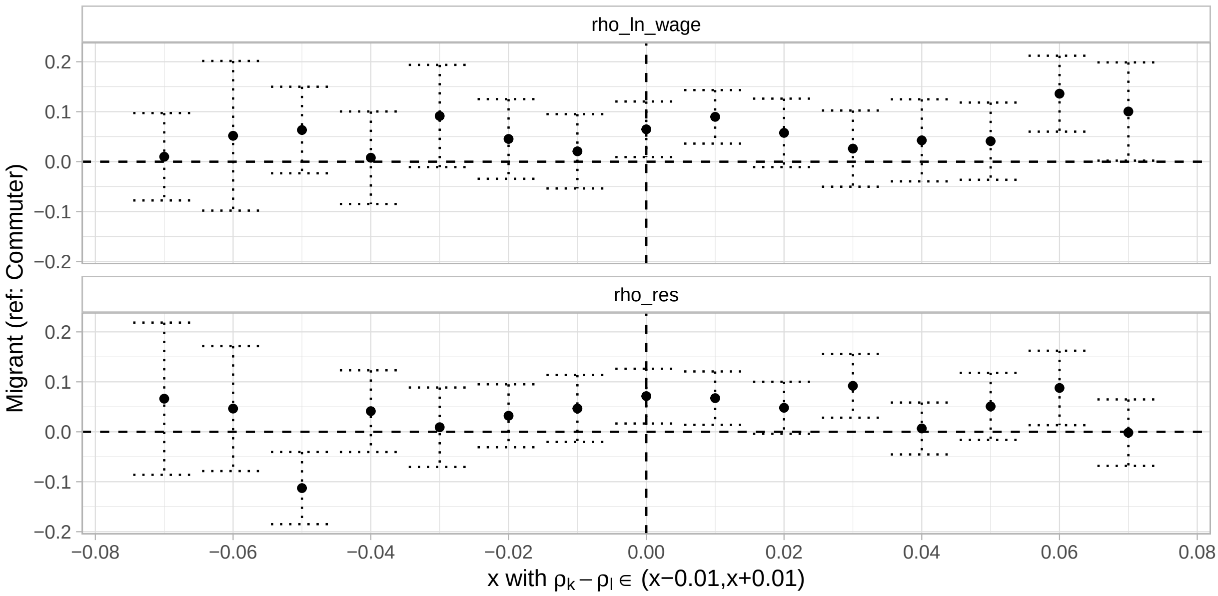


Figure 5: **Difference-in-difference with sample splits** *Notes: For the data, see Section 5.1. See definitions in Section 5.2. Mobility potentially occurs between time points $t-1$ and $t$. Models of column 4 in Table 2*

in their residuals than commuters in the sample where the common trend assumption is likely to be satisfied to a reasonable degree.

### 6.2.3 Model Restrictions

Selection on job match quality and selection on unobservable skills make different predictions for the sign patterns of the differences in residuals and their changes of contracted migrants and commuters. We can, thus, pit the models against each other and see whether selection on job match quality or selection on unobservable skills is more consistent with the data.

Given a source-destination pair, if the unobservables are not balanced accross the mobile, either the contracted migrants or the commuters have higher mean unobservable skills. Importantly, if unobservable skills are location-time-invariant, then whichever mobility group has more valuable unobservable skills pre-mobility should have more valuable unobservable skils post-mobility as well. First, both (6) and (7) are increasing in $E[\nu_i]$, and, thus, if the contracted migrants and commuters differ in their unobservable skills, both the pre-and post-mobility unobservable skills should be higher for the group with higher unobservable skills.

Consider the pre- and post-mobility residuals of the contracted migrants and commuters. For simplicity, suppose $\mu_\nu^m = \mu_\nu^c$. This restriction makes either of the models consistent with the data. Removing within skill wage dispersion from the model,

$\sigma_j^2 = \sigma_k^2 = 0$, so that there is no selection on job match quality yields, for $\rho_k - \rho_j > (<)\, 0$

$$E[u_{ij}|D=1] - E[u_{ij}|D=0] = \frac{1}{\sigma_\Delta}\sigma_\varepsilon^2(\rho_k - \rho_j)\rho_j[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] > (<)\, 0$$
$$E[u_{ik}|D=1] - E[u_{ik}|D=0] = \frac{1}{\sigma_\Delta}\sigma_\varepsilon^2(\rho_k - \rho_j)\rho_k[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] > (<)\, 0$$

Restricting $\sigma_\varepsilon^2 = 0$ so that there is no heterogeneity in unobservable skills and thus no selection on unobservable skills yields

$$E[u_{ij}|D=1] - E[u_{ij}|D=0] = -\frac{1}{\sigma_\Delta}\sigma_j^2[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] < 0$$
$$E[u_{ik}|D=1] - E[u_{ik}|D=0] = \frac{1}{\sigma_\Delta}\sigma_k^2[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] > 0$$

In Table 3, we observe $\hat{E}[u_{ij}|D=1] - \hat{E}[u_{ij}|D=0] < 0$ and $\hat{E}[u_{ik}|D=1] - \hat{E}[u_{ik}|D=0] > 0$ for both $\rho_k - \rho_j > 0$ and $\rho_k - \rho_j < 0$. For pre-mobility residuals the estimates are imprecise, but the coefficients correspond largely to the coefficients estimated in the whole sample. Thus, selection on job match quality explains the pattern in observed residuals whereas selection on unobservable skills does not.

Consider the change in residuals of the contracted migrants and commuters. For simplicity, suppose $\mu_\nu^m = \mu_\nu^c$. This restriction makes either of the models consistent with the data. Removing within skill wage dispersion from the model, $\sigma_j^2 = \sigma_k^2 = 0$, so that there is no selection on job match quality yields

$$E[u_{ik} - u_{ij}|D=1] - E[u_{ik} - u_{ij}|D=0] = \sigma_\varepsilon(\rho_k - \rho_j)[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] > 0$$

if $\rho_k - \rho_j > 0$ and

$$E[u_{ik} - u_{ij}|D=1] - E[u_{ik} - u_{ij}|D=0] = \sigma_\varepsilon(\rho_k - \rho_j)[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] < 0$$

if $\rho_k - \rho_j < 0$. Restricting $\sigma_\varepsilon^2 = 0$ so that there is no heterogeneity in unobservable skills and thus no selection on unobservable skills yields

$$E[u_{ik} - u_{ij}|D=1] - E[u_{ik} - u_{ij}|D=0] = \left(\sigma_k^2 + \sigma_j^2\right)^{\frac{1}{2}}[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] > 0$$
$$E[u_{ik} - u_{ij}|D=1] - E[u_{ik} - u_{ij}|D=0] = \left(\sigma_k^2 + \sigma_j^2\right)^{\frac{1}{2}}[\lambda(z_{ijk}^m) - \lambda(z_{ijk}^c)] > 0$$

In Table 3, we observe $\hat{E}[u_{ik} - u_{ij}|D=1] - \hat{E}[u_{ik} - u_{ij}|D=0] > 0$ regardless of the sign of $\rho_k - \rho_l$. Hence, again, the overall pattern of signs of the residual differences between the contracted migrants and commuters are better explained in terms of selection on job match quality than in terms of selection on unobservable skills.

| $\rho$ measure: | $sd(\hat{u}_{ih})$ | $sd(w_{ih})$ | $sd(\hat{u}_{ih})$ | $sd(w_{ih})$ |
|---|---|---|---|---|
| Sample: | $\rho_k - \rho_j > 0$ | | $\rho_k - \rho_j < 0$ | |
| Dependent variable: | | $\hat{u}_{ij}$ | | |
| Migrant | −0.0201 | −0.0196 | −0.0108 | −0.0166 |
| (ref: Commuter) | (0.0107) | (0.0110) | (0.0129) | (0.0124) |
| Dependent variable: | | $\hat{u}_{ik}$ | | |
| Migrant | 0.0329 | 0.0445 | 0.0385 | 0.0185 |
| (ref: Commuter) | (0.0095)*** | (0.0105)*** | (0.0123)** | (0.0122) |
| Dependent variable: | | $\hat{u}_{ik} - \hat{u}_{ij}$ | | |
| Migrant | 0.0531 | 0.0641 | 0.0494 | 0.0351 |
| (ref: Commuter) | (0.0109)*** | (0.0110)*** | (0.0143)*** | (0.0130)** |
| Cost controls | | | | |
| Distance to new job | Yes | Yes | Yes | Yes |
| Commuting distance | Yes | Yes | Yes | Yes |
| $j-k-l$ triad FE | Yes | Yes | Yes | Yes |
| $\hat{E}_{ik}[w]$, $\hat{E}_{ij}[w]$ | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 22,797 | 23,069 | 21,455 | 19,999 |
| Migrants | 6,345 | 6,314 | 5,827 | 5,858 |
| Commuters | 16,452 | 16,755 | 14,444 | 14,141 |
| $R^2$, $\hat{u}_{ij}$ | 0.0261 | 0.0267 | 0.0236 | 0.0222 |
| $R^2$, $\hat{u}_{ik}$ | 0.0263 | 0.0226 | 0.0248 | 0.0250 |
| $R^2$, $\hat{u}_{ik} - \hat{u}_{ij}$ | 0.0771 | 0.0689 | 0.0660 | 0.0652 |

Table 3: **Residual differences and difference-in-differences by sign of $\rho_k - \rho_l$.** *Notes: *p<0.05; **p<0.01; ***p<0.001. Data pooled from years 2011-2014. See Section 5.2 for the definitions of mobile, contracted migrants and commuters. For the computation of residuals, see Section 5.3. Personal controls contain housing type, family type, whether spouse working, previous commuting and migration experience, commuting distance in year $t-1$, control or ownership of a car, distance and distance squared to new job. White heteroskedasticity robust standard errors clustered at residence-source-destination municipality level in parenthesis.*

## 6.3 Mobility Costs

The differences between the residuals of contracted migrants and commuters have been interpreted as reflecting differences in their mobility costs. To support this reasoning, I now see if these differences are mediated by a more direct measures of mobility costs.

If different mobility costs drive the results, then when migration costs are smaller relative to commuting costs, the job match quality differences of contracted migrants and commuters should be smaller. Hence, I study the interaction effects of mobility mode and different mobility cost proxies. Table 4 presents the results. In the first column, the mobility cost proxy is whether the worker lives alone prior to mobility. Living alone reduces the relative migration costs. Thus, the difference in changes in residuals across contracted migrants and commuters should be smaller. This seems to be the case. Working spouse, on the other hand, increases migration cost relative to commuting cost. Hence, we observe

| Dependent variable: | $\hat{u}_{ik} - \hat{u}_{ij}$ | | | | |
|---|---|---|---|---|---|
| Cost proxy, $t-1$: | Lives alone | Spouse working | Lives rental | Owns a car | Distance to new job |
| Migrant | | | | | |
| * Cost proxy | $-0.0482$ | 0.0470 | $-0.0148$ | 0.0302 | $-0.0009$ |
| (ref: Commuter * Cost proxy) | (0.0209)* | (0.0195)* | (0.0224) | (0.0203) | (0.0077) |
| Cost controls | | | | | |
| Distance to new job | Yes | Yes | Yes | Yes | Yes |
| Commuting distance | Yes | Yes | Yes | Yes | Yes |
| $j-k-l$ triad FE | Yes | Yes | Yes | Yes | Yes |
| $\hat{E}_{ik}[w]$, $\hat{E}_{il}[w]$ | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 43,068 | 43,068 | 43,068 | 43,068 | 43,068 |
| Migrants | 12,172 | 12,172 | 12,172 | 12,172 | 12,172 |
| Commuters | 30,896 | 30,896 | 30,896 | 30,896 | 30,896 |
| $R^2$ | 0.0680 | 0.0682 | 0.0669 | 0.0674 | 0.0668 |

Table 4: **Residual difference-in-differences interacted with cost proxies.** *Notes:* $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$. *Data pooled from years 2011-2014. See Section 5.2 for the definitions of mobile, contracted migrants and commuters. For the computation of residuals, see Section 5.3. White heteroskedasticity robust standard errors clustered at residence-source-destination municipality level in parenthesis.*

a positive coefficient. Living rental reduces the relative cost of migration and as expected, we observe a negative effect, albeit not significantly different from zero. Owning a car reduces the cost of commuting and as expected, we observe a positive interaction effect, albeit not significantly different from zero. Longer distance to new job increases the cost of commuting relative to migration. As would be predicted, we observe a negative effect, albeit not significantly different from zero.

## 6.4 Robustness

### 6.4.1 The Threshold Distance of Mobility

The threshold distance in defining mobility of 50km is somewhat arbitrary, motivated to ensure that the choice between commuting and migration is not trivial and that there are both contracted migrants and commuters in the sample. Figure 6 shows that the estimated difference in job match qualities among the contracted migrants and commuters is robust to this threshold.

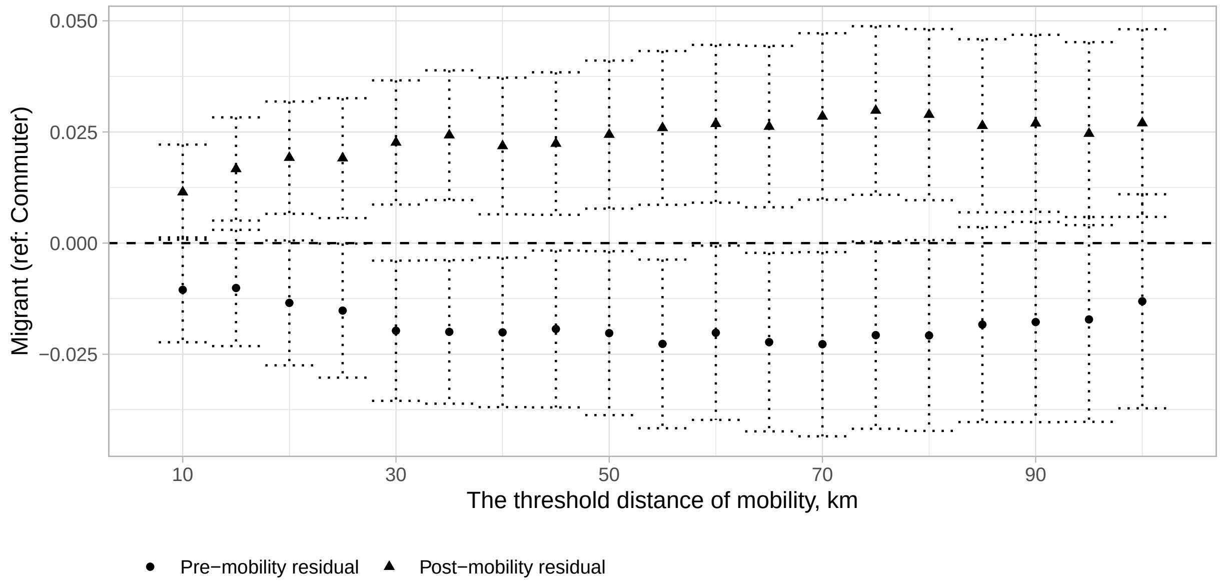


Figure 6: **Robustness with respect to the threshold distance of mobility.** *Notes: For the data, see Section 5.1. See definitions in Section 5.2. For the computation of residuals, see Section 5.3. Models of column 4 in Table 2. 95-percent confidence intervals are based on white heteroskedasticity robust standard errors clustered at residence-source-destination municipality level.*

# 7 Conclusions

Ignoring the possibility of contracted migration and interpreting all migration as speculative may lead biases in the interpretation of migrant selection results. First, no self-selection is required for contracted migrants to be positively selected when there is competition for jobs. Contracted migrants are selected by employers and profit-maximizing employers for each wage select the most productive applicants. Hence, even if a random sample searches for jobs interregionally, those who receive a job offer are not randomly selected. Second, as contracted migrants base their choice of migration on a realized wage offer, their observed wage distribution in the destination is not their expected wage distribution in the destination but their expected wage distribution truncated below at some threshold wage. The contracted immigrants may thus be drawing wage offers from the very same wage distribution as the natives, and so be equally able, but as the contracted migrants only accept wages that compensate for their migration costs, they realize higher wages than the natives. There is positive selection on destination job match quality. Third, as at the point of migration choice the destination wage of the contracted migrant is fixed, it is independent of the current wage. Hence, the observed wage distribution of contracted migrants in the source is not their expected wage distribution but their expected wage distribution truncated above. There is negative selection on source job match quality.

This paper takes a step toward including the possibility of contracted migration to analyses of migrant selection. The Roy-Borjas model extended with observability of source and destination location wages derived from a job search model that naturally highlights the wage dispersions in the source and destination location provides a framework to study contracted migration with many results and methods already familiar from the literature. While there is selection on unobservable skill depending on source and destination location difference in returns to skills, this selection is confounded by negative selection on job match quality in the source and positive selection on job match quality in the destination. Adding job match quality with zero autocorrelation in to the Roy-Borjas framework also provides a relevant interpretation for Borjas' (1987) "refugee sorting" selection pattern that combines negative selection in the source and positive selection in the destination. Rather than the narrow interpretation of migration from communist to non-communist regimes, a more useful interpretation arises when we view workers fleeing their bad job matches for better job matches.

I find evidence for the role job match quality plays by comparing two mobile groups with different mobility costs. While the classification of migrants into speculative and contracted migrants is limited by data, the approach taken here identifies a group of migrant that highly likely are contracted. This group of contracted migrants is the compared to the commuters who incur different mobility costs in order to control for the self-selection to interregional search and employer selection. Mobility costs create a scale effect on migrant selection as higher costs are compensated by stronger selection. Comparing two groups with different costs reveals what type of selection, selection on job match quality or selection on unobservable skills is magnified. Migrants have around 2-3 percent lower pre-mobility residuals and 2-3 percent higher post-mobility residuals than commuters. Selection on job match quality better explains this pattern than selection on unobservable skills.

# A Appendix

## A.1 Proofs of Propositions and Lemmas

**Lemma 1** (No speculative migration)**.** *Let a speculative migrant from $j$ to $k$ search for a job in $k$ in continuous time with unemployment income $b$ and accept the first offered job arriving at rate $\varphi$. Let she discount future with discount rate $r$. Then worker $i$ does not migrate speculatively iff* $\frac{\mu_k+\nu_{ik}-w_{ij}}{r} < \frac{w_{ij}-b+\pi_{ijk}}{\varphi}$.

*Proof of Lemma 1.* Let $rV_k = b - \pi_{ik} + \varphi[\frac{\mu_k+\nu_{ik}}{r} - V_k]$ be the asset value of migrating to $k$ and $rV_j = w_{ij} = w_{ij} - \pi_{ij} + \varphi[\frac{w_{ij}}{r} - V_j]$ the asset value of staying the in the current job in $j$. The values can then be written as

$$(r+\varphi)V_k = b - \pi_{ik} + \frac{\varphi}{r}(\mu_k + \nu_{ik}) \tag{24}$$

$$(r+\varphi)V_j = w_{ij} - \pi_{ij} + \frac{\varphi}{r}w_{ij} \tag{25}$$

and thus $V_k < V_j \iff \frac{\mu_k+\nu_{ik}-w_{ij}}{r} < \frac{w_{ij}-b+\pi_{ijk}}{\varphi}$. □

**Lemma 2.** *Let $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$, $\tilde{\Delta}_i \sim \mathcal{N}(E[\tilde{\Delta}_i], \sigma_\Delta^2)$ and $\Delta_i = \tilde{\Delta}_i - E[\tilde{\Delta}_i]$. Then*

$$E[x|\frac{\Delta_i}{\sigma_\Delta} > z] = \mu_x + \frac{E[x\Delta_i]}{\sigma_\Delta}\lambda(z), \tag{26}$$

*where $\lambda(z) \coloneqq \phi(z)/(1-\Phi(z))$ is the inverse Mill's ratio, where $\phi$, and $\Phi$ denote the density and distribution functions of the standard normal, respectively.*

*Proof of Lemma 2.* First note that

$$E\left[\frac{x}{\sigma_x}\Big|\frac{\Delta_i}{\sigma_\Delta}\right] = E\left[\frac{x}{\sigma_x}\right] + \frac{Cov[\frac{x}{\sigma_x}, \frac{\Delta}{\sigma_\Delta}]}{Var[\frac{\Delta}{\sigma_\Delta}]}\frac{\Delta}{\sigma_\Delta} = \frac{1}{\sigma_x}\mu_x + \frac{E[x\Delta_i]}{\sigma_x\sigma_\Delta^2}\Delta_i \tag{27}$$

Thus,

$$E[x|\frac{\Delta_i}{\sigma_\Delta} > z] = \sigma_x E[\frac{x}{\sigma_x}|\frac{\Delta_i}{\sigma_\Delta} > z] = \mu_x + \frac{E[x\Delta_i]}{\sigma_\Delta}E[\frac{\Delta_i}{\sigma_\Delta}|\frac{\Delta_i}{\sigma_\Delta} > z] \tag{28}$$

$$= \mu_x + \frac{E[x\Delta_i]}{\sigma_\Delta}\frac{\phi(z)}{1-\Phi(z)}, \tag{29}$$

where the second equality follows from Lemma 3. Defining $\lambda(z) \coloneqq \phi(z)/(1-\Phi(z))$ gives the result. □

**Lemma 3.** *Let $r, a \in \mathbb{R}$ and $z \sim f_z$, $x \sim f_x$. Then $E[z|x] = c + rx \implies E[z|x > a] = c + rE[x|x > a]$.*

*Proof of Lemma 3.* Suppose $E[z|x] = c + rx$. Then

$$\begin{aligned}
E[z|x > a] &= \int\limits_z z f_z(z|x > a)dz = \int\limits_z z \int\limits_{x>a} f_z(z|X = x)f_x(x|x > a)dxdz \\
&= \int\limits_{x>a}\int\limits_z z f_z(z|X = x)dz f_x(x|x > a)dx = \int\limits_{x>a} E[z|x]f_x(x|x > a)dx \\
&= \int\limits_{x>a}(c + rx)f_x(x|x > a)dx = c + r\int\limits_{x>a} x f_x(x|x > a)dx = c + rE[x|x > a]. \quad \square
\end{aligned}$$

*Proof of Proposition 1.* Let $\tilde{\Delta}_i = (\rho_k - \rho_j)\nu_i + q_{ik} - q_{ij}$ be the right-hand side of (MC),

$\Delta_i = \tilde{\Delta}_i - (\rho_k - \rho_j)\mu_\nu$ and

$$z \coloneqq \frac{1}{\sigma_\Delta}\left(\mu_j - \mu_k + \pi - (\rho_k - \rho_j)\mu_\nu\right) \tag{30}$$

such that (MC) can be written as $\frac{\Delta_i}{\sigma_\Delta} > z$. Then by Lemma 2

$$E[\rho_h \nu_i|(\text{MC})] = E[\rho_h \nu_i|\frac{\Delta_i}{\sigma_\Delta} > z] = \rho_h \mu_\nu + \frac{1}{\sigma_\Delta}\sigma_\nu^2(\rho_k - \rho_j)\rho_h \lambda(z) \tag{31}$$

and

$$E[q_{ij}|(\text{MC})] = -\frac{1}{\sigma_\Delta}\sigma_j^2\lambda(z), \quad E[q_{ik}|(\text{MC})] = \frac{1}{\sigma_\Delta}\sigma_k^2\lambda(z) \tag{32}$$

The result follows from the decomposition $w_{ih} = \mu_h + \rho_h \nu_i + q_{ih}$. □

*Proof of Proposition 2.* As Proposition 1 but the condition for relocation is now

$$w_{ij} < w_{ik} - \pi \iff \mu_{ij} - \mu_{ik} + \pi_{jk} < (\rho_k - \rho_j)\nu_i + q_{ik} - q_{ij}. \tag{33}$$

and thus let

$$z_{ijk} \coloneqq \frac{1}{\sigma_\Delta}\left(\mu_{ij} - \mu_{ik} + \pi_{jk} - (\rho_k - \rho_j)\mu_\nu\right). \tag{34}$$

The result follows from the decomposition $u_{ih} = \rho_h(\nu_i - \mu_h^\nu) + q_{ih}$. □

# References


Abramitzky, R. (2009). The effect of redistribution on migration: Evidence from the Israeli kibbutz. *Journal of Public Economics*, *93*(3-4), 498–511.

Bartel, A. P. (1979). The migration decision: What role does job mobility play? *The American Economic Review*, *69*(5), 775–786.

Bartolucci, C., Villosio, C., & Wagner, M. (2018). Who migrates and why? evidence from italian administrative data. *Journal of Labor Economics*, *36*(2), 551–588.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, 531–553.

Borjas, G. J., Bronars, S. G., & Trejo, S. J. (1992). Self-selection and internal migration in the united states. *Journal of urban Economics*, *32*(2), 159–185.

Borjas, G. J., Kauppinen, I., & Poutvaara, P. (2018). Self-selection of emigrants: Theory and evidence on stochastic dominance in observable and unobservable characteristics. *The Economic Journal*, *129*(617), 143–171.

Detang-Dessendre, C., & Molho, I. (1999). Migration and changing employment status: a hazard function analysis. *Journal of Regional Science*, *39*(1), 103–123.

Dostie, B., & Léger, P. T. (2009). Self-selection in migration and returns to unobservables. *Journal of Population Economics*, *22*(4), 1005–1024.

Emmler, J., & Fitzenberger, B. (2020). The role of unemployment and job change when estimating the returns to migration.

Flinn, C. J. (1986). Wages and job mobility of young workers. *Journal of Political Economy*, *94*(3, Part 2), S88–S110.

Garen, J. E. (1989). Job-match quality as an error component and the wage-tenure profiler a comparison and test of alternative estimators. *Journal of Business & Economic Statistics*, *7*(2), 245–252.

Gould, E. D., & Moav, O. (2016). Does high inequality attract high skilled immigrants? *The Economic Journal*, *126*(593), 1055–1091.

Ham, J. C., Li, X., & Reagan, P. B. (2011). Matching and semi-parametric iv estimation, a distance-based measure of migration, and the wages of young men. *Journal of Econometrics*, *161*(2), 208–227.

Hicks, J. R. (1932). *The theory of wages.* New York: Macmillan.

Hunt, G. L., & Mueller, R. E. (2004). North american migration: returns to skill, border effects, and mobility costs. *Review of Economics and Statistics*, *86*(4), 988–1007.

Kaestner, R., & Malamud, O. (2014). Self-selection and international migration: New evidence from mexico. *Review of Economics and Statistics*, *96*(1), 78–91.

Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2017). High-skilled migration and agglomeration. *Annual Review of Economics*, *9*, 201–234.

Liebig, T., & Sousa-Poza, A. (2004). Migration, self-selection and income inequality: an international analysis. *Kyklos*, *57*(1), 125–146.

Moraga, J. F.-H. (2011). New evidence on emigrant selection. *The Review of Economics and Statistics*, *93*(1), 72–96.

Moraga, J. F.-H. (2013). Understanding different migrant selection patterns in rural and urban mexico. *Journal of Development Economics*, *103*, 182–201.

Mortensen. (2003). *Wage dispersion: why are similar workers paid differently?* MIT press.

Parey, M., Ruhose, J., Waldinger, F., & Netz, N. (2017). The selection of high-skilled emigrants. *Review of Economics and Statistics*, *99*(5), 776–792.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, *3*(2), 135–146.

Saben, S. (1964). Geographic mobility and employment status, march 1962-march 1963. *Monthly Lab. Rev.*, *87*, 873.

Schultz, T. W. (1961). Investment in human capital. *The American economic review*, *51*(1), 1–17.

Silvers, A. L. (1977). Probabilistic income-maximizing behavior in regional migration. *International Regional Science Review*, *2*(1), 29–40.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of political Economy*, *70*(5, Part 2), 80–93.

Yankow, J. J. (2003). Migration, job change, and wage growth: a new perspective on the pecuniary return to geographic mobility. *Journal of Regional Science*, *43*(3), 483–516.