MEMORANDUM 1/2024

The Finnish Economic Policy Council

THE FINNISH ECONOMIC POLICY COUNCIL

Geographical and Occupational Mismatch in Finland over the Recent Era of Economic Crises: Estimates Using an Interconnected-Markets Approach

Veikko Uusitalo
Merja Kauhanen
Annika Nivala
Tuomo Suhonen

MEMORANDUM 1/2024

Geographical and Occupational Mismatch in Finland over the Recent Era of Economic Crises: Estimates Using an Interconnected-Markets Approach

Veikko Uusitalo
Merja Kauhanen
Annika Nivala
Tuomo Suhonen

The Finnish Economic Policy Council
VATT Institute for Economic Research

Helsinki 2024

MEMORANDUM 1/2024

Veikko Uusitalo, University of Helsinki and The Labour Institute for Economic Research LABORE, veikko.uusitalo@helsinki.fi

Merja Kauhanen, The Labour Institute for Economic Research LABORE, merja.kauhanen@labore.fi

Annika Nivala, VATT Institute for Economic Research, annika.nivala@vatt.fi

Tuomo Suhonen, The Labour Institute for Economic Research LABORE, tuomo.suhonen@labore.fi

The Finnish Economic Policy Council
Address:
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, November 2024

Geographical and Occupational Mismatch in Finland over the Recent Era of Economic Crises: Estimates Using an Interconnected-Markets Approach*

Veikko Uusitalo[1], Merja Kauhanen[1], Annika Nivala[2], and Tuomo Suhonen[1]

[1] *Labour Institute for Economic Research LABORE*
[2] *VATT Institute for Economic Research*

October 10, 2024

Abstract

In recent years, imbalances between labor demand and supply have increased in Finland, while the economy has been hit by several global and local shocks. We study how much of these imbalances can be explained by limitations in job seekers' geographical and occupational mobility by applying an interconnected-markets mismatch index approach which relaxes strong assumptions regarding job search behavior imposed in previous longitudinal studies. Our results suggest that geographical and occupational mismatch have only accounted for a small percentage of the lost matches: their estimated total contribution was, on average, around 12 percent until 2014, after which we observe a decline towards 5 percent by 2021. The results indicate that this development is mainly explained by the increased concentration of job seekers and open vacancies in the same regions and occupational groups rather than by job seekers becoming more likely to accept distant jobs.

Keywords: mismatch, labor market, job search, distance
JEL Codes: E24, J22, J23, J24, J61, J62, J63, R23

*This study was conducted for the Finnish Economic Policy Council. We are grateful for the financial support and feedback provided by the council. A special thanks to Niku Määttänen for his valuable advice. We also gratefully acknowledge the financial support from the OP Research Foundation for data expenses (project number 20230109) and the funding from the Palkansaajasäätiö Foundation for the early stage of the research.

1 Introduction

Labor market mismatch refers to a situation where open vacancies are filled slowly, even though there are many unemployed job seekers. This implies that employment is lower and unemployment higher than would be possible, and that some of the goods and services might be left unproduced due to firms' difficulties of finding suitable workers.

Labor market mismatch is commonly divided into geographical and occupational mismatch. The first type of matching problem arises because open vacancies are located too far from job seekers' homes and the latter one because open vacancies are for occupations that do not match with those of the job seekers. Recently, several studies from high-income countries have attempted to measure the extent of both types of matching problems (e.g. Şahin et al. 2014, Patterson et al. 2016, Marinescu and Rathelot 2018, Turrell et al. 2021, Alasalmi 2022, Pizzinelli and Shibata 2023). Overall, the findings suggest that geographical mismatch only plays a minor role in total unemployment, while occupational mismatch has been found to have significant economic consequences in some contexts (Şahin et al. 2014, Patterson et al. 2016).

In this paper, we examine the development of geographical and occupational mismatch in Finland between 2006 and 2021. During this period, Finland's economy was hit by the global financial crisis and the COVID-19 pandemic as well as idiosyncratic shocks related to the collapse of the Nokia mobile phone cluster and the declining paper industry. These events and other structural changes have potentially exacerbated mismatch between labor demand and supply.

As shown in previous studies (Pehkonen et al. 2018, Alasalmi 2022), there have been significant increases in the Finnish vacancy rate and unemployment rate between the early 2010s and the early 2020s, which have resulted in an outward shift of the Beveridge Curve. This development is visualized in Figure 1. The figure shows that unemployment increased significantly during and after the global financial crisis, remaining elevated until the mid-2010s. Subsequently, there was a a positive trend in the vacancy rate, and for a given level of unemployment, the vacancy rate has been higher than before. From a policy perspective, it is important to understand to what extent the implied decrease in overall matching efficiency has resulted from limitations in the geographical or occupational mobility of the Finnish workforce.

Figure 1: Beveridge Curve in Finland, 2006-2023, end of the year

Data source: Statistics Finland StatFin database. The number of open vacancies and the unemployment rate are from the Employment Service Statistics, and the number of employed workers is from the Labour force survey. The vacancy rate is defined as open vacancies / (open vacancies + employed workers).

To determine how many matches are lost due to the geographical or occupational misallocation of job seekers, we use the interconnected-markets mismatch index proposed by Marinescu and Rathelot (2018) as our primary approach. This approach relaxes a simplifying assumption made in recent studies, which have mainly relied on the mismatch index of Şahin et al. (2014), that labor markets can be divided into distinct geographical or occupational segments and that job seekers only apply for work within their own segment. This assumption often leads to results that are difficult to interpret due to their built-in sensitivity to the assumed scope of the labor market. By basing their analysis on job seekers' actual scope of job search estimated from online job board data, Marinescu and Rathelot (2018) have shown that, in many cases, their approach leads to markedly lower mismatch estimates in comparison to approaches assuming distinct labor markets.

While Marinescu and Rathelot (2018) only applied alternative mismatch indices for examining mismatch at a single point in time (April-June 2012), we are not aware of previous studies attempting to extend their analysis to a longitudinal context. Thus, our longitudinal analysis using the alternative mismatch indices of Marinescu and Rathelot (2018) and Şahin et al. (2014) sheds new light on how different assumptions of job seekers' behavior affect our view of the long-term development of geographical and occupational mismatch.

We also deviate from the original study of Marinescu and Rathelot (2018) by estimating job seekers' distaste for geographical and occupational distances using realized employment contracts available in our country-wide matched employer-employee data. It is worth noting that both types of data, the job board data used by Marinescu and Rathelot (2018) and our employment contract data, have their pros and cons regarding the estimation of job seekers' true distaste for distance. While the first data describe job seekers' *stated preferences* regarding acceptable jobs among those who used the job board in question, the latter data provide nationally representative information on the *revealed preferences* of the job seekers who accepted a job offer. Depending on the differences between job seekers' stated and revealed preferences as well as the preferences of the job seekers who are excluded from these datasets, the estimates using both types of data could be biased either upwards or downwards. In the case of our employment contract data, one type of bias results from the preferences of employers and their possible distaste for geographically or occupationally distant employees while making job offers. This may explain why our distaste estimates are higher than those obtained by Marinescu and Rathelot (2018). Another possible explanation is that job seekers are generally more mobile in the US than in Finland.

Our Poisson regression results suggest that geographical and occupational distances represent significant barriers to a job seeker's probability of being matched to a particular job. However, our mismatch index results indicate that these barriers matter surprisingly little in terms of aggregate unemployment, as geographical and occupational mismatch can only account for a small percentage of the lost matches: their estimated total contribution was, on average, around 12 percent until 2014, after which we observe a decline towards 5 percent by 2021. When examined separately, both geographical and occupational mismatch show a decreasing trend after the early 2010s. This development is likely driven by reduced distances between job seekers and open vacancies due to their increased concentration in specific regions and occupations. Then again, changes in job seekers' distaste for distance are found to be modest, and therefore, they unlikely play a major role in the development of mismatch.

Our results complement previous evidence on the role of mismatch in the employment dynamics during the recent global crises (Şahin et al. 2014, Patterson et al. 2016, Turrell et al. 2021, Pizzinelli and Shibata 2023). We find that, in Finland, occupational mismatch increased significantly and persistently during the global financial crisis, whereas geographical mismatch even slightly decreased. According to our findings, the COVID-19 pandemic again increased both geographical and occupational mismatch, but these effects appear to have been short-lived, as the levels of mismatch are observed to return to the pre-pandemic levels by the end of 2021. These findings are in line with those of Pizzinelli and Shibata (2023) regarding sectoral misallocation between job seekers and vacancies in the US and the UK.

Comparing results obtained by alternative mismatch indices, we find that the different measures of geographical mismatch converge close to one another at the end of the observation period, all of the measures pointing towards a very low level of geographical mismatch. However, there are larger and more systematic differences across the measures of occupational mismatch, and the interconnected-markets index of Marinescu and Rathelot (2018) somewhat consistently indicates a lower level of occupational mismatch than the alternative indices that assume separate occupation-specific labor markets. The indicated trends and volatility in mismatch also vary, to some extent, across different mismatch indices. Interestingly, we find that the interconnected-markets geographical mismatch index is, overall, less volatile than the comparison indices but responds more sensitively to the temporary COVID-19 shock.

By making a simplifying assumption that job seekers are only matched with jobs within their broad 1-digit occupational group, we further shed light on the evolution of occupational heterogeneity in geographical mismatch in Finland. The estimated occupation-specific trends in mismatch are mainly in line with the general development. However, in contrast with the general trend between 2006 and 2012, i.e. around the years of the global financial crisis and the European debt crisis, mismatch among professionals shows a strong increasing trend during this period. The second group which stands out is craft and related trades workers, for which we find an almost opposite pattern, with a sharp decreasing trend in mismatch between 2006 and 2012 followed by an increasing trend between 2012 and 2016.

Apart from applying different mismatch indices, our study contributes to the prior literature by providing new type of suggestive evidence of the validity of these indices. We conduct this evaluation by linking different measures of labor market tightness, which the indices are based on, to individual-level data on the duration of unemployment spells. We show that, after controlling for the heterogeneity of job seekers across municipalities, the generalized labor market tightness measure of Marinescu and Rathelot (2018) performs better at predicting job seekers' probability of moving out of unemployment than the conventional measures that do not account for mobility across regional units. These findings provide further empirical support for our approach.

The remainder of the paper is structured as follows. Section 2 introduces the data and the methods used for analyzing labor market mismatch. In Section 3, we present our main results regarding job seekers' distaste for distance and the development of geographical and occupation mismatch in Finland. Section 4 includes additional results regarding the sensitivity of the results and the explanatory power of different labor market tightness measures. Section 5 concludes.

2 Methodology

2.1 Data

Our analysis utilizes nationally representative longitudinal register data on job seekers and vacancies. Our primary data for analyzing labor market mismatch, the employment service data from the Ministry of Economic Affairs and Employment, allow for identifying all job seekers and vacancies registered at the public employment services at a daily frequency between 2006 and 2021. The data include information on job seekers' municipality of residence, occupation (the 5-digit ISCO code) and educational degrees as well as the characteristics of job vacancies, including their municipality-level location and occupation. For the analysis, we aggregate the daily observations for each municipality and 2-digit occupation at the month level. Our analyses also utilize Statistics Finland's data on the length of job seekers' unemployment spells which originate from the employment service database.

While the employment service data are relatively comprehensive and rich, they have certain caveats, particularly concerning the number of job vacancies. According to the annual survey of the Ministry of Economic Affairs and Employment (Räisänen 2023), in 2022, employers only used public employment services in 45 percent of recruitments to new positions and in 55 percent of other recruitments. Therefore, data from the public employment offices significantly underestimate the overall demand for new labor. According to Larja and Peltonen (2023), the highest undercoverage concerns specialist occupations. For instance, the number of ICT specialist recruitments has exceeded the number of corresponding open vacancies in employment offices by seven-fold. However, for certain occupational groups, such as sales and construction workers, the employment service data appear to overestimate the true demand, for instance, because the 'on-demand' worker positions posted by staffing companies are counted as open vacancies.

To analyze job seekers' distaste for geographical and occupational distances, we merge the individual-level employment service data on job seekers with matched employer-employee data from years 2004-2020, which allows for comparing job seekers' locations and occupations to those of their new employment contracts. For the Poisson regression estimation, introduced in sub-section 2.2, we calculate how many job seekers with municipality i and occupation o were hired for a job in municipality j and occupation m. Job seekers' municipality of residence is determined from the employment service data three months before the starting date of the employment contract or, if unavailable, from Statistics Finland's longitudinal full-population register data at the end of the previous year.

As we only have municipality-level location data, the distances between job seekers and jobs are defined via distances between municipalities' center points. When a job seeker is hired in her municipality of residence, the commuting distance is determined as half of the distance to the closest neighboring municipality. In 2006, there were a total of 431 municipalities in Finland. The number decreased to 309 in 2021 due to municipal mergers.[1] Currently, an average Finnish municipality has around 18 000 residents, with a median surface area of 760 km^2. Åland, the region located in the Finnish archipelago, is excluded from the analyses due to its distinct local labor market.

2.2 Distaste for geographical and occupational distances

Following Marinescu and Rathelot (2018), we first evaluate at which locations and occupations job seekers are likely to search for employment, taking into account their place of residence and occupational background. Similar to Marinescu and Rathelot (2018), we estimate job seekers' distaste for distance using a Poisson regression model which links the flows of job seekers between locations and occupations to the corresponding travel distances and occupational skill differences. However, due to using matched employer-employee data instead of job board data, we use the number

[1] Despite the municipal mergers, we conduct the analyses using the original municipalities instead of a harmonized classification to increase the accuracy of the distance measures. However, aggregating job seekers and vacancies at the level of the 2021 municipalities does not significantly alter the results.

of new employment contracts (i.e. realized matches) by location and/or occupation, instead of the number of job applications, as the outcome variable. As yearly data are used for the estimation of the Poisson regression models, time periods t refer to calendar years in this sub-section.

In the first step of our analysis, we estimate job seekers' distaste for geographical distance by using a Poisson regression model similar to that of Marinescu and Rathelot (2018), which assumes that the mean number of matches μ_{ijt} for job seekers in municipality i to jobs in municipality j in time period t is given by:

$$\mu_{ijt} = U_{it} V_{jt} \exp\left[\alpha_{it} + \lambda_{jt} + \omega_{ijt} + s_t(d_{ij})\right], \tag{1}$$

where U_{it} represents the total number of unemployed job seekers in municipality i in time period t and V_{jt} is the total number of open vacancies in municipality j in time period t. $U_{it}V_{jt}$ serves as an offset variable with a coefficient constrained to 1, which controls for changes in the size of i and j, i.e. the regional distribution of job seekers or vacancies. α_{it} and λ_{jt} denote the time-period-specific fixed effects of municipalities i and j. ω_{ijt} is an additional fixed effect that captures the difference in the urbanity of municipalities i and j, which is expected to impact the mobility between these municipalities apart from their distance. Municipalities are classified in the data as urban, semi-urban or rural. ω_{ijt} includes indicators for all six combinations of the municipality types.

$s_t(\cdot)$ is a spline function that captures the effect of distance on μ_{ijt}. We define 5 knots, at 20, 50, 100, 200 and 400 kilometers, between which the spline function is piecewise and linear. We use a spline function, as it allows, unlike other functional forms, the marginal decrease in the relative likelihood of employment to be significantly different for close distance intervals (e.g. 0–20 km) and far distance intervals (e.g. 200–400 km).

In the second step, we estimate job seekers' distaste for occupational distance by assuming the following functional form for the mean number of matches for job seekers in occupation o to jobs in occupation m, μ_{omt}:

$$\mu_{omt} = U_{ot} V_{mt} \exp\left[\alpha_{ot} + \lambda_{mt} + \alpha_{1t}\tau_{om} + \alpha_{2t} \cdot 100\sqrt{\sum_{h \in H} (\delta_{o_h} - \delta_{m_h})^2}\right]. \tag{2}$$

Equation (2) includes two measures for the distance between occupations o and m. τ_{om} is a categorical variable indicating how many differing digits there are between the ISCO codes of these occupations. The second measure corresponds to the occupational skill differences measures used by Marinescu and Rathelot (2018). However, whereas the original study used O*NET data for measuring skill differences, we construct our measure using skill information readily available in the Statistics Finland full-population register data: the educational degrees of workers in different occupations. To calculate this measure, we cross-tabulate Finland's residents' 2-digit occupations and educational degrees differentiated by the level and the field of education in each year between 2010 and 2021. The classification of educational degrees consists of around 400 level-field combinations.

Within each occupation m, we determine the percentages of workers with each combination of educational level and field δ_{o_h}, after which we calculate the differences in these percentages $(\delta_{o_h} - \delta_{m_h})$ for all occupation pairs (m, o) (for $m = o$, these differences are 0). Finally, we take the square roots of the summed squared differences to determine the overall skill differences between the occupations. The largest and smaller occupational skill differences estimated from the 2021 data are reported in the Appendix, Table A.2.

Finally, to jointly evaluate the distaste for geographical and occupational distance, we define the following equation, which includes factors from both equation (1) and equation (2), for μ_{ijomt}, i.e. the mean number of matches for job seekers in municipality i and occupation o to jobs in municipality j and occupation m:

$$\mu_{ijomt} = U_{iot} V_{jmt} \exp \left[\alpha_{iot} + s_t(d_{ij}) + \alpha_{1t}\tau_{om} + \alpha_{2t} \cdot 100 \sqrt{\sum_{h \in H} (\delta_{o_h} - \delta_{m_h})^2} \right], \quad (3)$$

where α_{iot} represents the combined fixed effect of the job seeker's home municipality and occupation. For computational feasibility, we do not include job-side fixed effects or differences in urbanity between municipalities i and j (ω_{ijt}) in equation (3).

To estimate equations (1), (2) and (3) via the Poisson pseudo-maximum likelihood procedure, we use the STATA package *ppmlhdfe* (Correia et al. 2020) and yearly data on job seekers and new employment contracts from 2004–2020. The estimation is conducted separately for each year t, and in each case, we use data from three consecutive years, t, $t-1$ and $t-2$, to ensure sufficient statistical power. As a downside, we need to assume that changes in mobility take place slowly. This might be a restrictive assumption, e.g., in the case of analyzing labor market dynamics during the COVID-19 pandemic, which might have caused a sudden shock to mobility.

The distaste for geographical distance is calculated yearly between 2006–2020. However, due to a change in Statistics Finland's occupational classification that took place in 2010, we cannot estimate the distaste for occupational distance for years preceding 2010 in a manner that would be comparable with the later estimates. Therefore, we use the distaste estimates from years 2010-2012 when calculating the occupational mismatch indices for years 2006–2010. The same restriction holds for the combined geographical and occupational mismatch indices as well. Since available data on employment contracts only extends to 2020, the distaste estimates of 2020 are used for 2021 as well.

2.3 Mismatch index

We use the interconnected-markets mismatch index developed by Marinescu and Rathelot (2018) to explain the extent to which limitations in geographical and/or occupational mobility contribute to the unemployment rate. For simplicity and computational efficiency, we use the simplified mismatch index proposed in the Online Appendix of Marinescu and Rathelot (2018). In the simplified index, job seekers do not take into the account the probability of getting a job offer when applying for a job. According to the authors' calculations, this simplified index gives similar but slightly larger estimates than their main model.

In the model, mismatch is defined as $I_{I,t} = 1 - M_t/M_t^*$, where M_t is the number of matches predicted by the model, and M_t^* is the maximum number of matches achievable in the absence of limitations in mobility. While, in this sub-section, we only present the index for geographical mismatch, Marinescu and Rathelot (2018) have illustrated that this index straightforwardly extends to measuring occupational mismatch as well. As we calculate the index using monthly employment service data, hereafter time periods t refer to months in this sub-section.

The model assumes that each employer has one open vacancy. Job seekers send $\bar{a}_t$ applications, of which q_t are are deemed valid (capable of generating an offer). Employers can make an offer to one person only, randomly selected if multiple valid applications are received for a vacancy. Job seekers, if receiving multiple offers, choose their most preferred one. If the offer is declined, the vacancy remains unfilled.

The job seekers utility from a job is $g_t(d_{i(u)j(v)})\varepsilon_{uvt}$, where g is a decreasing function capturing job seekers' preference for nearby vacancies. This function is built using the distaste-for-distance estimates explained in the previous subsection. ε_{uvt} includes all other factors influencing job seekers' utility.

Let p_{ijt} be the probability that a job seeker in municipality i applies for a vacancy in municipality j in time period t. This probability is equal to the amount of applications $\bar{a}_t$ times the job seekers' distaste for distance, divided by the sum of vacancies in all municipalities multiplied by their respective distastes for distance:

$$p_{ijt} = \bar{a}_t \frac{g_t(d_{ij})}{\sum_\ell g_t(d_{ij}) V_{\ell t}}. \tag{4}$$

The expected number of applications for a vacancy in j is $r_{jt} = \sum_k p_{kjt} U_{kt}$, where U_{kt} is the number of unemployed job seekers in k. The distribution for the amount of valid applications received by a vacancy in j follows a Poisson distribution with parameter qr_j.

The probability a job seeker receives a job offer with their application is q_t (valid application) times the probability that the employer chooses their application out of all valid applications. This occurs when there are more than 0 valid applications ($P(q_t r_{jt} > 0) = 1 - exp(-q_t r_{jt})$) and the particular application is chosen ($1/q_t r_{jt}$). Thus, denoting $\mathcal{R}(x) = (1 - exp(-x))/x$, the probability of a job offer is:

$$\pi_{jt} = q_t \mathcal{R}(q_t r_{jt}). \tag{5}$$

The number of job offers received by a job seeker in municipality i is distributed Poisson($\sum_k p_{klt} \pi_{\ell t} V_{\ell t}$). Therefore, the probability that the job seeker gets at least one offer is $1 - exp(-\sum_k p_{kl} \pi_\ell V_\ell)$. This is the expression for the job finding rate for a job seeker in municipality i in time period t. By substituting in the previously derived p_{klt} and $\pi_{\ell t}$, and multiplying with the number of job seekers in each municipality, we obtain M_t, the predicted number of job matches:

$$M_t = \sum_{k \in K} U_{kt} \left[1 - exp\left(-q_t \overline{a_t} \frac{\sum_\ell g_t(d_{k\ell}) V_{\ell t} \mathcal{R}(q_t \bar{a}_t \nu_{\ell t})}{\sum_\ell g_t(d_{k\ell}) V_{\ell t}} \right) \right], \tag{6}$$

where ν_{jt} represents the generalized inverse tightness, the ratio of unemployed job seekers to open vacancies in all municipalities multiplied by their respective distastes for distance from the perspective of municipality j:

$$\nu_{jt} = \sum_{k \in K} \frac{g_t(d_{kj}) U_{kt}}{\sum_\ell g_t(d_{k\ell}) V_{\ell t}}. \tag{7}$$

The maximum number of matches M_t^* that a social planner can achieve by relocating job seekers can be determined by assuming that job seekers have no distaste for distance, i.e., that $g_t(d_{k\ell}) = 1$ for all distances $d_{k\ell}$:

$$M_t^* = \bar{U}_t \left[1 - exp\left(-q_t \bar{a}_t \mathcal{R}\left(q_t \bar{a}_t \frac{\bar{U}_t}{\bar{V}_t} \right) \right) \right]. \tag{8}$$

The interconnected-markets mismatch index $I_{I,t}$ is then obtained by calculating $1 - M_t/M_t^*$:

$$I_{I,t} = 1 - \sum_{k \in K} \frac{U_{kt}}{M_t^*} \left[1 - exp(-q_t \overline{a_t} \frac{\sum_\ell g_t(d_{k\ell}) V_{\ell t} \mathcal{R}(q_t \bar{a}_t \nu_{\ell t})}{\sum_\ell g_t(d_{k\ell}) V_{\ell t})}) \right], \tag{9}$$

where $\mathcal{R}(x) = (1 - exp(-x))/x$.

The mismatch index is calculated monthly between 2006 and 2021 using the yearly estimates for the distaste for distance $g_t(\cdot)$. Because of data restrictions, the distaste estimates for 2020 are also used for the estimation of the 2021 monthly mismatch indices.

In the computation of the mismatch index, the term $q_t \bar{a}_t$ (probability of valid application multiplied by the number of applications) is calibrated monthly using data on unemployment spells. The actual job finding rate (percentage of unemployed workers becoming employed in a given month) is multiplied by the total number of job seekers to obtain a realistic number of matches. $q_t \bar{a}_t$ is determined by minimizing the squared difference of matches predicted by the model and the realistic number of matches. For some time periods, the minimization problem does not have a viable solution, in which case the latest available estimate is used.

When measuring occupational mismatch, the number of open vacancies and job seekers are biased since the data does not include occupation codes for all observations. Since labor market tightness is a key variable in the level of mismatch, this issue is addressed with a calibration procedure. In our indices including occupational analysis, the number of open vacancies is multiplied with a coefficient so that the labor market tightness used in the model matches the true tightness, keeping the regional distribution of vacancies and job seekers fixed.

As a point of comparison for the interconnected-markets index of Marinescu and Rathelot (2018), we use the mismatch index of Şahin et al. (2014), which is based on the Cobb-Douglas matching function. It relies on the assumption that regions and occupations form distinct labor markets and that job seekers only apply for jobs within their own labor market. Moreover, job seekers are assumed to have an equal probability of being matched with any open vacancy within their labor market, regardless of their geographical or occupational proximity. In the comparisons, we use the following simplified version of the index which assumes that the labor markets k are homogeneous in terms of matching efficiency:

$$I_{\c{S},t} = 1 - \sum_{k \in K} \left(\frac{V_{kt}}{V_t} \right)^{\eta} \left(\frac{U_{kt}}{U_t} \right)^{1-\eta}, \tag{10}$$

where η is a vacancy share parameter which determines how the number of matches depend on local labor market tightness. Following previous studies (e.g. Şahin et al. 2014, Alasalmi 2022, Pizzinelli and Shibata 2023), we assume that $\eta = 0.5$.

As the second alternative index, we use the distinct-markets index of Marinescu and Rathelot (2018) which only differs from the interconnected-markets index in that job seekers are assumed to apply for jobs exclusively within their local labor market. This amounts to assuming that $g(d_{ii}) = 1$ and $g(d_{ij}) = 0$ for $i \neq j$, in which case equation (9) simplifies as:

$$I_{D,t} = 1 - \sum_{k \in K} \frac{U_{kt}}{\overline{U}_t} \frac{1 - exp(-q_t \overline{a}_t \mathcal{R}(q_t \overline{a}_t U_{kt}/V_{kt})}{1 - exp(-q_t \overline{a}_t \mathcal{R}(q_t \overline{a}_t \overline{U}_t/\overline{V}_t)} \tag{11}$$

Marinescu and Rathelot (2018) show that this distinct-markets index often leads to significantly higher mismatch estimates than the interconnected-markets index. It differs from the index of Şahin et al. (2014) by using a different (urn-ball) matching function and by taking into account changes in the monthly job finding rate. Thus, these approaches may provide different results.

3 Results

3.1 Geographical mismatch

The results regarding job seekers' distaste for geographical distance, obtained by estimating equation (1) for each year, are shown for selected years in Table 1. The parameters $\gamma_1 - \gamma_6$ describe the change in the probability of employment between the knots of the piecewise-linear spline function at 20, 50, 100, 200 and 400 kilometers. For example, a one-kilometer increase in distance decreases the relative probability of employment by $\exp(\gamma_1)$. The results are very similar across years, indicating no large changes in distaste for geographical distance. Using the 2020 estimates, relative probability of employment is $\exp(20 \times -0.1228) = 8.6$ % at 20 kilometers, suggesting that probability of employment decreases rapidly with distance. For distances above 20 km, the parameters are summed, i.e., at 50 kilometers the probability is $\exp(20 \times -0.1228 + 30 \times (-0.1228 + 0.0541)) = 1.1$ %.

The results of Table 1 are visualized in Figure 2. The relative probability of employment decreases rapidly with distance, already approaching zero at 50 kilometers. The estimates are shown for years 2008, 2011, 2014, 2017 and 2020 to illustrate the change in mobility during the period of interest. Generally, the curve becomes slightly less steep over the years, indicating a modest decrease in the distaste for geographical distance. The distaste estimates of Marinescu and Rathelot (2018) obtained using the US job application data are included in Figure 2 for comparison. As illustrated by the flatter curve, their results indicate a significantly lower distaste for distance than our results based on Finnish register data.

The evolution of the interconnected-markets geographical mismatch index between 2006 and 2021 is shown in Figure 3. These results indicate that, depending on the observation period, limitations in geographical mobility have contributed to 0.5 %–5 % smaller number of matches as compared to the theoretical maximum number. There is large seasonal variation in the index values, with elevated values observed particularly in the summer season, potentially due to an increased preva-

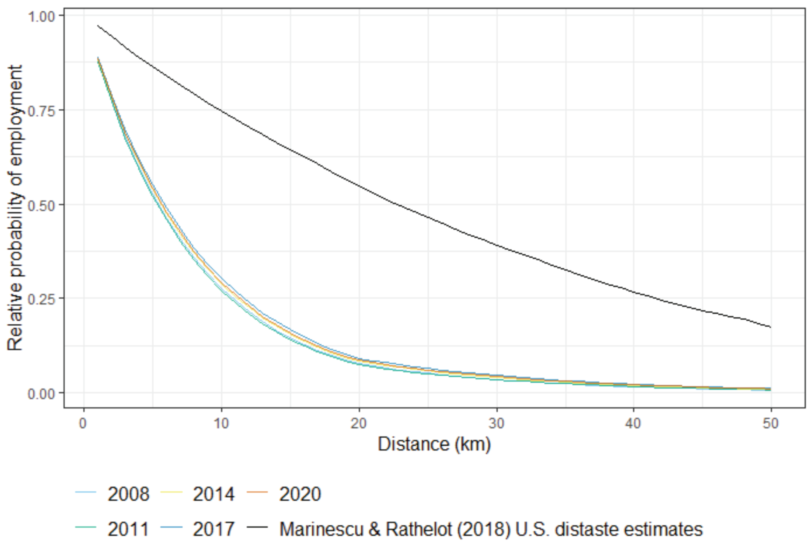

Figure 2: Job seekers' relative probability of being matched to a job as a function of the job's geographical distance: Poisson regression results

Notes: The figure is limited to 50 kilometers for better visual representation. The estimates of Marinescu & Rathelot (2018) from a model specification with job seeker fixed effects are included in the figure as a reference point.

Table 1: Distaste for geographical distance, selected years

Parameter	2008	2011	2014	2017	2020
γ_1	-0.1283***	-0.1304***	-0.1240***	-0.1194***	-0.1228***
	(0.0039)	(0.0044)	(0.0045)	(0.0047)	(0.0049)
γ_2	0.0508***	0.0564***	0.0519***	0.0507***	0.0541***
	(0.053)	(0.059)	(0.0060)	(0.0064)	(0.0065)
γ_3	0.0453***	0.0397***	0.0382***	0.0348***	0.0359***
	(0.0030)	(0.030)	(0.0031)	(0.0032)	(0.0032)
γ_4	0.0213***	0.0233***	0.0225***	0.0234***	0.0219***
	(0.0020)	(0.0020)	(0.0021)	(0.0020)	(0.0021)
γ_5	0.0071***	0.0067***	0.0071***	0.0059***	0.0064***
	(0.0013)	(0.0013)	(0.0014)	(0.0014)	(0.0014)
γ_6	0.0023***	0.0030***	0.0031***	0.0038***	0.0037***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0009)
N	204,864	222,967	220,821	236,130	209,026

Notes: Selected Poisson regression estimates of equation (1). The nodes of the spline function are at 20, 50, 100, 200 and 400 km. A one kilometer increase in distance decreases probability of employment by $\exp(\gamma_1)$ for distances under 20 km. For distances between 20 and 50 km, a kilometer increase decreases the probability by $\exp(\gamma_1 + \gamma_2)$ etc. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

lence of seasonal job openings. The seasonally adjusted trend remains relatively stable, approximately at 2.5-3.0 %, in the beginning of the observation period until 2014. Subsequently, mismatch starts to decrease, reaching a level of 1.5 % by the end of the period.

In Figures 4 and 5, the interconnected-markets mismatch index is compared to the two distinct-markets indices described in equations (10) and (11), which are calculated using three alternative definitions of the local labor market: region, sub-region and municipality. Using the comparison indices, the level of geographical mismatch varies based on the assumed scope of the labor market, being the higher the smaller the size of the local labor market is assumed to be.

In line with our our main index, the comparison indices show a decreasing trend

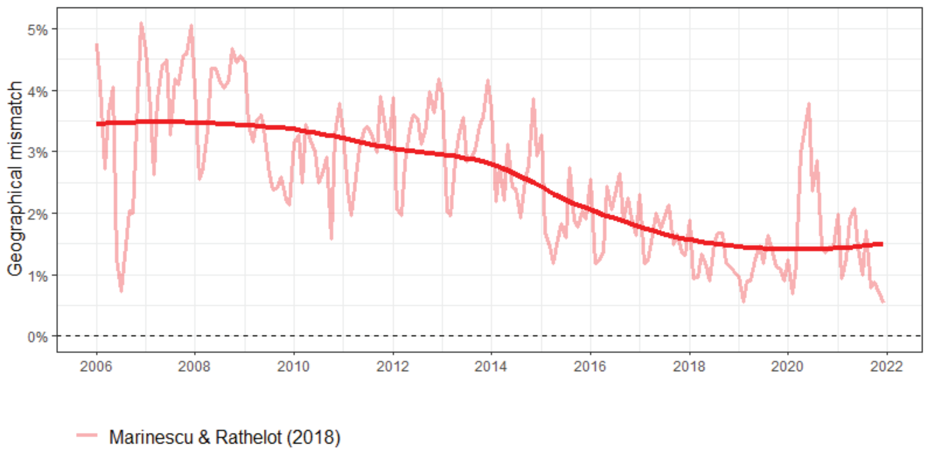

Figure 3: Geographical mismatch index

in mismatch. However, the indicated timing of the decrease varies, to some extent, between the indices of Şahin et al. (2014) and Marinescu and Rathelot (2018): while the first suggest that most of the decrease occurred already before 2012, the latter indicate a significant decrease from 2014 onwards. Importantly, we observe significant convergence in the trend lines of the different indices towards the end of the observation, where they fit within 2 percentage points. Thus, we can conclude that observations on the level of geographical mismatch have, in recent years, become somewhat insensitive to the choice of the mismatch index. Overall, the trend line for the interconnected-markets index aligns most closely with those of the sub-region-level Şahin et al. (2014) index and the region-level distinct-markets index of Marinescu and Rathelot (2018). Compared with the other indices, the municipality-level comparison indices indicate a higher levels of mismatch, particularly in the early observation period (2006–2008), where even 7% to 16% of hires are indicated to have been lost due to the geographical misallocation of job seekers and jobs. The municipality-level indices also exhibit a somewhat different trend in the early observation period due to a sharper decrease in the level of geographical mismatch in around 2009.

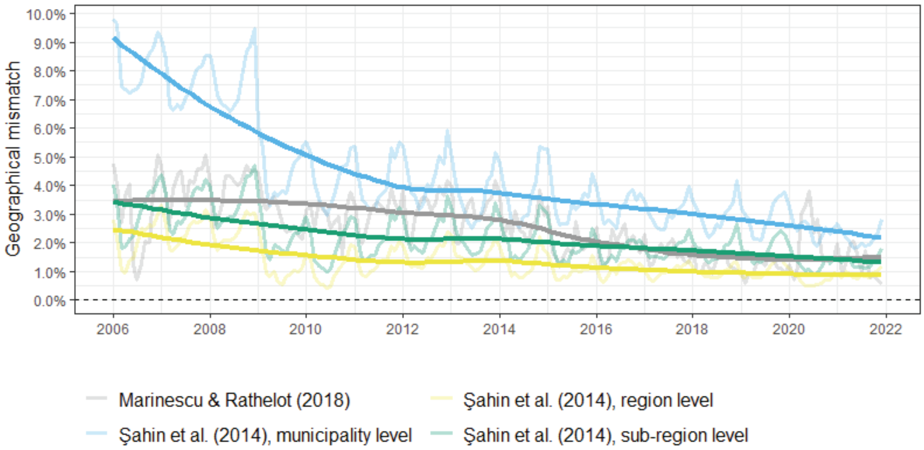

Figure 4: Geographical mismatch: comparison to the mismatch index of Şahin et al. (2014)

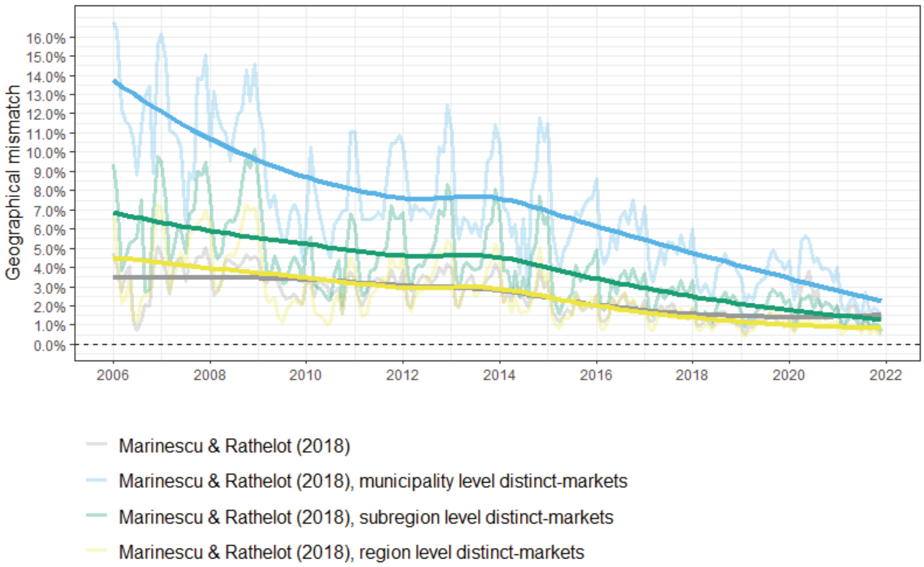

Figure 5: Geographical mismatch: comparison to the distinct-market mismatch index of Marinescu and Rathelot (2018)

Figures 4 and 5 further shows that, overall, the comparison indices indicate a more volatile development of geographical mismatch than the interconnected-markets index. However, whereas the volatility of the comparison indices remains roughly similar during the COVID-19 pandemic, the values of the interconnected-markets index of Marinescu and Rathelot (2018) clearly deviate from the trend line and indicate a sharp but short-lived increase in geographical mismatch in the spring of 2020. Then again, the interconnected-markets index indicates a more stable development of geographical mismatch during the years of the global financial crisis than the comparison indices.

3.2 Occupational mismatch

Table 2 shows the results of estimating job seekers' distaste for occupational distance (equation (2)), i.e. their willingness to switch to a different occupation. Coefficients α_{11} and α_{12} measure how much a job seeker's willingness to occupational mobility is affected by the difference between the job seeker's previous occupation code and the vacancy's occupation. Using the estimates of 2020, a one-digit difference in the occupation codes implies that the relative probability of employment decreases to $exp(-1.3559) = 26.3\%$. Similarly, if the occupation codes are completely different, this probability drops to $exp(-1.8665) = 15.5\%$. α_2 is the coefficient for the variable measuring the difference in workers' educational backgrounds of the previous and new occupations. Similarly to the estimated distaste-for-geographical-distaste parameters, Table 2 shows that the distaste-for-occupational-distance parameters α remain highly stable throughout the analysis period.

The occupation-only mismatch index is shown in Figure 6. Occupational mismatch ranges from less than 0.5 % to over 6 % during the observation period. The seasonally adjusted trend increases from 1 % in 2006 to 4 % in 2012, after which it starts to decrease, returning to a level of 1 % at the end of the observation period.[2]

[2] A possible caveat of the occupational mismatch estimates in Figure 6 is that they are based on the relatively broad 2-digit-level occupations and, therefore, might mask part of the occupational mismatch taking place across more narrowly defined occupational categories. As a robustness check, we also conducted the analysis of occupational mismatch using the 3-digit occupations. These results are highly identical to those in Figure 6, indicating only around a 0.5–1-percentage-point higher level of occupational mismatch.

Table 2: Distaste for occupational distance, selected years

Parameter	2014	2017	2020
α_{11}	-1.4162***	-1.5503***	-1.3559***
α_{12}	-1.8277***	-2.0153***	-1.8665***
α_2	-0.0546***	-0.0520***	-0.0524***
N	210,871	232,652	208,604

Notes: Selected Poisson regression estimates of equation (2). α_{11} and α_{12} are the coefficients for dummy variables indicating 1-digit and 2-digit differences, respectively, between the occupations of the job seeker and the vacancy. α_2 is the coefficient for a variable measuring the skill difference between the job seeker and the vacancy which is based on the educational backgrounds of workers in different occupations. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

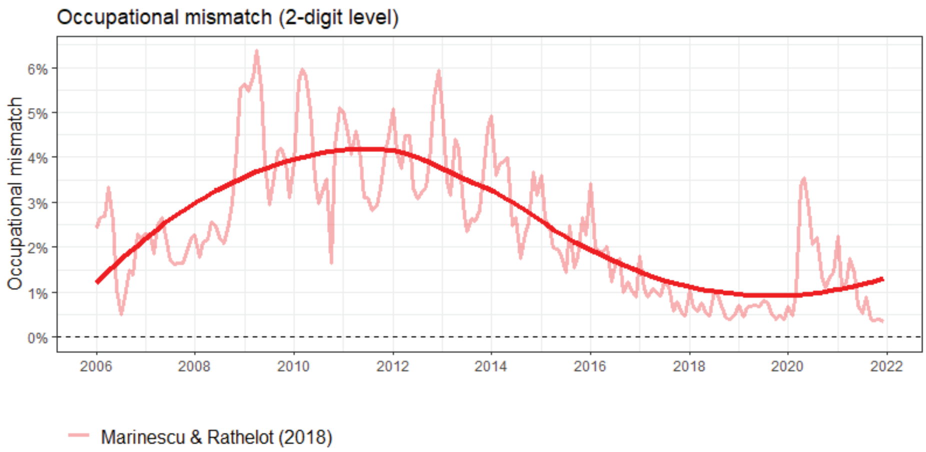

Figure 6: Occupational (2-digit-level) mismatch

For comparison, the distinct-markets indices of Şahin et al. (2014) (Figure 7) and Marinescu and Rathelot (2018) (Figure 8) are computed at the 1-digit and 2-digit levels of the occupational classification. Our interconnected-markets index is lower than any of the comparison indices. Assuming that job seekers only apply for vacancies within their own occupation code, even at the 1-digit level, can thus be too restrictive. Similar to the case of geographical mismatch, the distinct-markets ap-

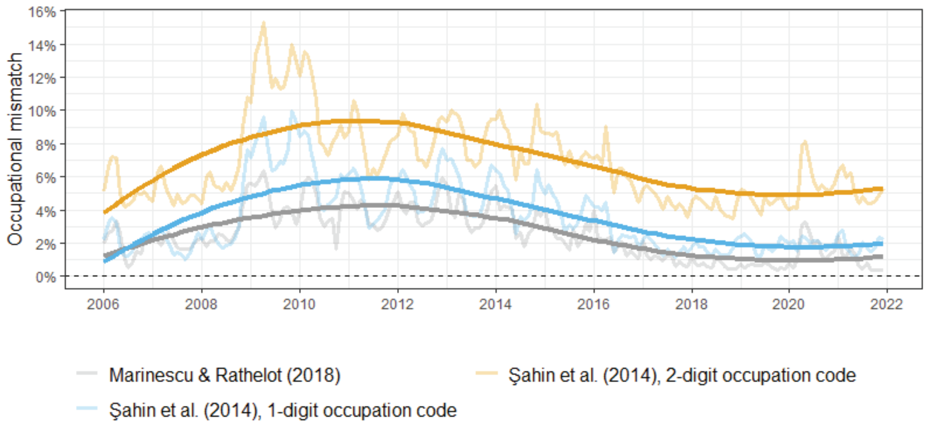

Figure 7: Occupational (2-digit level) mismatch: comparison to the mismatch index of Şahin et al. (2014)

proach of Marinescu and Rathelot (2018) results in the highest mismatch estimates. However, different indices mainly point towards similar temporal changes in occupational mismatch, even more so than in the case of analysing geographical mismatch. The differences between our main index and the Şahin et al. (2014) indices remain highly stable across the observation period, whereas the distinct-markets indices of Marinescu and Rathelot (2018) converge towards the main index in the latter half of the observation period.

3.3 Geographical and occupational mismatch

Table A.1 in the Appendix shows the results from the estimation of job seekers' distaste for both geographical and occupational distance simultaneously (equation (3)). Distaste for geographical distance is measured in the same way as previously, with parameters $\gamma_1 - \gamma_6$ describing changes in the spline function of distance at the different knots. We can see that the estimates for $\gamma_1 - \gamma_6$ are in the same ballpark but slightly lower than those in Table 1. The estimates for α_{11}, α_{12} and α_{22} are close to the results of Table 2.

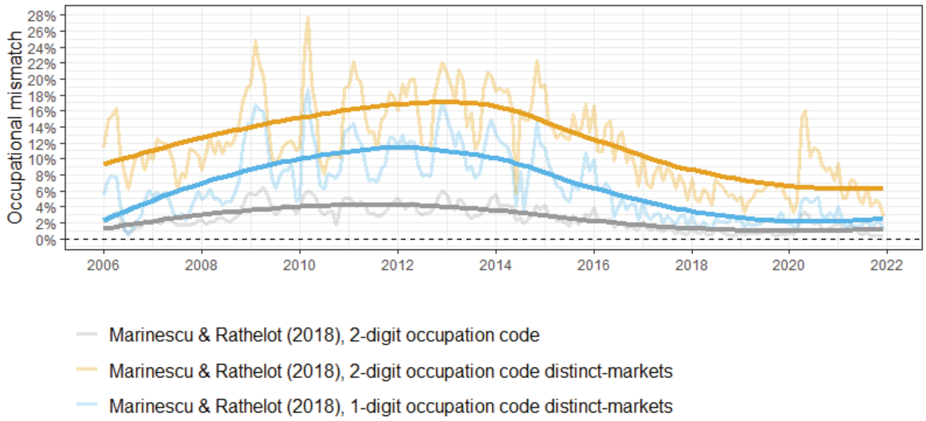

Figure 8: Occupational (2-digit level) mismatch: comparison to the distinct-market mismatch index of Marinescu and Rathelot (2018)

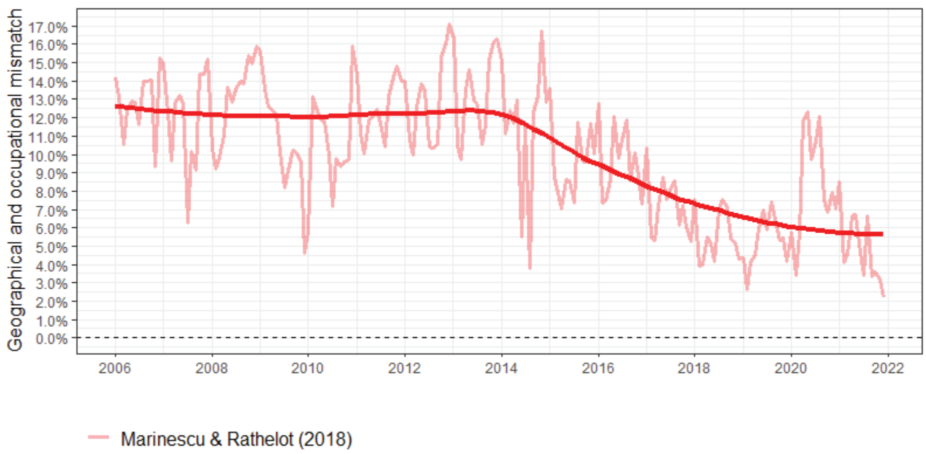

Figure 9: Geographical and occupational mismatch index

The combined geographical and occupational mismatch index is depicted in Figure 9. As expected, the combined index is higher than the separate indices for geographical and occupational mismatch, with values ranging from 2 % to 17 %. The trend of the index remains highly stable at around 12–13 % between 2006 and 2014. However, after 2014, the combined effect of geographical and occupational mismatch decreases significantly, falling below 5 % in 2021.

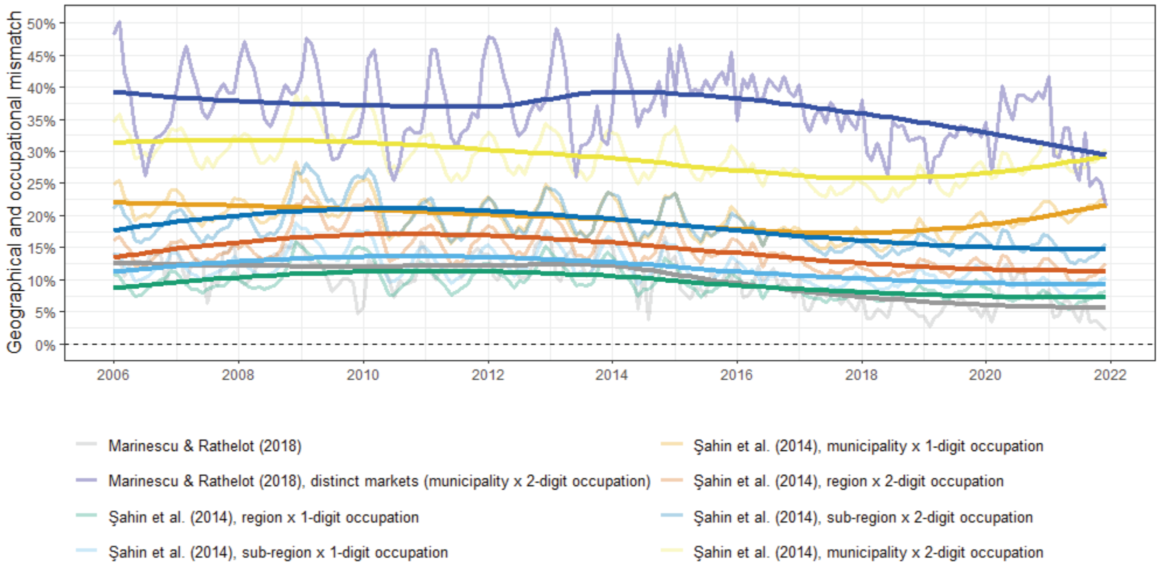

Figure 10: Geographical and occupational mismatch index: comparison of indices

In Figure 10, the results using the interconnected-markets index of Marinescu and Rathelot (2018) are again compared to several alternative indices. The index of Şahin et al. (2014) is calculated with several different labor market definitions, varying the levels of the occupational classification (1 or 2 digits) and the regional classification (municipality, sub-region or region). We observe that our main index gives lower estimates for the combination of geographical and occupational mismatch compared to most of the comparison indices, which likely overestimate the extent of mismatch due the restrictive assumption that job seekers only search within a specific geographical area and occupation. Our main index is closest to the least restrictive distinct-markets indices, where the definition of the local labor market is based on the 1-digit occupation code and either region or sub-region. The distinct-markets index of Marinescu and Rathelot (2018) based on the municipality- and 2-digit-occupation-level variation of labor market tightness, gives the highest estimates reaching up to 50 %. Our main index and the comparison indices mainly indicate a similar development in the combined geographical and occupational mismatch. However, the estimated change in the trend of the index around 2014 and the response of the index to the COVID-19 pandemic are more pronounced when using the approach of Marinescu and Rathelot (2018) than with the approach of Şahin et al. (2014).

3.4 Mismatch and economic crises

Previous research on mismatch unemployment is closely linked to the study of labor market dynamics during economic crises. For example, the studies of Şahin et al. (2014), Patterson et al. (2016) and Turrell et al. (2021) have been motivated by the large unemployment effects of the Great Recession in the US and the UK, while Pizzinelli and Shibata (2023) extended this analysis to the COVID-19 pandemic.

While previous sub-sections have described the general development of geographical and occupational mismatch, this sub-section examines more closely the years around the two aforementioned global crises: the 2007–2009 global financial crisis and the COVID-19 pandemic. The job destruction and creation caused by these crises were allocated differently across regions and occupations, which raises questions about the magnitude and the duration of the subsequent misallocation between job seekers and vacancies.

Şahin et al. (2014) studied the development of geographical and occupational mismatch in the US during the 2008 financial crisis, using the mismatch model derived in the article. The study finds that occupational mismatch increased significantly during the crisis, whereas geographical mismatch remained relatively constant. Pizzinelli and Shibata (2023) applied a modified version of the mismatch index by Şahin et al. (2014) to data during the COVID-19 pandemic in the US and UK. They found that the industry-level mismatch increased sharply at the beginning of the pandemic, but the increase was temporary and the level of mismatch normalized after a few quarters.

To illustrate the changes in our main mismatch models during these crises, we

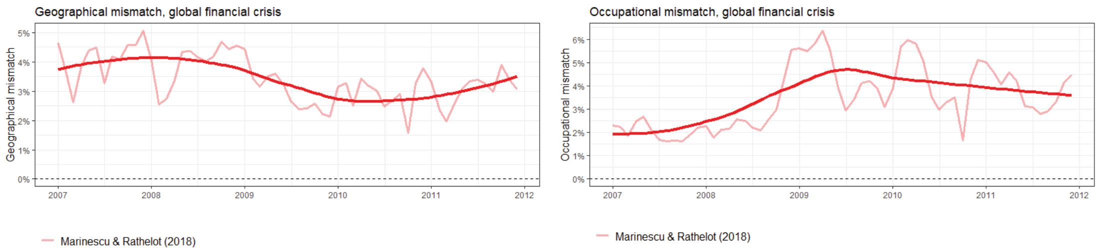

Figure 11: Mismatch during the global financial crisis

present separate graphs for the crisis periods. Since we estimate our distaste-for-distance parameters yearly, using data on two previous years as well, the index changes are potentially smaller than if job seeker's preferences were allowed to be more volatile.

Figure 11 shows the development of the geographical and occupational mismatch indices during the global financial crisis. Perhaps surprisingly, the figure shows that the role of geographical mismatch actually decreased slightly during this time period. As seen in Figure 4, the same conclusion is obtained when using the index by Şahin et al. (2014) as well. However, occupational mismatch shows a significant increase in late 2008 and early 2009. As seen in Figure 11, as well as Figure 6, occupational mismatch remained elevated for several years after the recession.

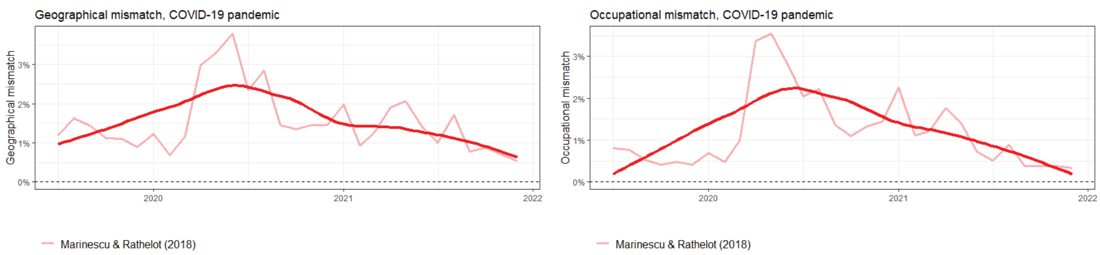

Figure 12: Mismatch during the COVID-19 pandemic

Figure 12 illustrates the development of our main mismatch models during the COVID-19 pandemic. According to the results, both geographical and occupational mismatch increased significantly at the onset of the pandemic in the spring of 2020. However, the mismatch levels appear to have returned back to pre-pandemic levels during 2021. The fairly short period of increased geographical and occupational mismatch is in line with the observations of Pizzinelli and Shibata (2023).

3.5 Mechanisms

The results above suggest that the trends in both geographical and occupational mismatch have shifted during the last decade. This implies that the share of unemployment attributable to mobility restrictions has changed, with the general trend being decreasing. In this section, we examine factors that could possibly explain the changes.

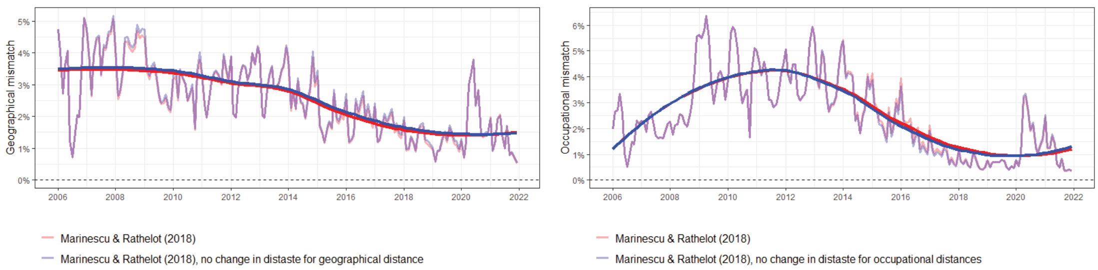

Figure 13: The development of geographical and occupational mismatch with and without allowing for the distaste for distance to change

Figure 13 shows how the indices would have developed if mobility had remained at the 2006 level. We can see that the level and changes of geographical mismatch are approximately similar to those implied by our main estimates. This suggests that changes in job seekers' job search behavior play almost no part in the decreasing role of geographical mismatch.

As we can only estimate changes in occupational distaste from 2010 onward, our possibilities for analyzing the sensitivity of the occupational mismatch index are more limited. Nonetheless, Figure 10 demonstrates that the variations in distaste-for-occupational distance after 2010 account for only a tiny proportion of the observed changes in occupational mismatch.

Since changes in job seekers' distaste for distance do not appear to explain the bulk of the changes in mismatch, it is possible that role of the geographical and occupational misallocation has decreased because job seekers and vacancies are increasingly closer to each other. Marinescu and Rathelot (2018) use the close proximity of job seekers and vacancies as the primary explanation for the low mismatch levels they find.

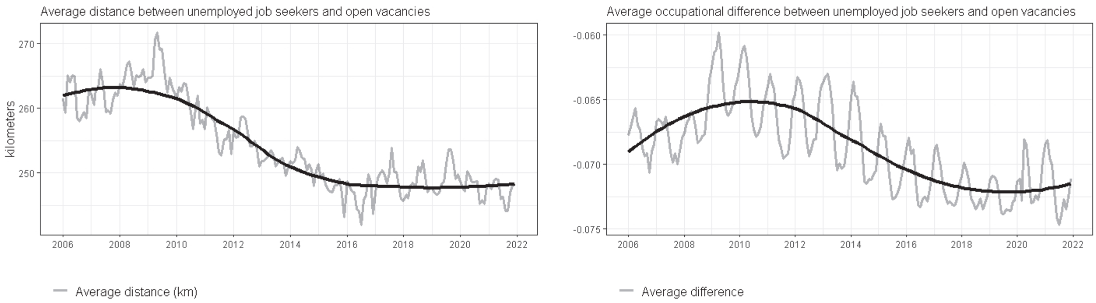

Figure 14: Average geographical and occupational distance of job seekers and open vacancies

Figure 14 shows how the average geographical and occupational distance between unemployed job seekers and open vacancies has developed during the observation period. The average geographical distance is calculated as the average distance in kilometers between all job seeker-vacancy pairs. Occupational distances are calculated similarly, but the distance between the 2-digit occupation codes of job seekers and open vacancies are measured using the method described in Section 2.2.

We can see that the average geographical distance has decreased during the time period, preceding the decrease in the geographical mismatch index. The change in the average distance is modest: from around 260 kilometers to less than 250 kilometers. However, since unemployed job seekers have a limited job search radius, distance to all open vacancies is not the most relevant statistic, and the change may be more significant for individual job seekers. The change in occupational distances also roughly resembles and precedes the development found in the occupational mismatch index, with an increase between 2006 and 2010, and a subsequent decrease during the 2010s.

It appears that the increased concentration of job seekers and vacancies may explain part of the changes in mismatch unemployment. To see if the change is more due to increased concentration of job seekers or open vacancies, we study their development separately. We examine the regional concentration of job seekers and open vacancies using the Herfindahl index:

$$HHI = \sum_i (s_i)^2, \tag{12}$$

where s_i is the share of either open vacancies or job seekers in municipality i. Since the shares are squared, higher shares increase the index disproportionately more than lower shares, so that an increase in HHI implies an increase in regional concentration.

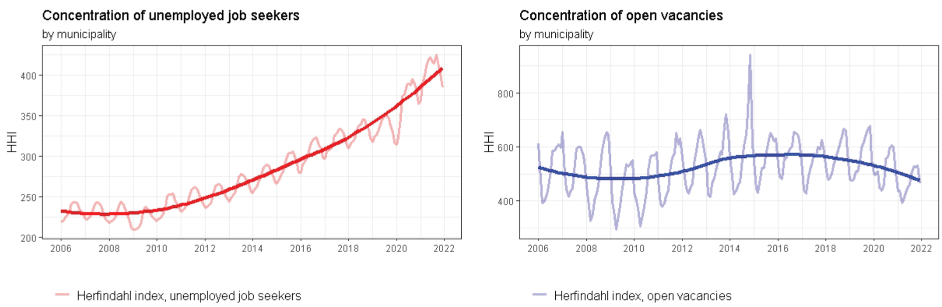

Figure 15: Regional concentration of job seekers and open vacancies

Figure 15 shows the development of the regional Herfindahl index. Since the number of municipalities in Finland has decreased during the time period, which may affect the results, the numbers of job seekers and vacancies are aggregated at the level of the 2021 municipalities. The concentration of unemployed job seekers has increased significantly during the observation period, whereas there is no clear increase in the concentration of open vacancies. Thus, it appears that the decreased average distance between unemployed job seekers and vacancies is mainly explained by the concentration of job seekers in the same municipalities. These observations are line with the previous observations of Alasalmi (2022) who investigated the changes in the share of the largest municipalities of all job seekers and vacancies.

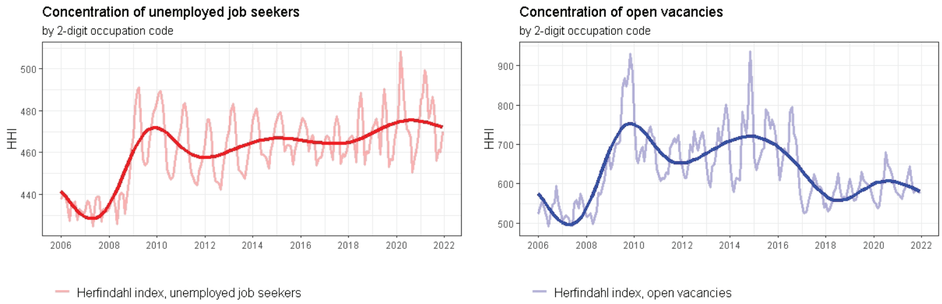

Figure 16: Occupational (2-digit-level) concentration of job seekers and open vacancies

We also calculate the HHI across the 2-digit occupation codes. Thus, in this case, s_i is the share of either open vacancies or job seekers in the 2-digit occupation i. These results in Figure 16 indicate a noticeable increase in the occupational concentration of job seekers and vacancies around the years 2008–2010, which is temporally related to the global financial crisis. Afterwards, the concentration of job seekers remains relatively constant, whereas the concentration of open vacancies shows a generally decreasing trend towards the end of the observation period. Thus, it seems that the increased occupational distances between job seekers and vacancies in the beginning of the observation period was caused by both vacancy- and job-seeker-side factors, whereas the decrease in the 2010s is more likely linked to the occupational diversification of open vacancies.

In summary, it appears that the estimated changes in geographical and occupational mismatch are more likely due to changes in the regional and occupational distribution of job seekers and vacancies rather than to changes in job seekers' probability of accepting distant jobs. In particular, the increased regional concentration of unemployed job seekers provides a compelling explanation for the decrease in geographical mismatch.

3.6 Geographical mismatch by occupational group

To study heterogeneity in geographical mismatch between occupational groups, we calculate the geographical mismatch index separately for different 1-digit occupation codes. To do this, we assume, for simplicity, that job seekers only apply for jobs within their previous occupational group. For this reason, the level of geographical mismatch is likely overestimated, but the results should still shed light on occupational heterogeneity. The yearly distaste estimates $g(\cdot)$ are calculated for each group separately, whereas the parameter $q_t \bar{a}_t$ is determined at the aggregate level.

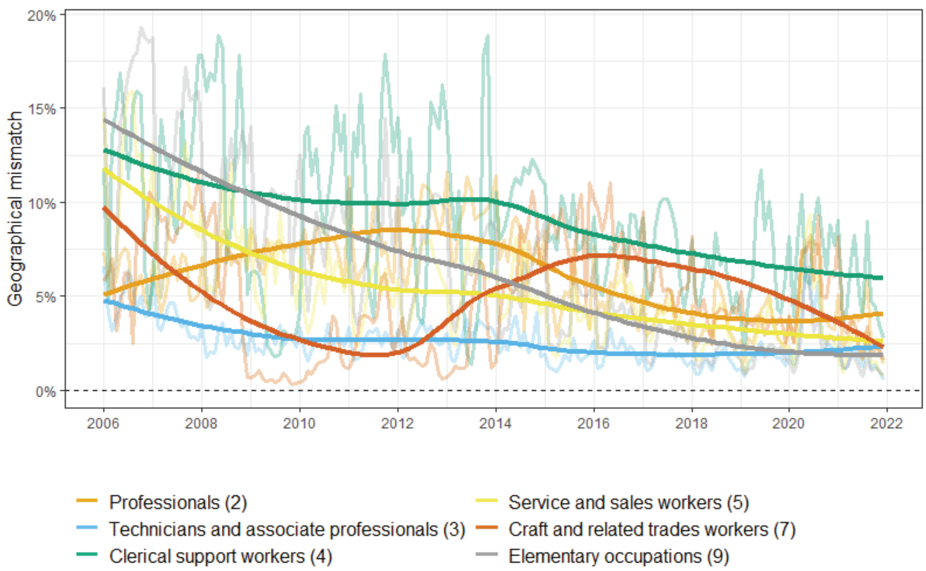

Figure 17: Geographical mismatch by occupational group

The geographical mismatch indices for selected occupational groups are depicted in Figure 17. We see that the estimated levels of mismatch are highly volatile for many of the occupational groups, particularly in the early observation period prior to 2014. Nevertheless, there are clear occupational differences both in the level of geographical mismatch and its development. For most of the time, mismatch has been the highest for the group of clerical support workers, which includes occupations that have suffered from labor oversupply. Mismatch has mostly been the lowest for technicians and associate professionals, for whom the level of mismatch has also remained relatively stable.

In Figure 17, we observe significant convergence of the mismatch indices across the groups towards the end of the observation period, coinciding with the general decreasing trend in geographical mismatch. After this convergence, all of the trend lines fit within 5 percentage points (2%–7%). According to the results, geographical mismatch has decreased especially for clerical support workers, service and sales workers and elementary occupations. These fields encompass a large portion of the working population, and it seems that these occupational groups explain the bulk of the decrease in the overall geographical mismatch index over the observation period.

The development of mismatch deviates, to some extent, from the overall development for two occupational groups: craft and related trades (including construction workers) and professionals. Among craft and related trades workers, mismatch decreased significantly during the global financial crisis, but returned to the initial level by the mid-2010s. Professionals experienced an opposite development, with an increase followed by a decrease. We cannot, with absolute certainty, identify the factors that drive these occupation-specific trends. However, as the construction industry was heavily affected by the global recession, the levelling-off regional differences in construction labor shortage are likely to partially drive the declining mismatch for craft and related trades workers in the early observation period. The recovery of the construction industry from the recession may again explain the increase in geographical mismatch afterwards. The increased mismatch among professionals in the late 2000s and early 2010s is again temporally related to both the financial crisis and the following difficulties of the Finnish technology sector, which offer potential explanations for this development.

Figure 18 shows the regional Herfindahl indices for job seekers and open vacancies within 1-digit occupations. According to these results, the regional concentration of job seekers has increased in all occupational groups, which likely explains at least part of the overall decrease in the occupation-specific mismatch indices. The

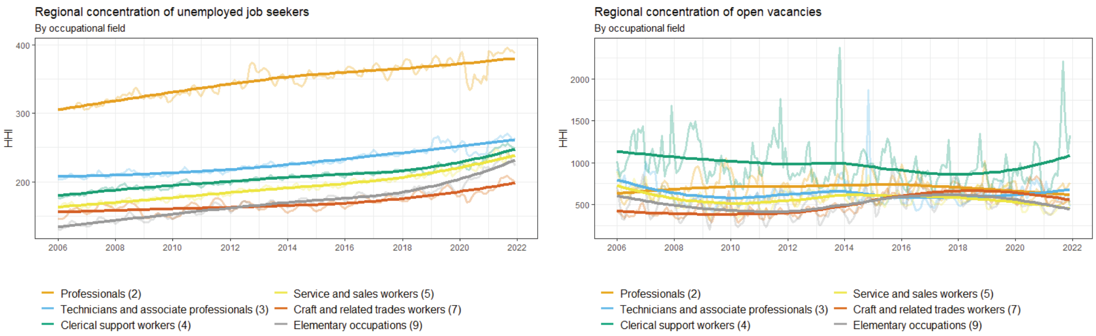

Figure 18: Regional Herfindahl index by occupational group

concentration of open vacancies has remained relatively stable for most occupations. However, this index increased for craft and related trades workers after 2012, which is temporally related to the increased regional mismatch in this occupational group. However, it does not seem that the development of mismatch for either professionals or crafts and related trades workers can be well explained by changes in the regional distribution of job seekers or vacancies within these occupational groups.

4 Validity and robustness

4.1 Sensitivity to distaste estimates

As we infer job seekers' preferences based on accepted job offers, i.e., a selected sample of their choices, our estimates for the distaste for distance and mismatch could be biased upwards or downwards. This could be the case, in particular, when measuring occupational mobility: while we attempt to measure how willing job seekers are to change occupations, we cannot separate job seekers' occupational preferences from employers' preferences. Therefore, the distaste for occupational distance could be overestimated.

To assess the sensitivity of our results, we recalculate our mismatch indices by varying the levels of distaste for distance. We multiply the baseline estimates with a varying coefficient to illustrate the indices' sensitivity to our mobility estimates. A coefficient higher than 1 implies decreased mobility. When the coefficient is smaller than 1, mobility is increased. Figure 19 demonstrates that the level of geographical mismatch is moderately sensitive to variations in the baseline distaste for distance. For example, increasing the distaste parameter by 25% results in, at most, a one-percentage-point increase in the level of geographical mismatch. These deviations do not significantly alter our view of the overall development of geographical mismatch.

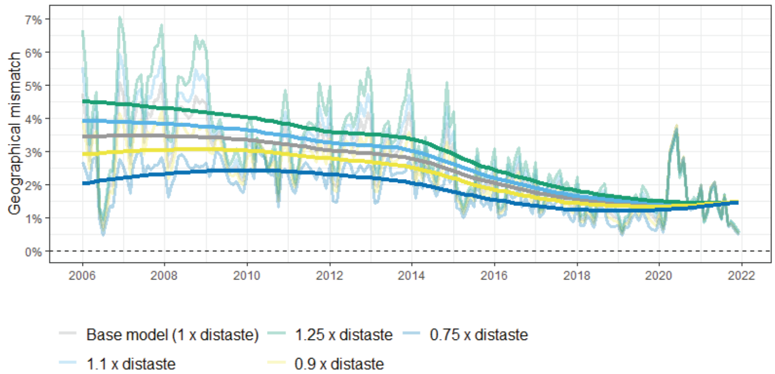

Figure 19: Geographical mismatch: sensitivity to distaste for distance

Figure 20 compares our main geographical mismatch index to one obtained by using the distaste-for-distance estimates of Marinescu and Rathelot (2018) based on US job application data and a Poisson regression model with job seeker fixed effects. As described in Figure 2 above, these distaste-for-distance estimates are significantly lower than our estimates based on Finnish employer-employee data. However, as demonstrated in Figure 20, our conclusions are somewhat robust to the estimated degree of distaste for distance, as the parameters of Marinescu and Rathelot (2018) result in mismatch estimates that only around 0.5–1 percentage points lower than our main estimates.

Due to the aforementioned issues in measuring the distaste for occupational distances with realized job matches, it is possible that the true level of occupational mismatch is below the estimated baseline level. However, Figure 21 shows that

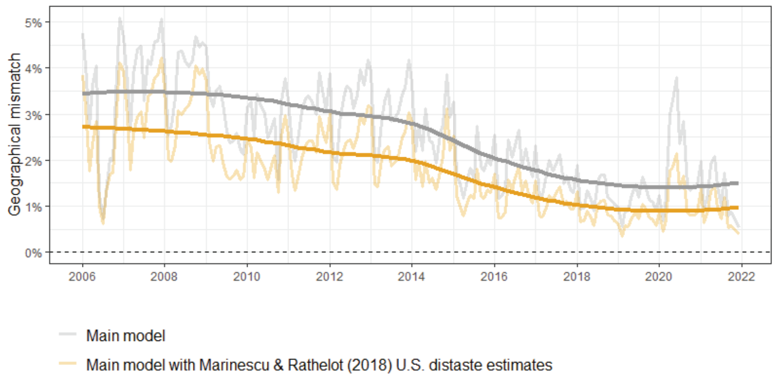

Figure 20: Geographical mismatch using the distaste-for-distance estimates of Marinescu and Rathelot (2018)

the implied level of occupational mismatch does not react dramatically to modest changes in the distaste for occupational distances. Thus, overall, it appears that that small to moderate measurement errors in the distaste for distance do not significantly alter the conclusions of our analysis.

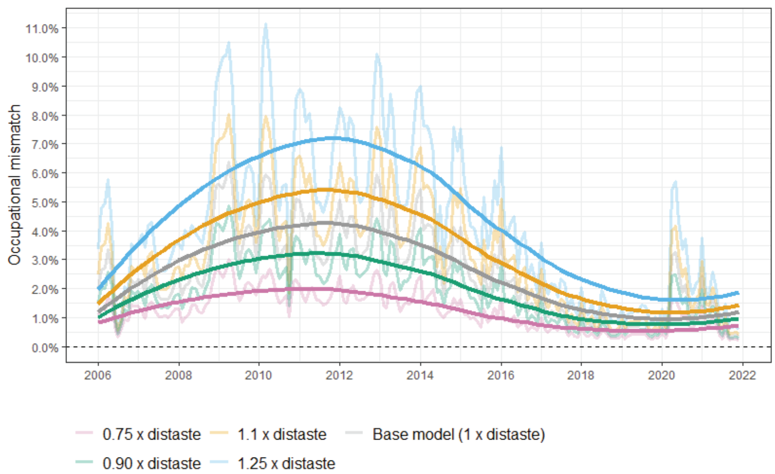

Figure 21: Occupational mismatch: sensitivity to distaste for occupational distance

4.2 Explanatory power of local labor market tightness measures

Being based on relatively weak assumptions on job seekers' job search behavior, the interconnected-markets approach of Marinescu and Rathelot (2018) provides a theoretically appealing framework for examining labor market mismatch. The purpose of this sub-section is to present empirical evidence for assessing the validity of different approaches. For this, we examine the explanatory power of alternative measures of local labor market tightness – which the mismatch indices are based on – over a relevant outcome measure: the duration of job seekers' unemployment spells.

The conventional approach for assessing the tightness of the local labor market in a specific region – the ratio of open vacancies to unemployed job seekers – amounts to taking into account only the vacancies and job seekers located in that region and placing an equal weight for all job seekers and vacancies, regardless of their location within the region. However, as job seekers are also likely to consider vacancies outside their region of residence, and to place a different weight for more and less distant jobs within the region, the conventional approach may not provide a very accurate representation of the local labor market tightness.

Marinescu and Rathelot (2018) introduced the *generalized labor market tightness* which accounts for the geography of job search in a more realistic way: when measuring labor market tightness for a specific municipality j in Finland, the numbers of job seekers and vacancies in all Finnish municipalities are taken into account, but these figures are weighted according to the distaste-for-distance estimates $g(\cdot)$. That is, the generalized labor market tightness captures the job opportunities and potential job seekers located in surrounding areas as well, taking into account job seekers' preference for nearby jobs. For a better graphical representation, we focus on the generalized *inverse* tightness, the ratio of job seekers to vacancies, which for area j is:

$$\nu_j = \sum_k \frac{g(d_{kj})U_k}{\sum_\ell g(d_{k\ell})V_\ell}. \tag{13}$$

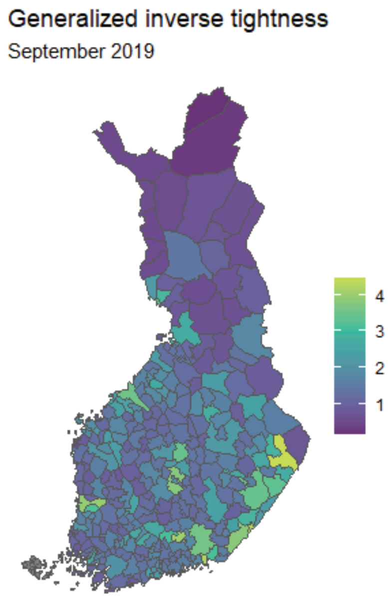

Figure 22: Generalized inverse tightness, municipality level

Figure 22 shows the generalized inverse tightness and the regular inverse tightness estimates at the municipality level for a chosen example period: September 2019. The ratio of unemployed job seekers to open vacancies ranges from less than 1 to over 4. Notably, generalized inverse tightness is particularly high in certain municipalities in eastern Finland, while Lapland demonstrates comparatively lower values. In Figure 23, we compare these outcomes to the conventional municipality-level inverse tightness on a merged scale. The figure shows that the generalized tightness measure indicates smoother differences across municipalities. Some municipalities have very high or low tightness when measured in the conventional way, but these extremities diminish when accounting for mobility to and from other municipalities.

Figure A1 compares the municipality-level generalized inverse tightness measure

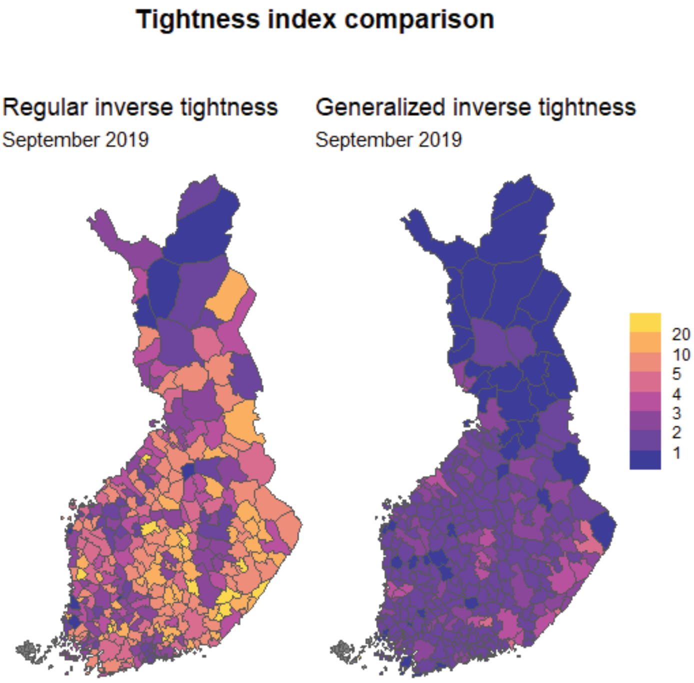

Figure 23: Generalized inverse tightness, comparison to regular inverse tightness

to the regular inverse tightness generated using region-level data on the number of job seekers and vacancies. In this case, the two approaches provide more similar estimates, but many distinct differences remain. For instance, the regular measure appears to significantly underestimate the tightness of the local labor market in the Lapland region.

To assess the predictive power of the generalized inverse tightness measure, we investigate its relationship with the duration of unemployment spells. Based on Statistics Finland register data, We compute the length of unemployment spells (in days) in Finland between 2011 and 2020. We only include unemployment spells that end due to employment. If an unemployment spell extends beyond the end of our data, December 2021, we extend the employment spell until the first of July 2022. Using OLS regression, we explain the length of a job seeker's unemployment spell by the generalized inverse tightness of the labor market in the job seeker's municipality of residence at the beginning of her unemployment spell. We conduct a similar analysis using the conventional measure for the municipality-level inverse labor market tightness as the independent variable.

As more job seekers and/or less open vacancies increase labor market competitiveness, there is reason to expect that the inverse tightness measures are positively associated with the duration of the unemployment spell. However, in the data, both the generalized and the conventional measure of inverse tightness are, in fact, negatively correlated with unemployment spell length. While this seems counter-intuitive, it is possible these correlations are confounded e.g. by a non-random sorting of job seekers into municipalities with more and less tight labor markets. To study this possibility, we estimate OLS regressions that control for job seekers' characteristics, including age, gender, language, education and occupation. The control variables are measured based on the latest available information prior to the beginning of each unemployment spell.

Table 3: Unemployment spell length and alternative measures of inverse labor market tightness

	Unemployment spell length					
	(1)	(2)	(3)	(4)	(5)	(6)
Generalized inverse tightness	-0.265***		0.169**		0.371***	
	(0.075)		(0.081)		(0.097)	
Regular inverse tightness		-1.19***		-1.54***		-0.877***
		(0.022)		(0.027)		(0.032)
Control variables:						
Age			✓	✓	✓	✓
Gender			✓	✓	✓	✓
Language			✓	✓	✓	✓
Educational level					✓	✓
Previous occupation					✓	✓

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In line with the raw correlations, the OLS regression results presented in the first two columns of Table 3 indicate a negative association between labor market tightness and the length of unemployment spell, when no controls are included in the regression. However, after introducing control variables, the coefficient estimate for the generalized inverse tightness measure shifts to positive, which is better in line with the theoretical predictions. In contrast, the coefficient estimate for the conventional tightness measure remains negative even after the inclusion of control variables. These results suggest that the generalized measure performs better at predicting relevant job seeker outcomes.

5 Conclusion

Our analysis using Finnish labor market data from 2006–2021 has contributed to a branch of empirical literature which has attempted to understand the role of barriers in geographical and occupational mobility as determinants of unemployment. Unlike the previous longitudinal analyses of geographical and occupational mismatch, we have applied the interconnected-markets approach of Marinescu and Rathelot (2018) for measuring mismatch, which is based on empirical estimates, rather than arbitrary assumptions, on job seekers' scope of job search.

A puzzling observation made in our analysis, as well as by Alasalmi (2022) is that, while the Beveridge Curve indicates a clear decline in the overall matching efficiency in Finland since the early 2010s, the contribution of geographical and occupational mismatch to the unemployment rate has been modest and has even decreased during the same period. These findings suggest that the increased labor market imbalances are largely attributable to other factors than limitations in job seekers' geographical and occupational mobility and that policies directed at increasing such mobility would, overall, have a very limited impact on unemployment.

The decreasing importance of geographical and occupational mismatch can be somewhat reliably traced to the reduced distances between job seekers and vacancies arising from their increased concentration in the same regions and occupations. However, we do not find evidence that job seekers' willingness to accept distant jobs would have changed substantially during the observation period, and thus, such changes cannot explain much of the reduced importance of distances.

If it is the case, as indicated by our results, that the geographical or occupational misallocation of job seekers and open vacancies does not explain the increased labor market imbalances in Finland, then what does? While providing a comprehensive answer to this question is out of the scope of our study, we can come up with at least three possible explanations. First, it is possible that, within the broad occupational groups, the mismatch between the demand and supply of skills has increased due to technological progress. Second, according to previous studies (e.g. Larja and Peltonen 2023), the increase in the vacancy rate has coincided with an increase in the share of part-time and fixed-term open vacancies. Thus, a reduction in the average quality of open vacancies together with job seekers' strong preferences for stable, full-time employment is likely to be a part of the story. Third, the labor market imbalances might be partly traceable to certain changes in the quality or the composition of the workforce over the past two decades, such as the rapidly increased share of immigrant workers – a group that has suffered from high unemployment.

References

Alasalmi, J. (2022). Selittävätkö alueellinen ja ammatillinen kohtaanto työmarkkinoiden kohtaannon heikentymistä? (in Finnish). *Finnish Labour Review*, 4/2022:22–30.

Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast poisson estimation with high-dimensional fixed effects. *Stata Journal*, 20(1):95–115.

Larja, L. and Peltonen, J. (2023). Työvoiman saatavuus, työvoimapula ja kohtaanto-ongelmat vuonna 2022. *TEM-analyyseja*, 113/2023.

Marinescu, I. and Rathelot, R. (2018). Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics*, 10(3):42—-70.

Patterson, C., Şahin, A., Topa, G., and Violante, G. L. (2016). Working hard in the wrong place: A mismatch-based explanation to the UK productivity puzzle. *European Economic Review*, 84(C):42–56.

Pehkonen, J., Huuskonen, J., and Tornberg, K. (2018). Kohtaanto työmarkkinoilla – havaintoja ja politiikkajohtopäätöksiä (in Finnish). *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja*, 15/2018.

Pizzinelli, C. and Shibata, I. (2023). Has COVID-19 induced labor market mismatch? Evidence from the US and the UK. *Labour Economics*, 81:102329.

Räisänen, H. (2023). Minne uudet työpaikat syntyivät vuonna 2022? *TEM-analyyseja*, 118/2023.

Turrell, A., Speigner, B., Copple, D., Djumalieva, J., and Thurgood, J. (2021). Is the UK's productivity puzzle mostly driven by occupational mismatch? An analysis using big data on job vacancies. *Labour Economics*, 71:102013.

Şahin, A., Song, J., Topa, G., and Violante, G. L. (2014). Mismatch unemployment. *American Economic Review*, 104(11):3529–3564.

Appendix

Table A.1: Distaste for geographical and occupational distances, selected years

Parameter	2014	2017	2020
γ_1	-0.1043***	-0.1007***	-0.1055***
	(0.0022)	(0.0022)	(0.0022)
γ_2	0.0124***	0.0136***	0.0159***
	(0.0030)	(0.0030)	(0.0031)
γ_3	0.0579***	0.0519***	0.0547***
	(0.0016)	(0.0016)	(0.0017)
γ_4	0.0213***	0.0233***	0.0230***
	(0.0008)	(0.0009)	(0.0009)
γ_5	0.0075***	0.0060***	0.0064***
	(0.0005)	(0.0005)	(0.0005)
γ_6	0.0015***	0.0026***	0.0024***
	(0.0003)	(0.0003)	(0.0003)
α_{11}	-1.6678***	-1.8688***	-1.4368***
	(0.0428)	(0.0437)	(0.0382)
α_{12}	-2.1970***	-2.3970***	-1.9948***
	(0.0469)	(0.0467)	(0.0406)
α_2	-0.0496***	-0.0462***	-0.0511***
	(0.0011)	(0.0011)	(0.0010)
N	212,254	234,524	210,032

Notes: Selected Poisson regression estimates of equation (2). The nodes of the spline function are at 20, 50, 100, 200 and 400 km. A one kilometer increase in distance decreases probability of employment by $\exp(\gamma_1)$ for distances under 20 km. For distances between 20 and 50 km, a kilometer increase decreases the probability by $\exp(\gamma_1 + \gamma_2)$ etc. α_{11} and α_{12} are the coefficients for dummy variables indicating 1-digit and 2-digit differences between the occupations of the job seeker and the vacancy, respectively. α_2 is the coefficient for a variable measuring the difference in the educational backgrounds of workers in the occupations of the job seeker and the vacancy. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.2: The largest and smallest skill differences between 2-digit occupations (2021)

Occupation 1	Occupation 2	Skill difference
Commissioned armed forces officers (01)	Health associate professionals (32)	75.1
Commissioned armed forces officers (01)	Personal care workers (53)	74.7
Commissioned armed forces officers (01)	Electrical and electronic trades workers (74)	73.2
Commissioned armed forces officers (01)	Food preparation assistants (94)	71.5
Commissioned armed forces officers (01)	Building and related trades workers, excluding electricians (71)	70.8
Commissioned armed forces officers (01)	Health professionals (22)	70.6
Commissioned armed forces officers (01)	Metal, machinery and related trades workers (72)	70.1
Commissioned armed forces officers (01)	Protective services workers (54)	69.2
Commissioned armed forces officers (01)	Agricultural, forestry and fishery labourers (92)	68.9
Health associate professionals (32)	Electrical and electronic trades workers (74)	68.1
Commissioned armed forces officers (01)	Cleaners and helpers (91)	68.0
Health associate professionals (32)	Personal care workers (53)	67.7
Commissioned armed forces officers (01)	Drivers and mobile plant operators	65.4
Health associate professionals (32)	Building and related trades workers, excluding electricians (71)	65.3
Health associate professionals (32)	Metal, machinery and related trades workers (72)	64.6
...	...	...
Assemblers (82)	Street and related sales and service workers (95)	14.3
Assemblers (82)	Refuse workers and other elementary workers (96)	13.7
Chief executives, senior officials and legislators (11)	Production and specialised services managers (13)	13.5
Chief executives, senior officials and legislators (11)	Administrative and commercial managers (12)	12.9
Food processing, wood working, garment and related trades workers (75)	Stationary plant and machine operators (81)	12.8
Assemblers (82)	Labourers in mining, construction, manufacturing and transport (93)	11.7
Stationary plant and machine operators (81)	Assemblers (82)	10.7
Labourers in mining, construction, manufacturing and transport (93)	Street and related sales and service workers (95)	9.0
Customer services clerks (42)	Numerical and material recording clerks (43)	8.3
General and keyboard clerks (41)	Numerical and material recording clerks (43)	8.3
General and keyboard clerks (41)	Customer services clerks (42)	7.9
Business and administration associate professionals (33)	Numerical and material recording clerks (43)	7.2
Labourers in mining, construction, manufacturing and transport (93)	Refuse workers and other elementary workers (96)	6.7
Administrative and commercial managers (12)	Business and administration professionals (24)	6.3
Street and related sales and service workers (95)	Refuse workers and other elementary workers (96)	4.9

Note: The skill differences between the occupations refer to the differences in workers' educational backgrounds included in equations (2) and (3).

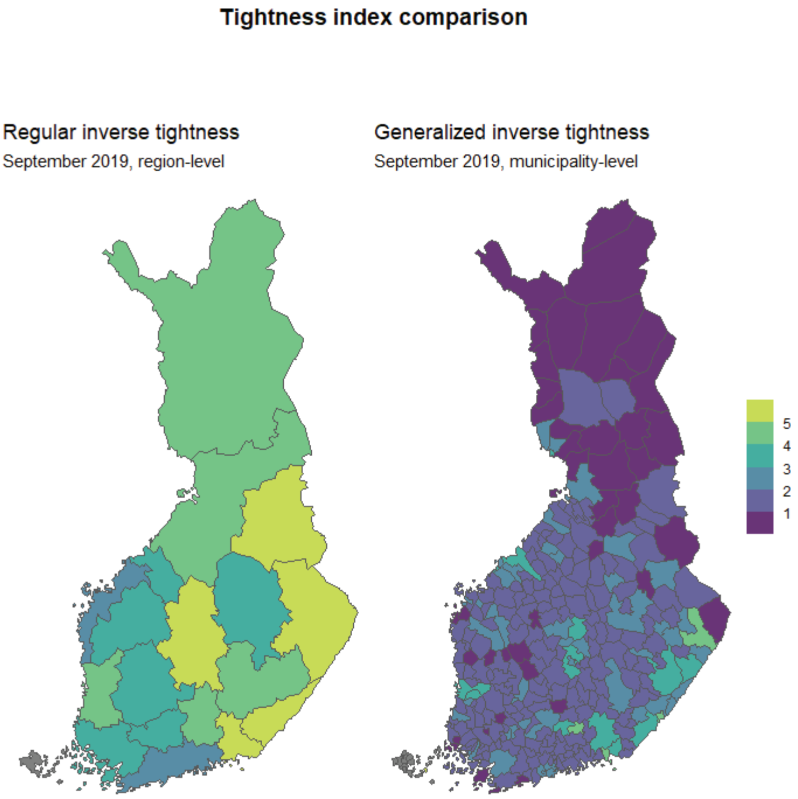

Figure A1: Inverse labor market tightness at the region level: comparison of indices