Growing Inequality:
a Novel Integration of
transformations research

Co-funded by the Horizon 2020 programme
of the European Union

D3.2 Worker resilience to technology shocks

WP3 Digital transformation: impact on skills and inequality

Author:	Emilie Rademakers
Associate Work Package:	WP3
Lead Beneficiary:	UU
WP leader:	University of Utrecht

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 101004494 — The contents of this publication are the sole responsibility of the GI-NI project Consortium and do not necessarily reflect the opinion of the European Union.

Document Summary

Document type:	*Report*
Title:	*Worker resilience to technology shocks*
Author/s:	*Emilie Rademakers*
Reviewer/s:	*Steven Dhondt*
Date:	*February 09, 2024*
Document status:	*Final version*
Keywords:	*Automation, unemployment, occupational reallocation, task overlap, job network*
Version:	*1.0*
Document level:	*Public*
DOI	http://dx.doi.org/10.2139/ssrn.4664997

Summary

We investigate the role of task similarity for the resilience of unemployed job seekers exposed to automation of routine tasks. Using a language model, we establish a novel job-to-job task similarity measure. Exploiting the resulting job network to define job markets flexibly, we find that only the most similar jobs affect job finding. Since automation-exposed jobs overlap with other highly exposed jobs, task-based reallocation provides little relief for affected job seekers. We show that this is not true for more recent software exposure, for which task overlap mitigates the distributional consequences. Our counterfactual simulation highlights the potential harm of increasing job mobility as it strengthens the divided exposure of job seekers to routine-task automation.

Resilience to Automation: The Role of Task Overlap for Job Finding *

Diego Dabed[1], Sabrina Genz[1,2,3], and Emilie Rademakers[1,2]

[1]Utrecht University
[2]TPRI
[3]IAB

Draft version: November, 2023

Abstract

We investigate the role of task similarity for the resilience of unemployed job seekers exposed to automation of routine tasks. Using a language model, we establish a novel job-to-job task similarity measure. Exploiting the resulting job network to define job markets flexibly, we find that only the most similar jobs affect job finding. Since automation-exposed jobs overlap with other highly exposed jobs, task-based reallocation provides little relief for affected job seekers. We show that this is not true for more recent software exposure, for which task overlap mitigates the distributional consequences. Our counterfactual simulation highlights the potential harm of increasing job mobility as it strengthens the divided exposure of job seekers to routine-task automation.

Keywords: automation, unemployment, occupational reallocation, task overlap, job network
JEL: J21, J24, J62, J64, O33

*We thank David Autor, James Bessen, Olivier Charlot, Maarten Goos, Meike Morren, Balasz Reizer, Pascual Restrepo, Anna Salomons, Johannes Schmieder, Marcel Smolka, Michael Stops, Francesco Vona, Wiljan van den Berge, Nicolas Ziebarth, seminar participants of the Technology and Policy Research Initiative, Utrecht School of Economics, and conference participants at the EALE 2023 conference, Pillars conference, GI-NI conference, Paris-Cergy workshop on Job Polarization and IAB TASKS VI conference for valuable comments and suggestions. Diego Dabed, Sabrina Genz and Emilie Rademakers gratefully acknowledge funding from Instituut Gak. Emilie Rademakers gratefully acknowledges funding from the European Union's H2020 research and innovation programme GI-NI under the grant agreement number 101004494 — The contents of this publication are the sole responsibility of the GI-NI project Consortium and do not necessarily reflect the opinion of the European Union.

1 Introduction

Automation technology shifts labor demand across occupation with different task content, away from tasks where automation technology has a comparative advantage (Autor et al., 2003). Uncovering the substitution of machines for workers in routine tasks, an extensive literature analyzes the decline in middle-skill employment in high income countries in the past decades, known as job polarization.[1] When workers have limited ability to adjust to these shifts in labour demand, routine-task automation generates winners and losers, both in the short and the long run.

Joblessness is an important dimension of how workers lose out in the short to medium run. Workers exposed to automation are more likely to be non-employed, and experience declining earnings (Cortes et al., 2017; Bessen et al., 2023). Following job displacement at the firm, exposed workers face stronger earning losses (Braxton and Taska, 2023), which is largely driven by more time spent in non-employment, persisting years after the job loss event (Blien et al., 2021; Goos et al., 2021). This is particularly worrisome given the documented scarring effects of job displacement (Jacobson et al., 1993; Davis and Von Wachter, 2011). Existing literature presents clear evidence of the disruptive impact of routine-task automation. However, it remains an open questions what the origins are of the adjustment costs that are causing this inequality in job loss.

In this paper, we study the role of occupational reallocation as a key factor of adjustment. If reallocation to other jobs with better prospects is very costly or simply unavailable to workers displaced from routine jobs, this can explain the persistence of distributional consequences and provide guidance to policy makers seeking to assist displaced workers. Given that the impact of routine-task automation has a strong occupational component, reallocation across occupations and the potential loss of human capital is crucial to understand how workers can shift away from exposed labor markets.[2]

The central argument that we make throughout this paper is that the existence of worker reallocation does not imply the reduction of distributional consequences from routine-task

[1] From Goos and Manning (2007) coining the term "job polarization", a large set of papers have shown the relevance of this phenomenon using data from different settings. A non-exhaustive list includes: Acemoglu and Autor (2011); Antonczyk et al. (2009); Autor et al. (2006); Autor and Dorn (2013); Autor et al. (2015); Cortes (2016); Dustmann et al. (2009); Goos et al. (2014); Gregory et al. (2022); Michaels et al. (2014) and Spitz-Oener (2006).

[2] A similar argument is made by (Traiberman, 2019) on the role of occupational human capital for the unequal effects of import competition on workers.

automation *per se*. This holds only when mobility connects affected to less affected markets with better job prospects, i.e., when exposure is not too correlated between connected markets. Whether this is true depends on who is exposed and how the connected structure between labor markets allows for reallocation. In other words, it is an empirical question whether workers displaced from routine jobs can shift towards more prosperous occupations. For the same reason, reallocation patterns from automation of routine-task content exposure do not necessarily transfer to more recent technologies such as artificial intelligence (AI).

This paper seeks to understand the unequal job finding of job seekers displaced from routine jobs through the lens of task-based similarity between jobs as a source for reallocation. To that end, we build two constructs from administrative data that are grounded in real job search. We observe the universe of newly unemployed job seekers and their work history in detailed occupations in Flanders, Belgium. Since we use administrative data on unemployed job seekers receiving benefits, these spells consist only of involuntary job loss. First, we apply a state-of-the-art language representation model to exploit detailed task descriptions for jobs maintained by the Public Employment Services (PES). This generates a network of jobs that are connected by similarity in task content. We interpret these connections as overlapping labor markets and representing the potential for reallocation. Second, we quantify the relevance of overlapping labor markets for unemployment duration by augmenting a standard matching function that allows for a flexible definition of the labor market, including overlap with markets that share task content (Goos et al., 2019). These constructs enable us to analyse technology exposure and its relationship to unemployment duration. Crucially, we evaluate the contribution of overlapping markets to this relationship by benchmarking our model to one without task overlap. To focus on the mechanism of job similarity, we employ a standard and proven measure of exposure to routine-task automation: Routine Task Index (RTI) (Acemoglu and Autor, 2011).

Next, we delve deeper into the conditions when overlapping markets can have an equalizing effect on job finding, using additional descriptive statistics and counterfactual simulations. We consider four pieces of evidence: (i) we manipulate the matching efficiency in overlapping markets; (ii) we explore the correlated exposure to RTI within the similarity network; (iii) we introduce additional tasks to exposed job seekers as a form of targeted re-training. Finally, (iv) we analyze a simulated reduction in vacancies along a different dimension to show that, under certain circumstances, overlapping markets can strongly equalize the effect of exposure. Finally, we compare the results from our benchmarking exercise on the exposure to automation of routine tasks with recent innovations in software, robots and AI as measured by current patents (Webb, 2020).

Our analysis yields three main findings. First, our estimation of a flexible matching function shows that only the most similar jobs outside of job seekers' own labor markets are relevant for job finding by the unemployed. Heterogeneity analysis reveals that this is driven by early career job seekers. However, these overlapping markets provide little relief from the link between reduced job finding and exposure to routine-task automation. This finding provides a novel understanding of the disruptive and persistent distributional effects of routine-task automation.

Second, we find that the reason for this lies in the structure of the similarity network. Labor markets of job seekers displaced from routine jobs overlap with other highly exposed markets that do not provide better job finding opportunities. Therefore, increasing the matching efficiency with the most similar markets increases job seekers' exposure to automation in close job markets and reduces their job finding probabilities. Even though we cannot anticipate general equilibrium effects given our reduced form matching function, our counterfactual simulations point to potential harm in recommendation algorithms that direct job seekers to explore job markets that are most similar to their previous job's task content. Furthermore, we stress the importance of correlated exposure by analyzing a simulated reduction in vacancies where exposure is less correlated among overlapping markets. In this case, we find a strong equalizing effect of overlapping markets on the gradient between unemployment duration and exposure of the job profile of job seekers.

Third, we show that the contribution of task overlap on job finding depends critically on the specific technological innovation. It depends on which jobs are most exposed to the technology and their embedding within the job network. For instance, for recent software innovations, overlapping markets provide discernible improvement in the expected job finding rates of exposed job seekers. Further, in contrast to routine-task exposure, current AI exposure predicts faster job finding.

Related literature We contribute to two strands of literature. A first studies the distributional consequences of automation for the labor market. Several papers show that automation-induced worker displacement tends to be task-specific and has shifted labor demand away from routine-intensive occupations (Autor et al., 2003; Autor and Dorn, 2013; Goos et al., 2014; Cortes, 2016; Gregory et al., 2022). Using either occupational exposure or direct measures of automation at the firm, evidence shows that affected workers experience earnings losses, importantly because of more time spent in non-employment (Bessen et al., 2023; Blien et al., 2021; Cortes et al., 2017; Goos et al., 2021).[3] While these papers stress the selective displacement effect of automation as a driver for unequal labor market outcomes, we contribute by providing evidence on the mechanism of reallocation.

Our paper relates in particular to recent work by Acemoglu and Restrepo (2022), who develop a general equilibrium model connecting task displacement to overall changes in wage inequality. Their model highlights the importance of task-based reallocation of workers as a mechanism, concluding that automation is responsible for the declining real wages of low-skilled workers over several decades. Acemoglu and Restrepo (2022) examine employment-to-employment changes over a period of 40 years and argue that task-based reallocation has contributed to the dispersion of negative wage effects from automation. We analyse job finding after unemployment as a relevant and complementary margin of inequality in the short to medium run. Further, we focus on task-based similarity between jobs as a source of reallocation.

Our reasoning on the importance of reallocation for the distributional consequences from automation is also related to recent research on the persistence of migration shocks. Borusyak et al. (2022) show that the estimation of geographic mobility responses to migration shocks is flawed without accounting for the spatial correlation of those shocks. They emphasize that actual mobility responses are underestimated when leaving out "the bilateral nature of location choices" (Borusyak et al., 2022, p. 3). That is, reallocation of workers following a migration shocks depends not only on the conditions of their own local labor markets, but also on the conditions of the labor markets they are able to reallocate to. Similarly, we highlight the importance of accounting for task-based correlation of exposure across occupations when studying the impact of technological change.

Second, we build on a long-standing literature that relates human capital to worker mobility across jobs (Poletaev and Robinson, 2008; Gathmann and Schönberg, 2010; Goos et al., 2019), which highlights that human capital is specific to the tasks workers perform on their jobs. Therefore, it is more costly to switch between jobs the more they differ in task content. Several papers look at observed job-to-job transitions and their associated wage changes and find sizeable costs associated with switching in support of this hypothesis (Cortes and Gallipoli, 2018; Gathmann and Schönberg, 2010; Macaluso, 2022; Poletaev and Robinson, 2008; Robinson, 2018).

[3] Using firm survey data on a combination of technological and organisational change, Battisti et al. (2023) do not find evidence of displacement effects for routine workers. They find that there is within-firm reallocation to more abstract work. This seems in line with the combined organisational change they define as the treatment.

The existing body of research builds mostly on occupational dictionaries to construct measures of task similarity between jobs and relies on the O*NET dictionary for US workers and job-to-job transitions (Dworkin, 2019; Schubert et al., 2021; del Rio-Chanona et al., 2020). Our contribution to this literature is twofold. First, we extend the approach of job markets as a network to explain the distributional consequences of automation for unemployed job seekers. Importantly, we look at job finding rates rather than job-to-job transitions to provide complementary evidence on the extensive margin of employment, as the reallocation patterns of unemployed workers are likely different. Second, we apply a language representation model to construct a novel measure of job-to-job similarity that captures similarity in task content beyond the exact textual overlap in task descriptions used in previous papers. Furthermore, we measure job descriptions at a very granular level that is grounded in the actual search activity of our sample. We are able to exploit the full work history reported to the PES rather than restrict to the last job before displacement.

The remainder of this paper proceeds as follows. Section 2 introduces our data. Section 3 presents our job network based on similarity in task content. Section 4 presents our empirical approach to estimate an extended matching function accounting for task overlap. Section 5 presents our main results on the implications for exposure to routine-task automation. Section 6 discusses possible explanations for our results in Section 5. Section 7 replicates our analysis for measures of more recent technology exposure. Section 8 discusses robustness analysis. Section 9 concludes.

2 Data

Job seeker profiles and task competency information The data for our analysis stem from a real job search environment. Our sample consists of the universe of newly registered unemployed job seekers receiving benefits with the Public Employment Services in Flanders, Belgium.[4] To be registered as "newly" unemployed receiving benefits, persons cannot have been unemployed in the past six months and lost their jobs involuntarily. Importantly, this means that recurring job seekers in between short spells, e.g., at a temp agency, are not captured. Because of the generous system of unemployment benefits in Belgium, labour market entrant looking for work can also apply for benefits and are part of our sample. Unemployed job seekers older than 55 are not required to search actively, and are therefore not included in our sample. Between

[4]Flanders is the Dutch-speaking region of Belgium, with 6.7 million inhabitants. Due to language differences, the PES is organized at the regional level. At the start of 2021, the unemployment rate in Flanders was 4.3%.

March 1, and September 9, 2021, we observe 33,352 individuals entering unemployment. Our main data are taken from the PES' online platform called "Mijn Loopbaan" (translated as "My Career"). It is mandatory to register on this platform and create a profile when a job seeker claims unemployment benefits.[5] The obligation to use the platform safeguards our analysis against common problems in studies using web platforms like voluntary selection into the online platform, user anonymity or sample attrition (Altmann et al., 2022). From the search profiles on the platform, we obtain the list of occupations that job seekers are interested to work in and their experience in these jobs. Therefore, the list does not contain jobs in which job seekers have experience, but no longer want to work in. We consider this to correctly identify the market for job search.[6] Since the data are based on job seekers' real CVs, the list may contain more than one job. Job seekers list three different jobs on average.[7] This sets us apart from previous literature on task-based mobility, which typically only looks at the last known occupation (Gathmann and Schönberg, 2010; Cortes and Gallipoli, 2018). In total, we observe 1,435 distinct jobs, according to a very detailed classification: 620 occupations which each up to four levels of experience: starter, less than two, between two and five, and more than five years of experience. Throughout this paper, we refer to a combination of an occupation and experience level as a job. Each of these jobs is linked to detailed task descriptions of the activities performed in the respective job. This task matrix was built on the classification by the French PES (ROME-v3) combined with information from expert surveys in Belgium.[8] Job seekers also see their task competencies when completing the profile.

Figure 1 shows the share of job seeker search profiles in our total sample by one-digit ISCO-08, indicated by the blue squares. Due to the universal coverage of our data, a large variety of occupations are included in job profiles. Our sample contains both higher skilled, and often non-routine, professional jobs, as well as lower skilled, more routine profiles such as clerical or

[5] While anyone can register on the platform, we focus on the unemployed given the universal coverage.

[6] We conducted a complementary survey with the job seekers in our sample. This survey was distributed via mail and sent five weeks after the start of unemployment. The response rate was around 8% on average. We asked job seekers to compare their use of the PES platform to other channels, like other job boards, social media and temp agencies. On average, the PES platform was used most often among all channels. It was also ranked highest in terms of preferences over where to search. Finally, for respondents who had already found a job, it was the second most important channel through which they found work, after temp agencies. The stated importance of the PES online platform in this survey supports the quality of the search profile data.

[7] We do not know the relative importance of the jobs listed in the profile. Therefore, we may be including information on previous job interests that are no longer relevant. Several search profiles have listed jobs dating from previous periods of unemployment. Given that profiles are mandatory and discussed with a counselor, we trust the information provided in the online profile. As a robustness, we drop jobs that were added to the profile more than 10 years ago. This does not change our main results.

[8] See Appendix A.1 for further discussion and examples.

elementary occupations.

Measures of technology exposure For our main analysis, we employ a commonly used measure for job seekers' exposure to routine-task automation: occupational routine-task intensity (RTI) based on O*NET 2009 (Acemoglu and Autor, 2011). The measure describes to what extent work activities can be fully described by a set of rules and procedures (i.e., well-defined codifiable steps) and carried out by machines (i.e., computers and other electronics). This measure is well-established in the literature and empirical evidence confirms its predictive power for computer adoption and subsequent worker outcomes (Autor et al., 2003; Autor and Dorn, 2013). We compare the automation of routine tasks to more recent waves of technological innovations. Specifically, we apply recently developed measures that capture the exposure of occupations to current patents for software, robots and AI (Webb, 2020).[9] Each measure is cross-walked from the SOC to ISCO-08 and merged at the four-digit ISCO-08 level. To calculate the respective technology exposure for each job seeker, we average each of the four measures across all jobs in the job seeker's profile.

Table 1 shows the mean RTI score by one-digit ISCO-08 occupation of the jobs in the job seeker profile. In line with previous literature, clerical support jobs, jobs related to plant and machine operating and elementary jobs are most exposed to routine-task automation.[10]

Administrative records We complement the main data with administrative records maintained by the PES. For each job seeker, we observe a list of person characteristics that we use as controls: age, gender, education level, a dummy for being a labor market entrant, migration background, location as NUTS2, urbanization of the place of residence, and Dutch proficiency. Summary statistics for our job seeker sample can be found in Table 2. Our sample is balanced in terms of men and women. Around 30% is an entrant to the labour market. Subsequently, the mean age is 30. We observe a large variety in educational background, with both high school, some college and university educated. Finally, 20% of our sample does not have Belgian nationality; and 10% reports having limited or no knowledge of Dutch. Importantly, we also

[9] See Appendix A.2 for more details on the construction of the technology exposure measures and their relation to each other.

[10] We find a higher RTI score of elementary occupations than others using RTI at the ISCO-08 level (Goos et al., 2014). This does not arise from the applied mapping from the SOC to ISCO-08 occupational codes, but originates from the specific work experience within one-digit ISCO-08. Job seekers in our sample with experience in elementary occupations tend to have experience in the particularly routine-intense jobs, such as occupations within two-digit ISCO-08 code 93: "Labourers in Mining, Construction, Manufacturing and Transport".

observe at a monthly frequency whether job seekers have found employment over a maximum of 18 months.[11]

Finally, we have information on the total stock of unemployed job seekers and vacancies in every month from the PES' online platform. Given the mandatory online registration, we capture the universe of unemployed job seekers.[12] Since vacancies are posted on the platform via two channels, it has a wide coverage. First, firms can post vacancies directly on the platform. Second, the PES has exchange agreements with major temporary agencies and other job boards. Importantly, vacancies are classified according to the same classification as used for jobs in the job seekers' profiles regardless of through which platform the vacancies were obtained.

Figure 1 plots the average tightness defined as the number of vacancies over the number of job seekers in a job market across all time periods and locations by one-digit ISCO-08 as indicated by the red circles. There is substantial variation in the tightness across jobs. Unsurprisingly, the market for professionals, in particular medical professionals, was very tight during the COVID-19 pandemic. However, we also see that a large share of our job seekers search for jobs in markets that have few vacancies per job seeker, such as the market for clerical support workers.

3 A network of jobs connected by task similarity

In this section, we build a network of jobs connected by task similarity. The network presents the availability of alternative, similar jobs for every job present in the search profiles of our sample. This is the first, and central building block to the analysis of this paper. The existence of such similarity across jobs creates an overlapping market structure which we use to understand the distributional consequences from routine-task automation.

Language model to quantify similarity in task content On average, every job in our data contains 13 tasks and every task occurs in two jobs in our task matrix. The occurrence of tasks across jobs is highly skewed: many tasks are specific only to one or two jobs and few tasks occur across many jobs. Previous studies have relied on the re-occurrence of exact task descriptions, e.g., the importance of activities as reported in the US Dictionary of Occupational Titles (DOT) or O*NET (Cortes and Gallipoli, 2018; Dworkin, 2019). Due to the skewed

[11] Around 15.3% of our sample is right-censored. We apply a duration model to handle these cases.

[12] The stock of job seekers comprises all jobs listed on the online profiles since job seekers may be searching in multiple job markets.

distribution of task occurrence, we would likely underestimate the overall similarity across jobs with an approach based on exact word matches.

To deal with the skewed distribution of task occurrence, we build a similarity measure that does not depend on the exact wording of the description. Rather, we distill the complex information recorded in the detailed task descriptions by applying the pre-trained language representation model MPNet (Song et al., 2020) fine-tuned for sentence similarity.[13] This self-supervised algorithm is trained to interpret sentences based on a large-scale text corpus by translating words into numerical vectors. Specifically, the algorithm first retrieves the meaning of words by looking at their surrounding words and their position in the sentence. Second, the model is fine-tuned to update the obtained vectors by focusing on the similarity computation of sentences. In our context, the obtained numerical vectors contain the semantic information of each task description. Therefore, task similarity comes from task descriptions containing the exact same words but also from task descriptions that describe similar activities. Task vectors are averaged up to the job level.[14] Job vectors are then compared using the standard measure employed by the literature on large language models: cosine similarity. We use cosine similarity throughout as our measure of task similarity between jobs.[15]

To bolster the validity of our measure, we compare it to two, more aggregate measures from the literature. First, we cross-walk the measure of occupational similarity between 37 occupation categories from the DOT by Cortes and Gallipoli (2018) to our job classification. Second, we use panel survey data from the EU-SILC project on Flemish adults between 2013 and 2020 to compute employment transition probabilities. However, this comparison can only be made at the two-digit ISCO-08 level due to sparse observations of employment transitions at lower levels.[16] Compared to these two measures, our job measure is much more granular: we can compute pairwise similarities for each of the 1,435 distinct jobs that are registered in search profiles. Therefore, we cannot directly compare the three measures at the level of detail that we observe in job search. After aggregating our job similarity measure to the two-digit ISCO-08 level, we find similar patterns in job similarity across all three. The Spearman rank correlation

[13]The code and more details on the fine-tuning are available at `https://huggingface.co/sentence-transformers/all-mpnet-base-v2`. MPNET builds on the more widely known BERT language model. Using the fine-tuning of BERT for sentences, SBERT, produces similar results.

[14]Using different types of aggregation lead to similar results.

[15]A visual representation and further discussion of our approach with MPNET to quantify similarity in task content can be found in Appendix A.3.

[16]Belgian administrative data does not contain information on occupation of employment and therefore does not provide us with an alternative way of estimating transition probabilities.

with our similarity is 0.476 and 0.505 respectively ($P<0.01$). This comparison highlights the strength of our approach to capture more nuances in the similarity between jobs, reflecting the jobs outside of their own search profile at which job seekers are able to direct their job search effort.

Task-based similarity between jobs Figure 2 displays a heatmap with all pairwise job similarities based on our constructed measure. We interpret this matrix as a network of jobs connected by task similarity. Since cosine similarity is symmetric when comparing a job pair, we only show the area below the diagonal. While we do not interpret the level of cosine similarity, we exploit the ordinal ranking of the entire similarity distribution to group overlapping jobs into four bins ranging from the 5% most similar jobs to the 35% least similar jobs. Darkly shaded areas represent highly similar job pairs in terms of semantic meaning; lightly shaded areas represent job pairs that have low similarity. Jobs are ordered by ISCO-08 on each axis, such that jobs close to the diagonal share several digits in their ISCO-08 codes. We provide labels along the axes and the diagonal at the one-digit ISCO level.

The dark areas close to the diagonal show that jobs with similar ISCO-08 codes are likely to have similar task content. However, we also find a lot of off-diagonal darkly shaded areas. There are several examples of shared task content between jobs belonging to the group of associate professionals (3) such as engineering technicians and many other technical jobs (7-8), represented by the dark vertical lines at the beginning of ISCO-08 code 3. There are also examples of jobs that have little to no overlap with other jobs, as indicated by the yellow vertical and horizontal lines. In the group of professionals (2) these are mostly entertainment jobs like singers or dancers that conduct specific tasks like vocal exercises or rehearsing of dance steps.

Broadly, two darkly shaded, large triangle areas at the top left and bottom right indicate a divide between jobs in ISCO-08 codes one to five and between ISCO-08 codes six to nine. This aggregate divide has been found in previous studies on job similarity, regardless of their methodology (Alabdulkareem et al., 2018; Cortes and Gallipoli, 2018; Dawson et al., 2021; Schubert et al., 2021) and reflects the boundary between cognitive and manual work, and college and non-college educated jobs.

In addition to these aggregate patterns, our approach detects variation in the level of job similarity even within narrowly defined ISCO-08 occupations. We leverage this granular variation, which matches the level of detail found in job seeker profiles, in order to predict job finding.

4 An extended matching function with task overlap

4.1 Empirical framework

To understand the capacity for task similarity among jobs to reduce the distributional impact of routine-task automation, we must first relate job finding to the conditions in your own and similar job markets. This allows us to uncover which ties within the job similarity network are actually relevant for job finding. For this purpose, we apply an extended version of a matching framework. Our empirical framework follows the work of Manning and Petrongolo (2017) and Goos et al. (2019), which we explain briefly. Further discussion on how the parameters link to the structural model of Manning and Petrongolo (2017) can be found in Appendix A.4.

Starting from a standard matching framework (Pissarides, 2000) with a Cobb-Douglas matching function, we expand the definition of a job market by differentiating between the tightness in the jobs in job seekers' own markets (i.e., jobs listed on their job profile in which they have experience and thus possess the corresponding set of task competencies) and the tightness in jobs that are to some extent similar (i.e., jobs in which they have no experience but which are partially similar in task competencies). Similar jobs are binned by percentiles of the cosine similarity distribution. The four bins are cut at the 95th, 75th and 35th percentile of the entire distribution of similarities, identical to the variation shown in Figure 2. These bins represent the tightness in "close" and "distant" neighbors in a network where jobs are connected through task similarity. We estimate a coefficient on each of these separate bins such that the relative search efficiency of similar jobs is estimated as a step function.[17] The log probability of finding a job for any job seeker can then be expressed as:

$$ln(T_i) = \alpha_1\, ln(V_i^{full} + \sum_{j=1}^{4} \gamma_j V_{ij}^{partial}) + \alpha_2\, ln(U_i^{full} + \sum_{j=1}^{4} \beta_j U_{ij}^{partial}) \tag{1}$$

with i = job seeker; j = the grouping of jobs by similarity; V_i^{full} = number of vacancies in own job markets; $V_{ij}^{partial}$ = number of vacancies in jobs with partial task overlap; U_i^{full} = number of other job seekers in own job markets; and $U_{ij}^{partial}$ = number of other job seekers in jobs with partial task overlap. As in the standard model, parameters α_1 and α_2 capture the positive and negative externalities from job search respectively. The parameters γ_j and β_j capture the relative matching efficiency of vacancies and job seekers in jobs within the similarity of bin j.

[17] We apply a range of sensitivity analysis to the choice of bin delineation in Section 8.

This efficiency should be interpreted relative to the vacancies and job seekers in the jobs for which the job seekers possess the full set of task competencies. For example, if $\gamma_j = 0.5$ this means that a vacancy in bin j has only half of the positive spillover on job finding as a vacancy in a job that is part of the job seeker's profile.

The focus of this paper is to understand to what extent overlap between job markets can equalize the exposure to automation. Therefore, we differentiate equation (1) by exposure to automation in the jobs for which the job seeker has experience. The channel through which this affects equation (1) is through lowering the arrival of vacancies in those occupations. For this discussion, it would be equivalent to assume that there are more unemployed job seekers in automation exposed occupations, or a combination of both. Instead, we assume,

$$\frac{\partial V_o}{\partial a_o} < 0 \tag{2}$$

where a refers to the automation intensity of the job market o.[18] Because an individual job seeker may have gathered experience over multiple occupations, we define a_i to be the exposure to automation across all jobs in which individual i has experience. Recent research has shown the importance of using full occupational histories for wage inequality (Vafa et al., 2023). Our analysis provides a potential mechanism for this importance. By looking beyond the occupation where the job seeker is displaced from, we consider to what extent workers become "entrenched" into a narrow set of jobs with similar task content. This provides them with limited capacity to adjust to routine-task automation.

For ease of exposition, we discuss three distinct scenarios of overlapping markets of our equation (1) to characterize the expression for the marginal effect of automation exposure on job finding:

$$\frac{\partial ln(T_i)}{\partial a_i} \tag{3}$$

First, consider a world in which jobs are completely insular, such that workers cannot find employment in jobs besides those where they have previous experience: $\gamma_j = \beta_j = 0 \quad \forall\, j = 1/4$. In this case, the associated change in job finding would fully depend on the change in vacancies in job seeker's own market. Because of the log specification, an increase in a_i is associated with

[18] In our data, a standard deviation increase in RTI of a job profile is associated with a 0.08 reduction in tightness, which is highly significant and 40% of mean tightness.

a loss in job finding due to the lowering of vacancies, proportional to the level of vacancies:

$$\frac{\partial ln(T_i)}{\partial a_i} = \frac{\alpha_1 \left[\partial V_i^{full}/\partial a_i\right]}{V_i^{full}} \tag{4}$$

Second, we can consider a world in which job markets are perfectly integrated, that is finding work in any type of job is possible without switching costs: $\gamma_j = \beta_j = 1 \quad \forall j = 1/4$. Then, the change in job finding no longer depends on i since every job seeker can equally search across all markets:

$$\frac{\partial ln(T)}{\partial a} = \frac{\alpha_1 \left[\partial V/\partial a\right]}{V} \tag{5}$$

where $V = V_i^{full} + \sum_{j=1}^{4} V_{ij}^{partial}$. This sum is the same for every individual since the total set of bins captures all existing job markets. Similarly, exposure to automation a would be the aggregate mean across all jobs.

Third, we return to a more realistic scenario in between the first two extreme scenarios, where some overlap exists. We leave out the superscripts $full$ and $partial$ to avoid cluttering.

$$\frac{\partial ln(T_i)}{\partial a_i} = \frac{\alpha_1 \left[\partial V_i/\partial a_i + \sum_{j=1}^{4} \gamma_j \left[\partial V_{ij}/\partial a_i\right]\right]}{V_i + \sum_{j=1}^{4} \gamma_j V_{ij}} \tag{6}$$

The expression is now a compound of the exposure in job seeker's own market and in overlapping markets, weighted by their relative importance γ_j. This second element is non-zero if the exposure in job seeker's own profile is somehow determinant of exposure to vacancies in the overlapping markets. More specifically, we consider:

$$\frac{\partial V_{ij}}{\partial a_i} = \frac{\partial V_{ij}}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial a_i} \tag{7}$$

We define a_{ij} to be the share of automation-exposed occupations in bin j for individual i. This expression captures the fact that exposure in job seeker's own market and markets in bin j may be correlated. That is $corr(a_i, a_{ij}) > 0$ such that $\frac{\partial a_{ij}}{\partial a_i} > 0$. In that case, the left-hand side of equation (7) is negative and the negative gradient of job finding on automation exposure becomes steeper. Consequently, overlapping markets prolong the time job seekers with more automation-exposed job profiles need to search for work.

Taken together, equations (6) and (7) show that inequality in job finding across job seekers with differential exposure to automation come from lower job prospects accumulated across own and overlapping markets. Therefore, the time job seekers need to search for work when they have more automation-exposed experience can be reduced by overlapping markets only if the correlation of exposure is negative between overlapping markets $\text{corr}(a_i, a_{ij}) < 0$. As we argued in the example above, the time spent in unemployment can also be reinforced by a positive correlation, or remain unaffected if the correlation is small. This effect is mediated by the relative efficiency of searching in overlapping markets: γ_j. A large γ_j implies that vacancies in overlapping markets j have similar effects on job finding as vacancies in job seeker's own market. In addition, the total change in job finding is larger the more positive spillovers are generated by vacancies: a large α_1 implies low overall search frictions. Whether exposure correlates across overlapping markets depends on the distribution of exposure and the structure of the job network, and motivates our empirical analysis.

Arguably, our parameters γ_j and β_j capture several, more structural parameters at the same time. The step function as shown in equation (1) can be seen as estimating in a reduced form setting the first "ripple effect" of the "propagation matrix" developed by Acemoglu and Restrepo (2022). Therefore, it captures the observed search patterns where there are barriers to search outside of job seekers' own market, causing $\gamma_j \wedge \beta_j < 1$. This can be driven by many elements. First, we consider these parameters of relative efficiency to capture the cost of task-based mobility and taste preferences that workers may have for the jobs in which they have experience. Although concerning geographic rather than task-based mobility, these elements mimic the modelling of Borusyak et al. (2022) for the location choice of workers. In their model, structural parameters on taste for a specific location and cost of mobility would decrease the share of outward mobility from your origin market, which is how we interpret small levels of γ_j and β_j. Second, in a setting of job search, parameters γ_j and β_j may also capture hiring practices by employers. Even if job seekers would have a low cost of moving across task-space, this may be limited by the willingness of employers to hire a job seeker who doesn't have experience in the exact job that is vacant. Finally, it may also reflect strong bundling of tasks in the production function and high costs of training. If the specific combination of tasks on the job contains significant complementary in the spirit of Lazear (2009), then hiring someone without experience in this specific combination of tasks required to perform the vacant job may be very costly as it requires essential training. For the purpose of this paper, the actually source of limited overlap between job markets is less important. Rather, we focus on how these barriers lead to unequal impact from routine-task automation for job finding.

4.2 Estimation approach

Because equation (1) is non-linear in parameters γ_j and β_j, we apply log-linearization as proposed by Manning and Petrongolo (2017) before estimating the duration analysis.[19] We use a log-logistic functional form to allow for non-monotonicity in duration dependence.[20] The stocks of vacancies and job seekers are measured at the start of the unemployment duration and are region specific at the NUTS2 level. We include person controls such as age, gender, detailed educational attainment, migration background, Dutch proficiency, urbanization of the place of residence, whether the job seeker is a labor market entrant, and region- and month of inflow-fixed effects. The coefficients presented in Table 3 are obtained from this estimation.[21]

5 The contribution of overlapping markets to unequal job finding

In this section, we first present the baseline estimated coefficients from equation (1), which informs us on the existence of overlapping markets. Next, we document to what extent these overlapping markets contribute to the inequality in job finding for job seekers exposed to routine-task automation. In subsection 5.2, we compare the predicted unemployment duration for job seekers with varying degrees of exposure to routine-task automation and benchmark the contribution of overlapping markets against a model without task overlap. The final subsection 5.3 performs a heterogeneity analysis by age groups.

5.1 Predicted job finding when accounting for overlapping job markets

Table 3 contains the estimated coefficients from equation (1). It confirms the standard assumption of externalities in job search: a doubling in vacancy stocks increases the job finding rate by 8.4%. The opposite holds for job seekers, albeit by a smaller magnitude (3.9%). We find a significant relationship for the stocks of vacancies and job seekers in overlapping jobs, but, strikingly, only

[19] To ensure that our results are not driven by the log-linearization, we also estimate a linear version of equation (1) as follows: $ln(T_i) = \alpha_1\, ln(V_i^{full}) + \sum_{j=1}^{4} \nu_j ln(V_{ij}^{partial}) + \alpha_2\, ln(U_i^{full}) + \sum_{j=1}^{4} \mu_j ln(U_{ij}^{partial})$. ν_j and μ_j then capture the compound effect of α_1 and γ_j, and α_2 and β_j respectively. We estimate this equation by putting in each bin separately and all bins combined. Although misspecified, the results are qualitatively similar.

[20] Our results are similar when using the Weibull proportional hazard functional form.

[21] See Appendix A.5 for more details. Note that the identifying variation for estimating our coefficients comes from the tightness in jobs and their neighbours in the network which vary at the level of job-month-provinces. Since the majority of the job seekers (roughly 60%) list multiple jobs in their search profile, clustering at the level of the first job listed combined with the month in which the job seeker entered unemployment and province information, would not be precise for those who list multiple jobs and too restrictive. However, clustering at the job-month-province level does not change the statistical significance of the estimated coefficients.

for the most similar jobs in the first bin. For these close neighbors, the relative efficiency is considerably lower at around 28%. Put differently, three to four more vacancies in most similar jobs have the same positive association with job finding as one more vacancy in a job in which the job seeker has experience. As mentioned in section 4.1, a variety of reasons might contribute to these significant search frictions among similar jobs such as a lack of information, lock-in effects, and hiring practices related to job experience, among others. The corresponding coefficient on job seekers has the same magnitude although we lack power to estimate it precisely.[22] This is our first finding: job markets overlap based on task similarity. However, it is only the most similar jobs, i.e., the closest neighbors in the network, that matter for job finding.

This result provides a novel understanding to previous findings that task-based mobility costs appear sizeable (Gathmann and Schönberg, 2010; Cortes and Gallipoli, 2018; Macaluso, 2022). First, we take into account that even equally similar jobs may have unequal job prospects in terms of tightness. Second, we find sizeable search frictions even among jobs that are very similar in terms of task content when accounting for semantic meaning. Interestingly, our estimates for the relative efficiency of overlapping job markets are very similar to Goos et al. (2019) even though our overlapping markets contain similarity based on semantic meaning in addition to exact textual overlap only. This suggests we capture relevant additional information.

5.2 Implications for exposure to routine-task automation

Next, we test how task-based mobility helps account for the distributional consequences for job seekers exposed to automation. Do overlapping labor markets provide an alternative for highly exposed job seekers trying to find work? Figure 3 plots the expected unemployment duration for job seekers ranked according to their automation exposure captured by the average routine-task intensity (RTI) in their job profile. First, we plot the expected unemployment duration from our extended model accounting for task overlap as discussed in equation (6), after using the estimated coefficients as presented in Table 3 ("with overlap"). Consistent with the routine replacing hypothesis (Gregory et al., 2022), job seekers with routine intensive profiles are more likely to remain unemployed longer: a standard deviation increase in RTI implies approximately five days longer spent in predicted unemployment on average. This corresponds to around 5%

[22]This is driven, in part, by the condition on additional control variables that absorb part of the variation in tightness in similar markets. In particular, we control for one-digit ISCED educational attainment, which contains six categories. Suspending education from the set of control variables yields a similarly sized coefficient, but with smaller standard errors.

of the median expected duration. Second, we benchmark our result against a restricted model which includes only jobs listed in the job seeker's search profile, i.e., not taking task overlap into account ("without overlap") as presented in equation (4). Importantly, we consider whether the gradient, the difference between routine and non-routine job seekers in unemployment duration, differs when allowing for task overlap. However, we find no change in the slope when contrasting the extended with the restricted model. It appears that the tightness in the neighboring markets for routine jobs is not sufficiently different to offer an escape hatch. In terms of our equation (7) this would imply that our term $\sum_{j=1}^{4} \gamma_j \left[\partial V_{ij}/\partial a_i\right] \approx 0$. First, we know that only the most similar bin of jobs is relevant: $\hat{\gamma}_j \approx 0$, for $j = 2, 3, 4$. Therefore, this result implies $\gamma_1 \left[\partial V_{i1}/\partial a_1\right] \approx 0$. From our estimation, we know $(\gamma_1 \,\&\, \beta_1) \neq 0$, but small in magnitude. This means that there is a small, and at least non-negative correlation of exposure to RTI in job seeker's own and most similar job market: $\frac{\partial a_{i1}}{\partial a_i} \nless 0$.

This is the second finding of our paper: the observed longer unemployment duration of job seekers with routine intensive job profiles can be understood because only very similar markets are relevant for job finding at a low efficiency ($\gamma_1 \,\&\, \beta_1 \ll 1$) and these do not provide automation-exposed job seekers with better prospects. This mechanism provides a new lens to understand the disruptive and persistent distributional consequences of routine replacing technology for job finding.[23]

This conclusion is complementary to that of Acemoglu and Restrepo (2022) who find a sizeable reduction in wage inequality because of worker reallocation. Differences might be due their focus on job-to-job transitions among the working population in the US across the last four decades. We analyze newly unemployed job seekers since this particular group provides complementary evidence on the extensive margin of employment in the short-run. Furthermore, we concentrate on a fine-grained task-based similarity between jobs as a key source of reallocation.

5.3 Heterogeneity analysis by age

There are two main reasons why our results are likely heterogeneous by age of the job seeker. First, as mentioned by Autor and Dorn (2009), older workers are over-represented in jobs exposed to routine-task automation. Young workers entering the labor market are more likely to enter growing occupations instead. Second, older workers are often less mobile for two reasons related

[23] Another possible interpretation of our result is that even after allowing for overlapping job markets, automation exposure still creates a mismatch between labor supply and demand across different jobs. Despite different approaches, this conclusion is comparable to del Rio-Chanona et al. (2020).

to human capital theory: i) they have accumulated more investment in their current job; ii) they have fewer years to enjoy the benefit of switching to a different job.

In Table 4 we split our main analysis into two age groups: older and younger than 35.[24] We find significant results for γ_1 and β_1 for the sample of those younger than 35, implying that overlapping markets matter for this group of job seekers. Surprisingly, also for early career workers overlapping job markets remain limited to jobs with high degree of overlap. For those older than 35, we cannot reject the hypothesis of perfectly insular occupational markets. The point estimates in column (2) are not within the expected range of $[0; 1]$ or sign, but are highly insignificant. Our results suggest that we can expect job seekers older than 35 to experience more concentrated effects from exposure to automation. This is in line with heterogeneity analysis made in Bessen et al. (2023), who find that more time spent out of employment after an automation event is mostly driven by workers older than 50.

In addition, we check whether this heterogeneity changes the implications for exposure to routine-task automation and find this is not the case. The coefficients in Table 4 imply that the reallocation for later career job seekers is too small to be detected. Therefore, tightness in very similar occupations does not change the gradient between job finding and RTI of the job profile. When we check this separately for early career job seekers, we still find that the observed overlap with similar job markets leaves the gradient unchanged. Even though we observe overlapping markets for early career job seekers, these do not provide better job finding opportunities for those exposed to routine-task automation.

Taken together, the heterogeneity analysis suggests that the inequality in job finding from routine-task automation is pervasive across the age distribution. Overlapping markets do not provide any relief, not even for early career workers.

6 Why do overlapping markets not provide equalizing effects?

Our main result in Figure 3 is that task-based mobility does not mitigate the distributional consequences of routine-task automation exposure by itself. In this section, we analyze why this is the case through discussion of the network structure and three counterfactual simulations that

[24]Note from the number of observations in column (1) and (2) that we have a considerable share of early career job seekers. There are two main reasons for this. First, because different rules on receiving unemployment benefits apply once persons are older than 55, our sample is limited to those younger than 55. Second, in Belgium, labor market entrants are eligible for unemployment benefits even though they have no working experience. Since our observation period covers summer months, the outflow from secondary and tertiary education is also captured in our sample.

manipulate i) the relative matching efficiency; ii) the task competencies of routine job seekers as measured by the top 33% of the RTI distribution (Acemoglu and Restrepo, 2022); and iii) the number of vacancies along a different dimension: job titles that are specific to manufacturing. Each of these provides further insight into when overlapping markets *can* or *cannot* alleviate distributional consequences from exposure to declining demand and how policy could weigh on this. While we cannot anticipate further general equilibrium effect of these interventions here, incorporating dynamic adjustments is an important direction for future research.

There are generally two possible reasons why overlapping markets do not provide equalizing effects. First, the overlap with other markets in the job network is not strong enough, i.e., the relative matching efficiency is too small. Second, the network structure itself is not connecting affected job markets to less affected ones. We analyze each of these in turn and discuss their relevance for policy making.

6.1 Is the relative matching efficiency of similar markets too low?

Table 3 demonstrates that close neighbors are only predictive for job finding to a limited extent: the relative efficiency of additional vacancies and job seekers in similar jobs is small, around a third. The low relative matching efficiency potentially reflects job switching costs posing a barrier for job seekers to access similar overlapping markets (Cortes and Gallipoli, 2018). To understand the importance of low relative efficiency for our results, we simulate an extreme scenario without any barriers to access close neighbors. Specifically, we set the parameters γ_1 and β_1 in equation (1) to one. That is, we allow the tightness in most similar jobs to have the same influence as the tightness in job seekers' own jobs. This can be interpreted as a reduction in information friction by giving job seekers better guidance on job search outside of their initial occupation. Alternatively, it can be a result of a policy that targets hiring practices to be more open about hiring candidates with different occupational backgrounds or certification of person-specific instead of job-specific task competencies. Finally, one may consider this as providing training to job seekers to access a broader set of vacancies. As a result, especially highly routine jobs experience an increase in the predicted unemployment duration and the gradient between duration and RTI exposure increases. This simulation worsens the prospects of routine job seekers since they get more exposed to the correlated automation exposure and similar tightness conditions among overlapping jobs. This implies that simply improving reallocation to job markets with similar task content, for example through a recommender algorithm on a job platform, will not help job seekers exposed to routine-task automation.

6.2 Does the network structure determine the equalizing effects of overlapping markets?

To understand the importance of the job structure for the equalizing effects of overlapping markets, we provide three pieces of evidence. First, we explore a key element from our discussion of equation (7): to what extent does the exposure to routine-task automation in your own market correlate with exposure in similar markets: $corr(a_i, a_{ij})$. We can measure this directly in our data. Second, we ask whether routine job seeker profiles would benefit from containing more non-routine task competencies and discuss what kind of policies could support this. Finally, we simulate an industry-specific shock to highlight the specific occupational nature of routine-task automation.

The correlation of routine-task automation exposure between own and most similar job markets To understand what drives the irrelevance of overlapping markets for job seekers exposed to routine-task automation, Table 5 explores the structure of the job network that determines the correlation component of equation (7). To differentiate routine and non-routine job profiles, we follow the convention to classify job profiles in the top 33% of the RTI distribution as routine (Acemoglu and Restrepo, 2022).

The striking difference between routine and non-routine job search profiles is their set of close neighbors, as described by the top panel of Table 5. While job seekers with a routine search profile have slightly more close neighbors in the first most similar bin compared to non-routine job seekers, routine job seekers have a much higher share of other routine jobs among their close neighbors compared to non-routine job seekers, around 49% relative to 21%. As a result, routine jobs are segregated from non-routine jobs in terms of close neighbors in the job network. This difference in the embedding in the network helps explain why overlapping job markets do not make routine job seekers more resilient: their accessible overlapping jobs are highly exposed to automation, too. These close neighbors provide them with fewer job opportunities compared to non-routine neighbors: the average tightness of close neighboring jobs is lower for routine profiles. Interestingly, we do not find stark differences between routine and non-routine job profiles along other characteristics of the jobs listed in the search profiles such as the number of jobs listed in the profile or the average number of tasks within a job. We find that routine profiles list jobs with slightly lower diversity of tasks within a job. This is measured by comparing the task embeddings within a job using cosine similarity. Lower diversity of tasks might contribute to the segregation as lower diversity in the task content may lead to smaller variation in the content of close neighbors. As expected the tightness in the jobs listed in the search profiles is lower for routine job profiles than for non-routine job profiles.

A plausible explanation for the segregation of jobs in the network is a segregation in how tasks are bundled within jobs. Routine jobs do not overlap with non-routine jobs, most likely because they do not contain enough non-routine tasks. And vice versa for non-routine jobs. The unstructured nature of the task descriptions in our data impedes such an analysis at the task level since we cannot specify routineness at the task level. Consistently, Alabdulkareem et al. (2018) document such segregation using O*NET, referring to it as skill polarization. The contribution of this paper is to show how this relates to inequality in job finding.

Providing routine job seeker profiles with non-routine task competencies One plausible approach to resolve the segregation between routine and non-routine jobs may be to alter the task competencies of routine job seekers by providing them with the most common non-routine tasks. To simulate the implications of this intervention, we add the 15 most common tasks among jobs with negative RTI to the most routine job profiles.[25] Examples among the 15 most common tasks are "coordinating the activities of a team", "monitor and analyze data on the activities of the service or organization propose a progress path" and "determine and monitor the budget of an organization". Therefore, these added tasks do not resemble a single set of skills but rather a bundle of common non-routine tasks. We interpret this mainly as a form of training in non-routine task content for routine job seekers. An alternative interpretation is that it resembles within-occupational shifts in task content towards more non-routine content with on-the-job learning. However, note that this training is general and crude, as it is not tailored to the profile of the job seeker. The important element of an effective retraining policy is that it allows routine job seekers to have high levels of overlap with job markets that have higher ex-ante job finding rates. Consistently, this simulation decreases the predicted unemployment duration for those job seekers with the most routine job profiles. As expected, the average predicted median duration when accounting for task overlap declines.

This simulation exercise emphasizes how a training policy could facilitate the equalizing

[25]These findings are comparable to adding the 15 most common tasks among non-routine jobs weighted by tightness to the most routine job profiles.

effects of overlapping markets. By increasing the overlap with less affected markets, the inequality in job finding is diminished.

The equalizing effect from overlapping markets in a case of industry specific exposure Given that only very similar jobs matter for job finding, it may seem that the role for overlapping markets to equalize exposure is limited by definition. To understand whether this is the case, we simulate a reduction in vacancies along a different dimension. We identify the specialization of jobs in manufacturing by checking for the words "production" and "industry/industrial" in the job title. Around 15% of our job seeker sample is identified to have at least one such job title in their profile and the mean share of such jobs in a profile is 9%. We simulate a reduction of 20% in the vacancies with such a title and calculate the predicted change in the unemployment duration for every job seeker. To understand the equalizing effect of overlapping markets for this simulated shock we compare the prediction from our full model to a model without overlap, as done for our main result in Figure 3. There is a positive gradient between the share of manufacturing jobs in the profile and the difference in the predicted unemployment duration by construction. The absolute size of the gradient is proportional to the simulated reduction in vacancies. However, Figure 4 also shows this gradient is about twice as large when restricting to a model without overlap. Note that the binscatter contains fewer data points given that the share of manufacturing jobs in a job search profile has a limited set of values.

This simulation shows that even when only the most similar jobs matter for job finding, there can still be a considerable equalizing effect from overlapping markets for a relative decline in vacancies. Equation (6) informs us under which condition this is the case: lower correlation of exposure among close neighbors. We consider two reasons why this could be the case for our chosen dimension of manufacturing specialized jobs. First, this exposure is meant to capture jobs that have a specific manufacturing industry attachment. If these manufacturing job titles also cover task content that is more general, then the exposure need not be as correlated. For example, if job seekers have experience as a production line manager, their profile likely contains tasks related to the production process. However, general managerial tasks are likely to provide additional overlap with other jobs that are not specific to the manufacturing industry and production lines alone. Thus, this simulated exposure focuses more on the industry rather than the occupation component which contributes to the lower correlation of exposure among close neighbors. Second, manufacturing specialized job titles are less pervasive. Shocks that affect a wide array of jobs are more likely to generate correlated shocks. Therefore, RTI exposure is more likely to be correlated among neighbors than the simulated shock in this section.

Taken together, this exercise shows that even in a job network with sparse close neighbors being relevant for job finding, overlapping markets *can* have an equalizing effect. The key condition which needs to be fulfilled is for exposure to be less correlated among close neighbors. We provide the example of an industry-specific shock.

7 Are other waves of technological change different?

Our result in Figure 3 reveals the interaction between exposure to routine-task automation and overlapping job markets for unemployed job seekers. Importantly, the distributional consequences may be different for other waves of technological innovations depending on which jobs are most exposed to the technology and their embedding within the job network. To highlight this, we apply three measures for exposure to recent technologies: artificial intelligence (AI), software, and robots, using the measures obtained from current patent data by Webb (2020). For each of these technologies, we repeat the comparison between the predicted unemployment duration when accounting for overlapping job markets and the benchmark model without task overlap.

Figure 5 presents us with two interesting findings. First, the gradient between technology exposure and predicted unemployment duration differs across technology types. There is a slight negative association between unemployment duration and AI exposure, which corresponds with the high share of high-skill jobs among jobs most exposed to AI (Webb, 2020). More importantly, accounting for the impacts of overlapping job markets has different predicted unemployment implications for different technologies. In particular, the positive association between software exposure and predicted unemployment duration declines when including overlapping job markets. The increase in predicted unemployment duration for one standard deviation increase in software exposure declines by 38% (from +2.9 days to +1.8 days) when accounting for task overlap. In contrast to the results on automation exposure captured by RTI, accounting for task overlap lowers the inequality in job finding among job seekers exposed to software automation.

To understand the difference in equalizing effects, Table 6 explores the correlation in exposure for each of these technologies among job seekers' search profiles and their most similar job markets. The degree of job segregation for these alternative technological innovations are different compared to routine-task automation and the degree varies across technology type.

Highly AI-exposed job seekers have a 40% share of other highly AI-exposed jobs among their close neighbors compared to not exposed job seekers, which is around 20%. Still, the difference in share of high exposed neighboring jobs is smaller than RTI exposure (49% to 22%, as shown in Table 5). For the highly-software exposed job profiles the most striking difference is the number of close neighbors connected to the search profiles. Highly software-exposed job seekers have on average 212 close neighbors compared to on average 124 close neighbors for not exposed job profiles. Importantly, these close neighbors provide different opportunities: the share of exposed neighbors lies just under the level of segregation for routineness, and the tightness of neighboring jobs is considerably better for highly exposed profiles than non exposed profiles. For this reason, overlapping jobs considerably lower the positive gradient between software exposure and unemployment duration. Exposure to robots is somewhat more segregated, but still comes with slightly better tightness in neighboring jobs. Hence, we find a decline in the gradient between robot exposure and unemployment duration when accounting for overlapping markets, albeit small. This could help explain why strong local and occupational displacement effects from robots have been found (Acemoglu and Restrepo, 2019; Acemoglu et al., 2020).

Overall, the degree of technology exposure among most similar jobs together with the tightness in neighboring job markets determines to what extent overlapping markets provide better prospects, which differs across types of technology.

8 Robustness to the choice of bin delineation for similar markets

The step function inside equation (1) is meant to capture a non-linear relationship which is likely to be more continuous and complex in reality. Although spline functions have been developed to account for non-monotonic, continuous relationships between two variables, we aim to estimate the return to the interaction between task similarity and tightness. Therefore, we estimate a simplified step function. This mimics the idea of a spline and similarly depends on arbitrary choices concerning bin definitions. Any measurement error associated with our bin delineation is likely to attenuate our estimates.[26] For this reason, we have created a range of alternative bin specifications. Columns (2) to (13) of Table 7 displays estimation results with different definitions of how we bin overlapping markets. The coefficients of our main analysis can be

[26] Another possibility is that our language model is not able to precisely distinguish similarity among similar job pairs in our very granular job classification such that there is a broader set of cut-offs that is consistent with close-range distance in task similarity.

found in column (1) for comparison. In columns (2) to (7) we change the definition of our first bin, trying to zoom in on the overlap with very similar neighbours. Columns (8) to (10) change multiple bin cut-offs and number of bins to test for the importance of how we condition on bins overall. Finally, columns (11) to (13) look at the importance of a more restrictive definition of similarity, where we drop jobs that are very dissimilar, i.e. at the lowest end of the similarity distribution. Overall, our results are qualitatively similar. Regardless of the definition of bins, the main conclusion remains that only the most similar jobs are important for job finding.

Regardless of their exact definition, we find that the coefficients on bins delineated along the first 10 percentiles of the distribution are statistically indistinguishable from each other. For example, making the first bin smaller (columns (4) and (5)) or splitting it into two separate bins (columns (6) and (7)) does not result in the highest coefficient on tightness in the most similar jobs (e.g., at the 98th - 99th percentile), followed by a smaller but still significant coefficient on tightness in jobs at follow-up percentiles of similarity (e.g., 97th - 95th). With the very disaggregate information we constructed on job similarity, we also lack statistical power when estimating the returns to tightness in these very small bins, in particular, related to variation in the number of job seekers.[27]

9 Concluding remarks

We present evidence on the importance of task overlap for job finding and its implications for exposure to automation. We build our evidence in several steps. After constructing a novel task similarity measure, we show that only the most similar jobs outside of job seekers' own job profiles matter for job finding using an extended matching function specification. By benchmarking our model to a model without overlap, we are able to test the importance of overlapping markets for the positive gradient found between unemployment duration and being displaced from routine jobs.

Overlapping markets provide little relief from reduced job finding for routine intensive job seekers since tightness in the neighboring markets is not sufficiently different to offer an escape hatch. The positive gradient between exposure and predicted unemployment duration remains unchanged when restricting the model to job search without overlapping markets. Thus, task-

[27]Because of the importance of the first bin for job finding, we also perform a robustness check on what jobs are most similar. In particular, we consider that lower experience levels of jobs listed in the profile are by construction included in the first bin with perfect overlap. Given that some job seekers may not find this appealing, we remove these cases of overlap from the first bin. However, our results remain unchanged.

based mobility does not diminish the distributional consequences of routine-task automation by itself.

We show that lowering the cost of mobility even increases their exposure to routine-task automation in close job markets. Therefore, there is potential harm in recommendation algorithms that direct job seekers to explore job markets that are most similar to their previous job's task content. Our findings confirm the necessity to include the market opportunities in the recommendation system for job seekers (Altmann et al., 2022; Behaghel et al., 2022; Belot et al., 2019, 2022; Le Barbanchon et al., 2023). To access jobs with better prospects, policymakers may provide targeted retraining to job seekers exposed to automation. For example, our findings promote the efforts made by sectoral employment programs such as "WorkAdvance" (Katz et al., 2022) and directed job search interventions by the Public Employment Services (Behaghel et al., 2022). Importantly, such programs retrain job seekers in less prospective markets to access local firms in tight labor markets.

Taken together, to understand workers' resilience to automation it is necessary to identify exposed jobs and examine how these are connected to other jobs. Using a task-based measure of job similarity, our results highlight the role of job segregation within the network. Jobs highly exposed to automation of routine tasks are segregated from less exposed and better prospect jobs, yielding concentrated low job finding rates for directly affected job seekers. However, we illustrate that recent advancements in technologies such as AI and software are less disruptive since they affect persons with better job finding perspectives or affect jobs that are connected to jobs with better prospects. More generally, the surrounding structure of the job network is relevant for the dispersion of the effects of any task-based labor market shock such as COVID-19, offshoring, or trade. As discussed in our final simulation, overlapping markets can have an equalizing effect if the correlation of exposure is low enough among close neighbors.

References

Acemoglu, D. and Autor, D. (2011). Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D., Lelarge, C., and Restrepo, P. (2020). Competing with Robots: Firm-Level Evidence from France. *AEA Papers and Proceedings*, 110:383–388.

Acemoglu, D. and Restrepo, P. (2019). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, pages 000–000.

Acemoglu, D. and Restrepo, P. (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*, 90(5):1973–2016.

Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., and Iyad Rahwan (2018). Unpacking the Polarization of Workplace Skills. *Science Advances*, 4(7):eaao6030.

Altmann, S., Glenny, A., Mahlstedt, R., and Sebald, A. (2022). The Direct and Indirect Effects of Online Job Search Advice. IZA Discussion Paper No. 15830, IZA Institute of Labor Economics.

Antonczyk, D., Fitzenberger, B., and Leuschner, U. (2009). Can a Task-Based Approach Explain the Recent Changes in the German Wage Structure? *Jahrbücher für Nationalökonomie und Statistik*, 229(2-3):214–238.

Autor, D. and Dorn, D. (2009). This Job Is "Getting Old": Measuring Changes in Job Opportunities Using Occupational Age Structure. *American Economic Review*, 99(2):45–51.

Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–1597.

Autor, D. H., Dorn, D., and Hanson, G. H. (2015). Untangling Trade and Technology: Evidence from Local Labour Markets. *The Economic Journal*, 125(584):621–646.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2006). The polarization of the u.s. labor market. *American Economic Review*, 96(2):189–194.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Battisti, M., Dustmann, C., and Schönberg, U. (2023). Technological and Organizational Change and the Careers of Workers. *Journal of the European Economic Association*, 21(4):1551–1594.

Behaghel, L., Dromundo Mokrani, S., Gurgand, M., Hazard, Y., and Zuber, T. (2022). Encouraging and Directing Job Search: Direct and Spillover Effects in a Large Scale Experiment.

Belot, M., Kircher, P., and Muller, P. (2019). Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice. *The Review of Economic Studies*, 86(4):1411–1447.

Belot, M., Kircher, P., and Muller, P. (2022). Do the Long-Term Unemployed Benefit from Automated Occupational Advice During Online Job Search? *SSRN Electronic Journal*.

Bessen, J., Goos, M., Salomons, A., and van den Berge, W. (2023). What Happens to Workers at Firms that Automate? *The Review of Economics and Statistics*, pages 1–45.

Blien, U., Dauth, W., and Roth, D. H. W. (2021). Occupational routine intensity and the costs of job loss: Evidence from mass layoffs. *Labour Economics*, 68:101953.

Borusyak, K., Dix-Carneiro, R., and Kovak, B. (2022). Understanding Migration Responses to Local Shocks. Working Paper, Accessed at: https://papers.ssrn.com/abstract=4086847.

Braxton, J. C. and Taska, B. (2023). Technological Change and the Consequences of Job Loss. *American Economic Review*, 113(2):279–316.

Cortes, G. M. (2016). Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data. *Journal of Labor Economics*, 34(1):63–105.

Cortes, G. M. and Gallipoli, G. (2018). The Costs of Occupational Mobility: An Aggregate Analysis. *Journal of the European Economic Association*, 16(2):275–315.

Cortes, G. M., Jaimovich, N., and Siu, H. E. (2017). Disappearing Routine Jobs: Who, How, and Why? *Journal of Monetary Economics*, 91:69–87.

Davis, S. J. and Von Wachter, T. (2011). Recessions and the Costs of Job Loss. *Brookings Papers on Economic Activity*, 42(2 (Fall)):1–72.

Dawson, N., Williams, M.-A., and Rizoiu, M.-A. (2021). Skill-Driven Recommendations for Job Transition Pathways. *PLOS ONE*, 16(8):e0254722.

del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F., and Farmer, J. D. (2020). Automation and Occupational Mobility: A Data-Driven Network Model. Manuscript, Accessed at: http://arxiv.org/abs/1906.04086.

Dustmann, C., Ludsteck, J., and Schönberg, U. (2009). Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, 124(2):843–881.

Dworkin, J. D. (2019). Network-Driven Differences in Mobility and Optimal Transitions Among Automatable Jobs. *Royal Society Open Science*, 6(7):182124.

Gathmann, C. and Schönberg, U. (2010). How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, 28(1):1–49.

Goos, M. and Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *Review of Economics and Statistics*, 89(1):118–133.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–2526.

Goos, M., Rademakers, E., and Röttger, R. (2021). Routine-Biased technical change: Individual-Level evidence from a plant closure. *Research Policy*, 50(7):104002.

Goos, M., Rademakers, E., Salomons, A., and Willekens, B. (2019). Markets for Jobs and Their Task Overlap. *Labour Economics*, 61:101750.

Gregory, T., Salomons, A., and Zierahn, U. (2022). Racing With or Against the Machine? Evidence on the Role of Trade in Europe. *Journal of the European Economic Association*, 20(2):869–906.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings Losses of Displaced Workers. *The American Economic Review*, 83(4):685–709.

Katz, L. F., Roth, J., Hendra, R., and Schaberg, K. (2022). Why Do Sectoral Employment Programs Work? Lessons from WorkAdvance. *Journal of Labor Economics*, 40(S1):S249–S291.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, 3(3):303–320.

Lazear, E. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy*, 117(5):914–940.

Le Barbanchon, T., Hensvik, L., and Rathelot, R. (2023). How can AI improve search and matching? Evidence from 59 million personalized job recommendations.

Macaluso, C. (2022). Skill Remoteness and Post-Layoff Labor Market Outcomes. Working Paper, Accessed at: https://experts.illinois.edu/en/publications/skill-remoteness-and-post-layoff-labor-market-outcomes.

Manning, A. and Petrongolo, B. (2017). How Local Are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review*, 107(10):2877–2907.

Michaels, G., Natraj, A., and Van Reenen, J. (2014). Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years. *Review of Economics and Statistics*, 96(1):60–77.

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory*. MIT Press, Cambridge, MA, second edition.

Poletaev, M. and Robinson, C. (2008). Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000. *Journal of Labor Economics*, 26(3):387–420.

Robinson, C. (2018). Occupational Mobility, Occupation Distance, and Specific Human Capital. *Journal of Human Resources*, 53(2):513–551.

Schubert, G., Stansbury, A., and Taska, B. (2021). Employer Concentration and Outside Options. Working Paper, Accessed at: https://papers.ssrn.com/abstract=3599454.

Song, K., Tan, X., Qin, T., Lu, J., and Liu, T.-Y. (2020). MPNet: Masked and Permuted Pre-training for Language Understanding. *Advances in Neural Information Processing Systems*, 33:16857–16867.

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2):235–270.

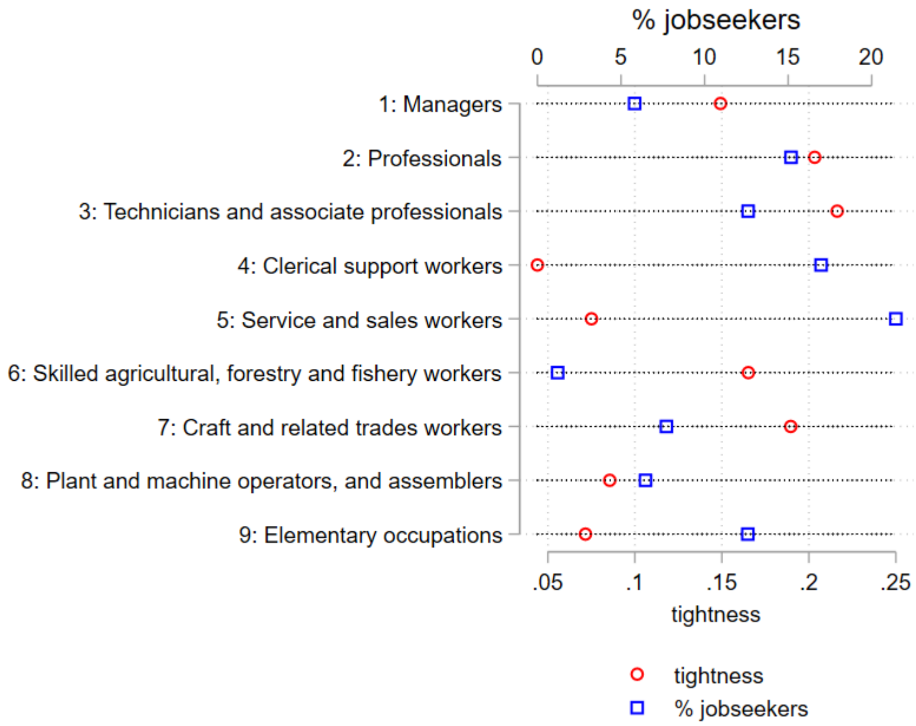
Traiberman, S. (2019). Occupations and Import Competition: Evidence from Denmark. *American Economic Review*, 109(12):4260–4301.

Vafa, K., Athey, S., and Blei, D. M. (2023). Decomposing Changes in the Gender Wage Gap over Worker Careers.

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. Working Paper, Accessed at: https://www.michaelwebb.co/webb_ai.pdf.

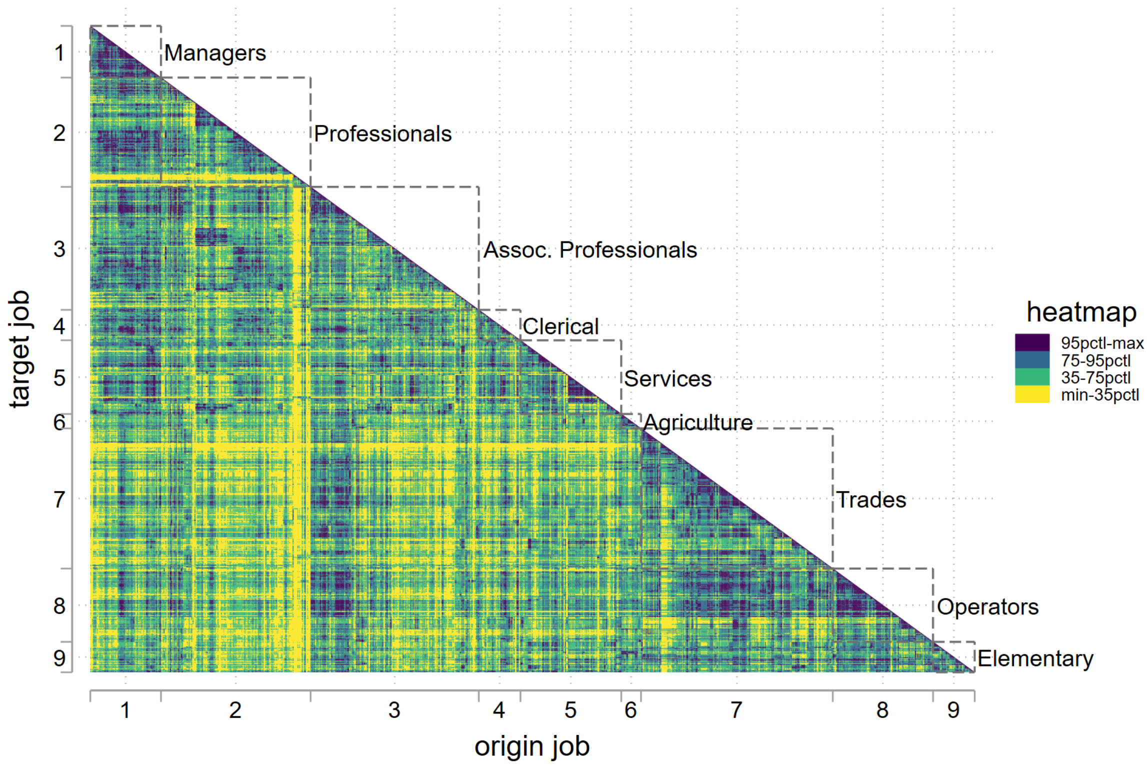
Figures and Tables

Figure 1: Tightness and the sample % of job seekers, on average by 1 digit ISCO08



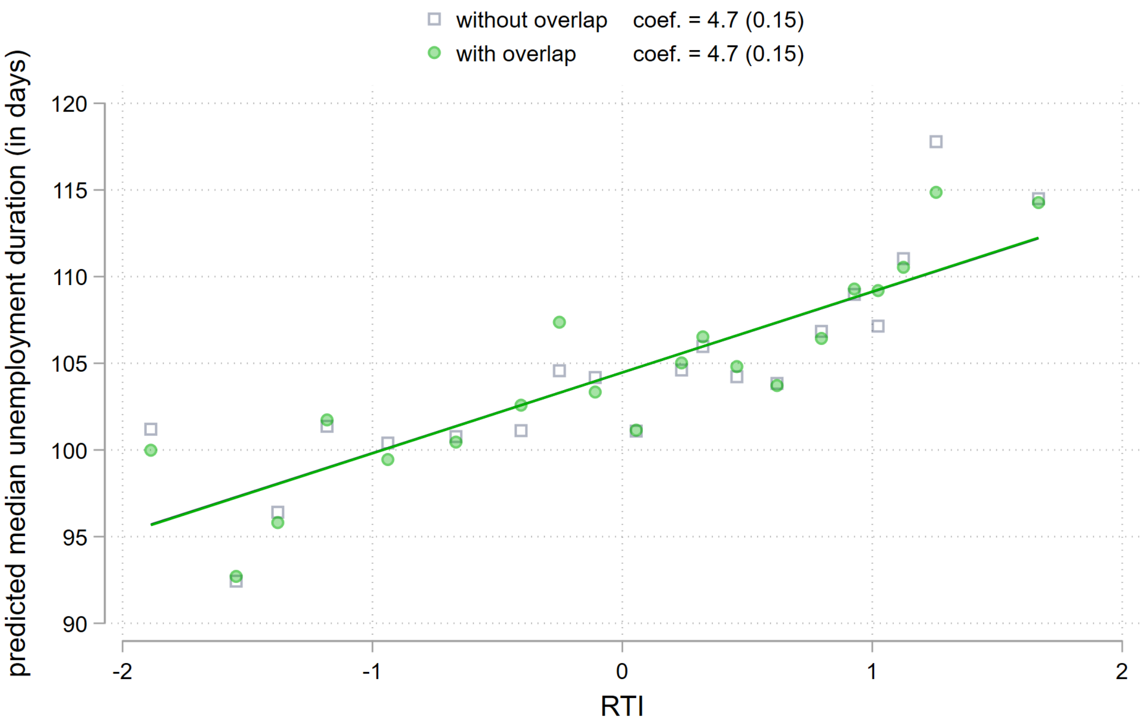
Source: VDAB: Mijn Loopbaan. *Notes:* The blue squares represent the share of our sample of job seekers across one-digit ISCO-08. The red circles represent the average tightness (vacancies/job seekers) in a job, by one-digit ISCO-08.

Figure 2: Cosine similarity between jobs ranked by ISCO-08



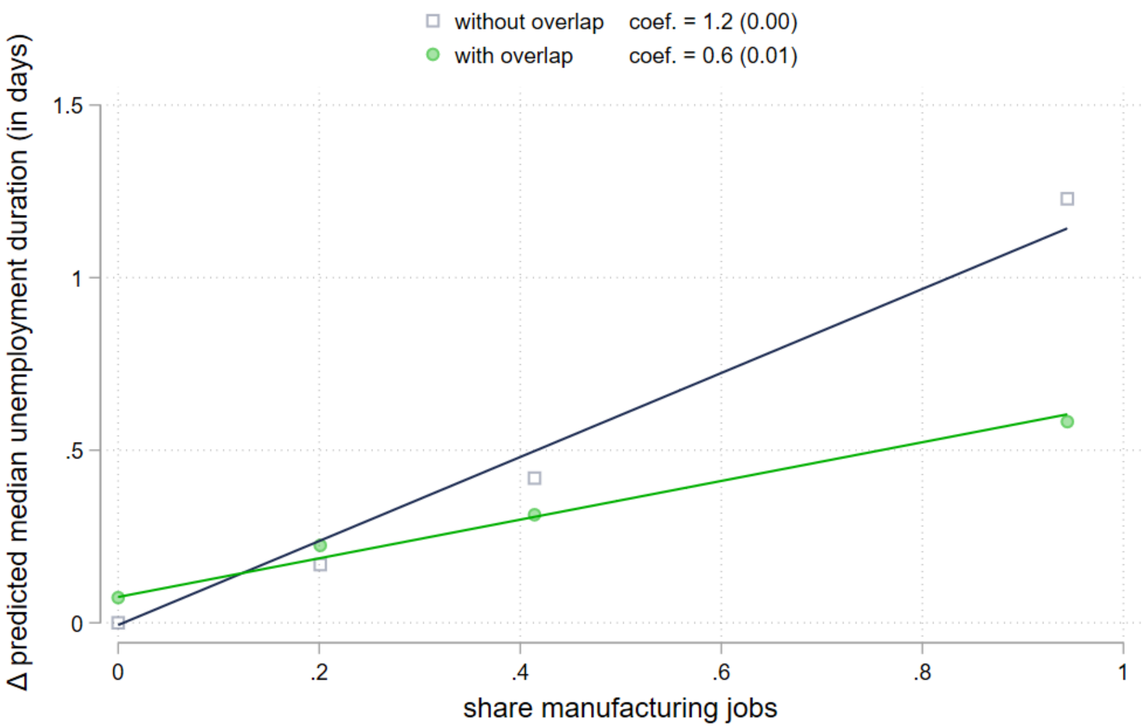
Source: VDAB: Mijn Loopbaan. *Notes:* Similarity between jobs based on cosine similarity after applying MPNet pre-training fine-tuned for sentence similarity at the task level (Song et al., 2020). The shades in heatmap are binned by the percentiles of the distribution of similarity. These bins correspond to the bins used for estimation in Table 3. Jobs of the x- and y-axis are ordered by four-digit ISCO-08 occupation codes.

Figure 3: Difference in predicted unemployment duration with versus without overlap by RTI



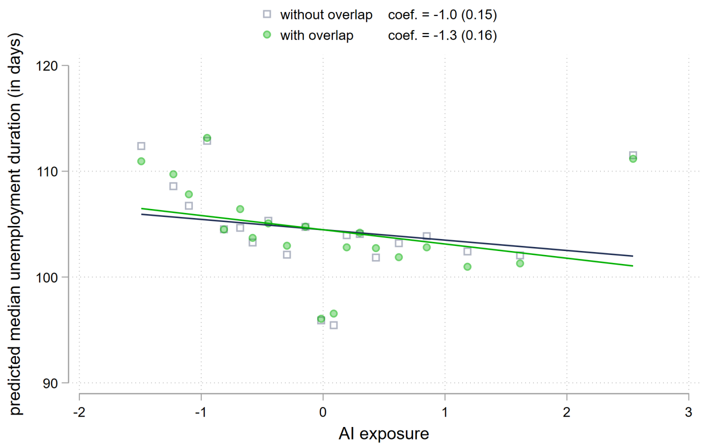
Source: VDAB: Mijn Loopbaan. *Notes:* The gray squares represent binscatters (n=20) of the predicted median unemployment duration (measured in days) without allowing for task overlap between jobs. The gray line represents the linear prediction across RTI, which predicts 4.6 days longer unemployment duration with a one standard deviation increase in the RTI measure. The green circles represent binscatters (n=20) of the predicted median unemployment duration (measured in days) with allowing for task overlap between jobs. The green line represents the linear prediction across RTI, which predicts 4.7 days longer unemployment duration with a one standard deviation increase in the RTI measure. RTI is constructed following Acemoglu and Autor (2011) and cross-walked to the four-digit ISCO-08 level.

Figure 4: The simulated difference in predicted unemployment duration with versus without overlap by manufacturing job title share

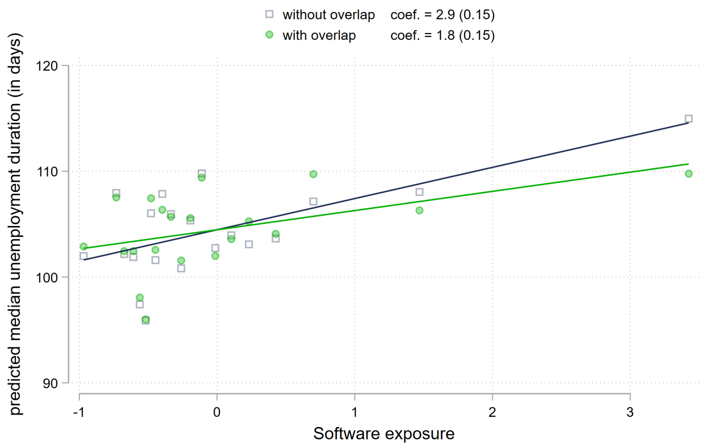


Source: VDAB: Mijn Loopbaan. *Notes:* The gray squares represent binscatters (n=4) of the predicted median unemployment duration (measured in days) without allowing for task overlap between jobs. The gray line represents the linear prediction across the share of manufacturing specialized job titles in the job profile, which predicts 1.2 days longer unemployment duration when the share increases from 0 to 1. The green circles represent binscatters (n=4) of the predicted median unemployment duration (measured in days) with allowing for task overlap between jobs. The green line represents the linear prediction across the share of manufacturing specialized job titles in the job profile, which predicts 0.6 days longer unemployment duration when the share increases from 0 to 1. Specialization of jobs in manufacturing is measured by identifying the words "production" and "industry/industrial" in the job title. We manipulated the stock of vacancies to be 20% lower in those jobs.

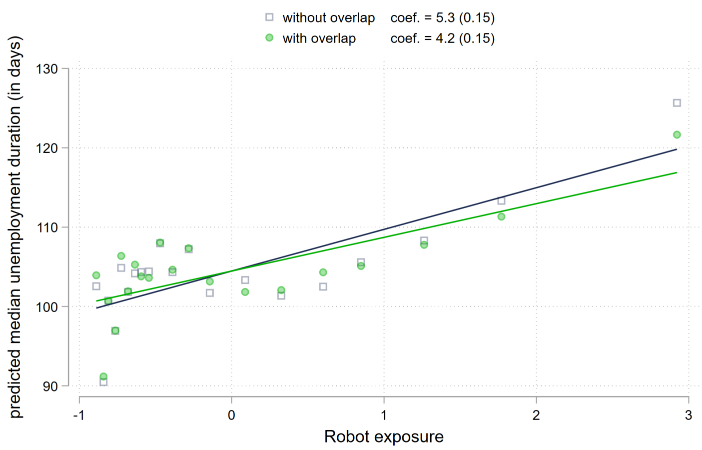
Figure 5: Exposure to recent technological innovations



(a) AI exposure



(b) Software exposure



(c) Robot exposure

Source: VDAB: Mijn Loopbaan. *Notes:* The gray squares represent binscatters (n=20) of the predicted median unemployment duration (measured in days) without allowing for task overlap between jobs. The gray line represents the linear prediction across the respective technology exposure without allowing for task overlap. The green circles represent binscatters (n=20) of the predicted median unemployment duration (measured in days) with allowing for task overlap between jobs. The green line represents the linear prediction across the respective technology exposure when allowing for task overlap. The technology exposures are constructed following Webb (2020) and cross-walked to the four-digit ISCO-08 level.

Table 1: Summary statistics on mean RTI score across jobs listed in the search profiles

one-digit ISCO-08	Mean RTI (1)
Managers	-3.35
	(1.12)
Professionals	-2.07
	(1.32)
Technicians and associate professionals	-0.43
	(2.54)
Clerical support workers	2.21
	(0.61)
Service and sales workers	0.64
	(1.64)
Skilled agricultural, forestry and fishery workers	-0.11
	(0.54)
Craft and related trades workers	0.63
	(1.88)
Plant and machine operators, and assemblers	0.81
	(1.52)
Elementary occupations	2.50
	(1.08)
Observations	33,352

Source: VDAB: Mijn Loopbaan. Standard errors in parentheses. The table shows the mean standardized RTI score across all jobs listed in the search profiles of job seekers. We construct the RTI information for four-digit ISCO-08 codes and show in this table the mean across one-digit ISCO-08 codes. Since 60% of the job seekers we observe in our sample list multiple jobs in their search profiles, the mean takes into account all the listed jobs specified those job seekers who list multiple jobs.

Table 2: Summary statistics on job seeker sample

	Mean (1)	S.D. (2)
Females	52.11	49.96
Age in years	30.66	9.90
Share of labor market entrants	30.39	46.00
Nationality		
Belgium	78.33	41.20
Europe	5.54	22.87
non-Europe	16.13	36.78
Knowledge Dutch		
None	2.58	15.87
Limited	6.52	24.69
Good	14.57	35.28
Very Good	76.33	42.51
Location		
Urban	23.41	42.34
Suburban	68.73	46.36
Rural	7.86	26.91
Education		
Primary	5.55	22.90
Some secondary	13.54	34.22
Secondary	19.02	39.25
Professional	26.51	44.14
Bachelor	22.29	41.62
Master	13.08	33.72
Observations	33,352	

Notes: Summary statistics for newly unemployed in Flanders between 01/03 and 10/09/2021. Labor market entrants refer to students who have just graduated from either secondary or tertiary education and are currently looking for jobs. Education is the highest attained. The specification of municipality of residence into urban, suburban and rural was done using the LAU correspondence at the niscode provided by Eurostat. Knowledge of Dutch is self-reported. Source: PES Flanders, Belgium.

Table 3: Importance of task overlap in job finding probability

Hazard rate of finding a job	Coefficient (1)	Std. err. (2)
search externality related to		
vacancies	0.0847***	(0.0137)
job seekers	−0.0396***	(0.0123)
rel. eff. of vacancies in jobs with similarity		
most similar $[max - 95pctl)$	0.2813***	(0.0721)
$[95pctl - 75pctl)$	0.0022*	(0.0012)
$[75pctl - 35pctl)$	−0.0003	(0.0002)
least similar $[35pctl - min)$	−0.0000	(0.0001)
rel. eff. of job seekers in jobs with similarity		
most similar $[max - 95pctl)$	0.2821	(0.2066)
$[95pctl - 75pctl)$	0.0004	(0.0056)
$[75pctl - 35pctl)$	−0.0005	(0.0014)
least similar $[35pctl - min)$	−0.0001	(0.0002)
Observations	33,352	
Person controls	✓	
Time FE	✓	
Region FE	✓	

Source: VDAB: Mijn Loopbaan. *Notes:* *, **, and *** stand for 10, 5 and 1% statistical significance respectively. Standard errors in parentheses. The estimation uses a log-logistic functional form and controls for person characteristics such as age, gender, region of birth, Dutch proficiency, urbanization of place of residence, and labor market relevant characteristics such as education, channel of registration on the platform, and whether job seeker is currently registered as a student. The coefficients are backed out from equation 1. See Appendix A.5 for more details.

Table 4: Results with overlap by age

	Age categories	
	younger than 35 years	35 years and older
Hazard rate of finding a job	(1)	(2)
search externality related to		
vacancies	0.0811***	0.1175***
	(0.0149)	(0.0315)
job seekers	−0.0457***	−0.0229
	(0.0133)	(0.0294)
rel. eff. of vacancies in jobs with similarity		
$[\,max\,-\,95pctl\,)$	0.1971***	5.4380
	(0.0457)	(40.1809)
$[\,95pctl\,-\,75pctl\,)$	0.0013	0.1056
	(0.0009)	(0.7840)
$[\,75pctl\,-\,35pctl\,)$	−0.0002	−0.0141
	(0.0001)	(0.1048)
$[\,35pctl\,-\,min\,)$	−0.0001	0.0009
	(0.0000)	(0.0086)
rel. eff. of job seekers in jobs with similarity		
$[\,max\,-\,95pctl\,)$	0.1575**	−0.0652
	(0.0748)	(0.0850)
$[\,95pctl\,-\,75pctl\,)$	−0.0017	−0.0190
	(0.0027)	(0.0141)
$[\,75pctl\,-\,35pctl\,)$	0.0004	0.0018
	(0.0007)	(0.0016)
$[\,35pctl\,-\,min\,)$	−0.0001	0.0004
	(0.0001)	(0.0012)
Observations	23,198	10,154
Person controls	✓	✓
Time FE	✓	✓
Region FE	✓	✓

Source: VDAB: Mijn Loopbaan. *Notes:* *, **, and *** stand for 10, 5 and 1% statistical significance respectively. Standard errors in parentheses. The estimation controls for person characteristics such as age, gender, region of birth, Dutch proficiency, urbanization of place of residence, and labor market relevant characteristics such as education, channel of registration on the platform, and whether job seeker is currently registered as a student.

Table 5: Summary statistics on the structure of the job network

	Routine profiles (1) Mean	Non-Routine profiles (2) Mean
Characteristics of neighbors in most similar bin 1		
No. of neighboring jobs	179.61	140.25
	(131.06)	(112.73)
Share of routine neighboring jobs (in %)	49.03	21.81
	(18.77)	(17.72)
Tightness of neighboring jobs	0.32	0.37
	(0.29)	(0.35)
Tightness of routine neighboring jobs	0.22	0.24
	(0.25)	(0.47)
Characteristics of jobs listed in own profile		
No. jobs listed	2.57	2.52
	(1.77)	(2.04)
No. of tasks in jobs listed	9.42	9.89
	(5.02)	(6.90)
Similarity between tasks in jobs listed	0.33	0.39
	(0.08)	(0.11)
Tightness in jobs listed	0.02	0.10
	(0.28)	(0.62)
Observations	11,116	22,236

Source: VDAB: Mijn Loopbaan. Standard errors in parentheses. The table shows summary statistics for the job network based on job profiles listed in the own search profiles and characteristics of the most similar jobs for job seekers which be define to include jobs up the 95th percentile of the overall cosine similarity distribution across job similarity pairs. Column (1) displays the average characteristics for routine job profiles and Column (2) displays the average characteristics for non-routine job profiles. To differentiate routine and non-routine job profiles, we follow the convention to classify job profiles in the top 33% of the RTI distribution as routine (Acemoglu and Restrepo, 2022). No. of tasks represent the average number of tasks across all jobs listed on the own job profile. The similarity between tasks gives the average similarity between tasks within each job listed in the search profile. The similarity between tasks gives the average similarity between tasks within each job listed in the search profile. The similarity is measured by comparing task embeddings using cosine similarity. The tightness across jobs listed in the own profile gives the average number of vacancies divided by the number of job seekers in each job listed in the search profile. The tightness of neighboring jobs gives the average number of vacancies divided by the number of job seekers in most similar jobs contained in the first bin and tightness of routine neighboring calculates the tightness only for routine jobs among the most similar jobs in the first bin.

Table 6: Summary statistics on structure of the job network for other waves of technological change

	AI		Software		Robots	
	Highly exposed	Not exposed	Highly exposed	Not exposed	Highly exposed	Not exposed
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	Mean	Mean	Mean	Mean	Mean
Characteristics of neighbors in most similar bin 1						
No. of neighboring jobs	189.03	135.60	212.88	124.59	205.59	127.26
	(132.63)	(109.88)	(139.11)	(98.36)	(137.89)	(101.30)
Share of highly exposed neighboring jobs (in %)	39.78	20.12	48.59	17.98	51.11	12.48
	(14.64)	(12.63)	(14.36)	(14.81)	(15.67)	(14.10)
Tightness of neighboring jobs	0.40	0.33	0.36	0.35	0.36	0.35
	(0.35)	(0.32)	(0.34)	(0.33)	(0.33)	(0.33)
Tightness of highly exposed neighboring jobs	0.41	0.44	0.32	0.35	0.33	0.50
	(0.81)	(0.62)	(0.36)	(0.78)	(0.37)	(0.75)
Characteristics of jobs listed in own profile						
No. jobs listed	2.48	2.57	2.71	2.45	2.80	2.40
	(1.78)	(2.04)	(1.93)	(1.97)	(1.92)	(1.96)
No. of tasks in jobs listed	10.34	9.43	9.86	9.67	10.09	9.56
	(7.68)	(5.53)	(6.42)	(6.30)	(6.35)	(6.33)
Similarity between tasks in jobs listed	0.38	0.37	0.36	0.38	0.34	0.39
	(0.14)	(0.09)	(0.12)	(0.10)	(0.07)	(0.12)
Tightness in jobs listed	0.10	0.06	0.06	0.08	0.08	0.07
	(0.64)	(0.47)	(0.46)	(0.57)	(0.57)	(0.52)
Observations	11,095	22,257	10,871	22,481	11,117	22,235

Source: VDAB: Mijn Loopbaan. Standard errors in parentheses. The table shows summary statistics for the job network based on job profiles listed in the own search profiles and characteristics of the most similar jobs for job seekers which be define to include jobs up the 95th percentile of the overall cosine similarity distribution across job similarity pairs. Column (1) displays the average characteristics for highly AI exposed job profiles and Column (2) displays the average characteristics for job profiles not highly exposed to AI. To differentiate highly AI-exposed and not AI-exposed job profiles, we classify job profiles in the top 33% of the AI score distribution as highly AI-exposed. We apply the same procedure to differentiate between highly software-exposed job profiles in Column (3) from not software-exposed profiles in Column (4) and to differentiate between highly robot-exposed job profiles in Column (5) from not robot-exposed profiles in Column (6). No. of tasks represent the average number of tasks across all jobs listed on the own job profile. The similarity between tasks gives the average similarity between tasks within each job listed in the search profile. The similarity is measured by comparing task embeddings using cosine similarity. The tightness across jobs listed in the own profile gives the average number of vacancies divided by the number of job seekers in each job listed in the search profile. The share of highly exposed neighbors gives the share of highly exposed jobs among the most similar jobs in the first bin. The tightness of neighbors gives the average number of vacancies divided by the number of job seekers in most similar jobs contained in the first bin and tightness of highly exposed neighbors calculates the tightness only for highly exposed jobs among the most similar jobs in the first bin.

Table 7: Alternative specifications for the importance of task overlap in job finding probability

	Baseline	Alternative bin 1						Alternative bins			Cut data at the end		
		cut-off variation				splitting bin 1		alternative bins			drop lowest similarities		
	95pctl	93pctl	94pctl	96pctl	97pctl	at 98pctl	at 99pctl	split bin 3	5 equal bins	10 equal bins	lowest 5%	lowest 7%	lowest 10%
Hazard rate of finding a job	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
search externality related to													
vacancies	0.0847***	0.1088***	0.0973***	0.0898***	0.0928***	0.0789***	0.0687***	0.0843***	0.2075***	0.1427***	0.0875***	0.0876***	0.0874***
	(0.0137)	(0.0149)	(0.0143)	(0.0131)	(0.0123)	(0.0116)	(0.0110)	(0.0137)	(0.0252)	(0.0189)	(0.0136)	(0.0136)	(0.0135)
job seekers	−0.0396***	−0.0623***	−0.0528***	−0.0405***	−0.0407***	−0.0285***	−0.0223***	−0.0381***	−0.1654***	−0.1020***	−0.0404***	−0.0402***	−0.0400***
	(0.0123)	(0.0137)	(0.0131)	(0.0114)	(0.0104)	(0.0097)	(0.0086)	(0.0123)	(0.0243)	(0.0174)	(0.0123)	(0.0123)	(0.0123)
rel. eff. of vacancies in jobs with similarity													
most similar *bin*1	0.2813***	0.3185***	0.2819***	0.3090***	0.3326***	0.3054***	0.2634***	0.2765***	0.2697***	0.3261***	0.2909***	0.2914***	0.2905***
	(0.0721)	(0.0799)	(0.0640)	(0.0741)	(0.0756)	(0.0655)	(0.0548)	(0.0703)	(0.0567)	(0.0749)	(0.0741)	(0.0741)	(0.0738)
*bin*2	0.0022*	0.0028**	0.0025**	0.0015	0.0008	−0.0018	−0.0008	0.0030**	0.0032	0.0379**	0.0023*	0.0024*	0.0024*
	(0.0012)	(0.0014)	(0.0012)	(0.0011)	(0.0006)	(0.0066)	(0.0022)	(0.0014)	(0.0069)	(0.0158)	(0.0013)	(0.0013)	(0.0013)
*bin*3	−0.0003	−0.0001	−0.0002	−0.0002	−0.0002	0.0012*	0.0007	−0.0008	−0.0024	0.0015	−0.0004*	−0.0004*	−0.0004**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0007)	(0.0005)	(0.0009)	(0.0036)	(0.0015)	(0.0002)	(0.0002)	(0.0002)
*bin*4	−0.0000	−0.0001	−0.0000	−0.0000	−0.0000	−0.0002	−0.0001	−0.0002	0.0004	0.0003	0.0000	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0005)	(0.0008)	(0.0041)	(0.0001)	(0.0001)	(0.0002)
*bin*5						0.0000	−0.0000	−0.0001	−0.0002**	−0.0013			
						(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0035)			
*bin*6										−0.0021			
										(0.0017)			
*bin*7										−0.0013			
										(0.0013)			
*bin*8										0.0018			
										(0.0011)			
*bin*9										−0.0001			
										(0.0006)			
least similar *bin*10										−0.0001			
										(0.0001)			

continued on next page ...

... *continued from previous page*

	Baseline	Alternative bin 1						Alternative bins			Cut data at the end		
		cut-off variation				splitting bin 1		alternative bins			drop lowest similarities		
	95pctl	93pctl	94pctl	96pctl	97pctl	at 98pctl	at 99pctl	split bin 3	5 equal bins	10 equal bins	lowest 5%	lowest 7%	lowest 10%
Hazard rate of finding a job	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
rel. eff. of job seekers in jobs with similarity													
most similar *bin*1	0.2821	0.7374	0.3751	0.2510*	0.2286**	0.1296**	0.0919**	0.2595	−0.3998**	0.9013	0.2590	0.2585	0.2591
	(0.2066)	(0.9050)	(0.2722)	(0.1512)	(0.1154)	(0.0547)	(0.0405)	(0.1857)	(0.1559)	(0.9380)	(0.1724)	(0.1725)	(0.1744)
*bin*2	0.0004	0.0189	0.0120	0.0006	0.0008	0.0120	0.0005	0.0102	0.0015	0.4930	0.0010	0.0028	0.0044
	(0.0056)	(0.0318)	(0.0139)	(0.0030)	(0.0007)	(0.0084)	(0.0017)	(0.0085)	(0.0182)	(0.4989)	(0.0052)	(0.0057)	(0.0062)
*bin*3	−0.0005	−0.0001	−0.0007	−0.0006	−0.0004	0.0007*	0.0000	−0.0136	0.0190	0.0210	−0.0008	−0.0015	−0.0021
	(0.0014)	(0.0027)	(0.0016)	(0.0011)	(0.0004)	(0.0004)	(0.0001)	(0.0092)	(0.0115)	(0.0472)	(0.0014)	(0.0017)	(0.0020)
*bin*4	−0.0001	−0.0002	−0.0001	−0.0001	−0.0001	−0.0004*	−0.0000	0.0066	−0.0048*	−0.0309	−0.0000	0.0001	0.0003
	(0.0002)	(0.0005)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0047)	(0.0029)	(0.0453)	(0.0003)	(0.0003)	(0.0005)
*bin*5						−0.0001	−0.0000	−0.0006	0.0003	−0.0265			
						(0.0001)	(0.0000)	(0.0004)	(0.0003)	(0.0346)			
*bin*6										−0.0236			
										(0.0316)			
*bin*7										0.0077			
										(0.0116)			
*bin*8										0.0071			
										(0.0102)			
*bin*9										−0.0007			
										(0.0040)			
least similar *bin*10										−0.0009			
										(0.0014)			
Observations	33,352	33,352	33,352	33,352	33,352	33,352	33,352	33,352	33,352	33,352	33,352	33,352	33,352
Person controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Source: VDAB: Mijn Loopbaan. *Notes:* *, **, and *** stand for 10, 5 and 1% statistical significance respectively. Standard errors in parentheses. The estimation controls for person characteristics such as age, gender, region of birth, Dutch proficiency, urbanization of place of residence, and labor market relevant characteristics such as education, channel of registration on the platform, and whether job seeker is currently registered as a student. Column (1) repeats the main estimation results. In column (2) the cut-off between bin 1 and bin 2 lies at the 93th percentile. In column (3) the cut-off between bin 1 and bin 2 lies at the 94th percentile. In column (4) the cut-off between bin 1 and bin 2 lies at the 96th percentile. In column (5) the cut-off between bin 1 and bin 2 lies at the 97th percentile. In column (6) bin 1 is split at the 98th percentile such that bin 1 goes from the max. until 98th percentile and bin 2 goes from 98th percentile until the 75th percentile while the remaining three bins remain unchanged compared to the baseline estimation. In column (7) bin 1 is split at the 99th percentile such that bin 1 goes from the max. until 99th percentile and bin 2 goes from 99th percentile until the 75th percentile. In column (8) the baseline bin 3 is split into two equally sized bins of each 20 percentiles while the remaining three bins remain unchanged compared to the baseline estimation. Column (9) splits the entire similarity distribution into five equally sized bins of each 20 percentiles. Column (10) splits the entire similarity distribution into ten equally sized bins of each 10 percentiles. In column (11) the lowest 5% of the similarity distribution are cut-off such that bin 4 contains 35th percentile until the 5th percentile. In column (11) and column (12) the lowest 7% and 10% of the similarity distribution are cut-off respectively.

Appendix

A.1 Task descriptions

Table A.1 contains two examples of jobs and their task descriptions listed in the task competency matrix called "COMPETENT" provided by the PES. First, we list the task descriptions of a translator with less than two years of experience (= experience level 2). The described activities relate to translation and language use. Several tasks get added to the eight tasks listed for translators with higher experience levels. The task competency matrix lists 24 distinct tasks for translators with at least five years of experience.

The second example lists the tasks of a medical-technical assistant with more than five years of experience (= experience level 4). The job contains 17 tasks that relate to assisting in medical procedures and organizing medical administrative files. For starters in this job (= experience level 1), only the first eleven tasks are listed. This demonstrates the importance of defining jobs as a combination of occupation and experience levels. Less experienced workers are only expected to conduct baseline activities on their job. Workers with more than five years of experience are able to conduct a greater variety of tasks, especially tasks associated with greater responsibility such as performing budget management of the organization. This example illustrates the monotonic increase in the number of tasks listed for a job with increasing experience levels. Jobs with the lowest experience level (= experience level 1) contain on average five tasks. Jobs with less than two years of experience (= experience level 2) and jobs between two and five years of experience (= experience level 3) execute on average eight tasks. The most experienced jobs with more than five years of experience (= experience level 4) conduct on average 23 distinct tasks.

A.2 Measures for exposure to technologies

We employ a commonly used measure for job seekers' exposure to automation. We use the routine task intensity (RTI) of occupations following Acemoglu and Autor (2011). This measure builds on the assumption that jobs that are intensive in routine tasks can be automated and displaced by machines. Routine work activities can be fully described by a set of rules and procedures (i.e., well-defined codifiable steps) such that they can be carried out by machines (i.e., computers and other electronics). These activities are distinct from non-routine tasks that are challenging to automate since they require creativity, intuition, problem-solving, adaptability, and responsiveness to unscripted in-person interactions. The RTI measure is based on O*NET

2009 and positive values indicate a job has a high intensity in routine tasks while negative values indicate high importance of non-routine tasks. The RTI is cross-walked from the SOC to ISCO-08 and merged at the four-digit ISCO-08 level. To calculate the RTI for each job seeker, we average the RTI across every job in the job seeker's profile.

We also employ three measures for job seekers' exposure to recent technological innovations. We use the suitability of occupations to be automated by artificial intelligence (AI), software and robots based on (Webb, 2020). These measures use the text of patents to identify activities that the technology can do and quantify the exposure of occupations to the specific type of technology by searching in the text of job descriptions for similar tasks. The AI exposure measure describes to what extent work activities can be conducted by machine learning algorithms such as supervised learning and reinforcement learning algorithms. The software exposure index captures the suitability of job tasks to be performed by software, i.e., computer programs that implement "if-then" procedures manually specified by humans. The robot exposure index measures the extent to which job activities can be carried out by industrial robots in manufacturing. These measures are computed by matching verb-noun pairs from patent descriptions obtained from the Google Patents Public Data to job descriptions listed in the O*NET 2017 (Version 22.0). The Google Patents Public Data is a large-scale dataset of patent publications updated quarterly since 1834 and provided by IFI CLAIMS Patent Services. Since the total number of patents has drastically increased over time (Kelly et al., 2021), patents for recent technical innovations are likely dominating the patents used for the verb-noun pairs. We cross-walked the resulting measures to ISCO-08 and merged at the four-digit ISCO-08 level. To calculate the respective technology exposure for each job seeker, we average the technology exposures across every job in the job seekers' profile.

While RTI is often presented as capturing exposure to software and robots, we find their occupational exposure does not coincide. This is most likely because measures based on current patents capture the more recent innovations. Out of the three, robot exposure correlates most with RTI as both relate to automation of production work, e.g., production operators. Still, robot exposure contains the automation of tasks of occupations such as forklift drivers, which are not considered to be routine since they require adaptability. Similarly, we find that software exposure partially overlaps with RTI. For example, software from current patents has the potential to automate the work activities of hand packers, a job typically described to be routine intensive. However, there is also a sizeable exposure of skilled professional jobs to software, suggesting this measure captures very recent advancements. Finally, AI is the least related to RTI, with even a slight negative correlation coefficient at the four-digit ISCO-08 level. This is plausible since AI targets tasks that are also located in skilled professional occupations. For example, chemical engineers, and opticians and optometrists have a high exposure to AI.

A.3 Using a language model to quantify similarity in task content

We illustrate our approach by means of the examples in Figure A.1. In the first step, the textual information of the task descriptions for the jobs archivist, translator and medical-technical assistant are transformed into numerical vectors based on pre-trained language representation model MPNet. While the jobs translator and medical-technical assistant share the task "administration and document management", the job archivist does not overlap in terms of exact task descriptions with translators or medical-technical assistants. In the second step, we average the task vectors within each job.[28] In the last step, we compare job vectors pairwise using cosine similarity. Using exact textual overlap would conclude there is no similarity between the job archivist and translator but indicate some similarity between translators and medical-technical assistants due to the one common task. However, our language model accounts for semantic meaning and detects that the tasks listed for translators relate to texts, books and documents which are very similar to archivists' tasks. In contrast, besides the document management task, the other tasks listed for medical-technical assistants are very specific medical tasks.[29] Consequently, the resulting task-based similarity between the jobs archivist and translator is higher compared to the similarity between translator and medical-technical assistant when accounting for semantic meaning in addition to exact textual overlap. This example highlights the ability of the language model to capture similarities between jobs beyond exact textual overlap in the detailed task descriptions.

A.4 The job search model allowing for overlapping markets in a nutshell

In this subsection, we provide an intuitive discussion of the job search model allowing for overlapping markets and highlight how our estimation approach relates to the structural model developed by Manning and Petrongolo (2017). The core of the structural model by Manning and Petrongolo (2017) is an optimal search strategy across *geographic space* to characterize

[28] In practice, different aggregation methods lead to very similar ordinal results of our similarity measure.
[29] The full list of tasks is shown in Table A.1.

local labour markets which provides the theoretical foundation for our understanding of search strategies across jobs that are differentiated in the *task space*. We implement a reduced-form estimation approach as discussed by Manning and Petrongolo (2017) in Appendix D. Nevertheless, the reduced form estimates are linked to the structural parameters of the model, which we explain here in brief terms. As part of their reduced-form analysis, Manning and Petrongolo (2017) provide empirical evidence that the structural parameters vary with the reduced-form estimates in line with the model predictions. This further strengthens the validity of inferring information from this reduced-form matching function.

In the original model by Manning and Petrongolo (2017), workers decide whether to work in their own local labor market or to commute to jobs in adjacent labor markets across geographic space. We provide an analogous interpretation with respect to task space. At the core of our job search model with overlapping markets, similarity between jobs arises from similarity in task content. The more similar two jobs, the smaller their distance in task space. Job seekers face the trade-off between searching for jobs in which they have experience in or "commute" to jobs that are similar in task content but in which they do not have experience in. Unlike geographic space, occupational task space does not have to be laid out smoothly but may be much more disjointed.

To optimize job search, job seekers trade off the cost of applying to a greater number of vacancies with the expected utility of finding a job. In the model by Manning and Petrongolo (2017) workers are located in areas a, but can decide to commute to other areas b. In our application this translates to job seekers having experience in occupations a, but can apply to other occupations b. The utility of working in occupation b after having experience in a is expressed as $\Omega_{ab}\varepsilon_i$. The first component depends on the probability that the application is successful and the attractiveness of working in b for a person with experience in a, which we specify in more detail later. The second component is an idiosyncratic utility component, which is assumed to be i.i.d. across job seekers and vacancies, and may capture individual specific preferences for working in a given occupational labor market, among other things.

Job seekers are assumed to have a cost function for sending N applications as follows:

$$C(N) = \frac{c}{1+\eta} N^{1+\eta} \tag{A.1}$$

The maximization of expected utility implies that a job seekers applies to all vacancies that have an expected utility net of marginal cost $C'(N)$.[30] By assuming that ε is Pareto distributed with exponent k, the probability that a job seeker with experience in a applies to a vacancy in b has a closed form expression:

$$\Pr\left(\Omega_{ab}\varepsilon_i \geq cN_a^{\eta}\right) = \left(\frac{\Omega_{ab}}{cN_a^{\eta}}\right)^k \tag{A.2}$$

This can be summed across all possible vacancies across destinations b and solved for N_a. The total number of applications sent by a given job seeker with experience in a is given as:

$$N_a = \left(c^{-k} \sum_b V_b \, \Omega_{ab}^k\right)^{\gamma} \tag{A.3}$$

where $\gamma = 1/(1 + \eta k)$, and relates to the returns to scale in the matching function.[31] To make this equation operational, Manning and Petrongolo (2017) assume the expected utility of working in occupation b after having experience in a has the following functional form:

$$\Omega_{ab} = p\left(\frac{\sum_a N_{ab} U_a}{V_b}\right) W_b f_{ab} \tag{A.4}$$

where N_{ab} is the number of workers in a that apply to jobs in b and $p\left(\frac{\sum_a N_{ab}U_a}{V_b}\right)$ is the probability of an application being successful and is assumed to depend *negatively* on the number of applicants per vacancy: $\frac{\sum_a N_{ab}U_a}{V_b}$. W_b is the wage offered by jobs in occupation b and f_{ab} represents the intrinsic attractiveness of a job in occupation b for a job seeker with experience in occupation a. In the original model, Manning and Petrongolo (2017) see the intrinsic attractiveness as primarily a function of distance. The further a worker needs to commute, the less attractive it becomes. Note that the cost of distance cannot be separately identified from the parameter k. The larger k , the more varied the idiosyncratic characteristic of specific jobs are, and the more job seekers might be willing to accept jobs further away. Observationally, this would appear the same as a decline in the cost of distance. In our application, one could consider the role of task distance between occupations as two-folded. First, it could lower the intrinsic attractiveness of working in an occupation b because it implies greater adaptation on the side of the job seeker. Second, and more complicated to the problem of the job seeker, it might also directly affect the probability of

[30] Importantly, the assumption is made that the probability of more than one application being successful is infinitesimal. Otherwise, the worker's decision becomes far more complicated as vacancies cannot be ranked by the expected utility offered. Further discussion can be found in footnote 6 of Manning and Petrongolo (2017, p.2883).

[31] More specifically, it represents the marginal cost of an application and therefore determines how many new applications would arise after doubling vacancies and job seekers.

getting hired. For example, job seekers might be less considered for hiring because of lack of experience in the exact job title. See Section 4.1 for further discussion on how we interpret the cost of task distance.

The relationship that we test in the main text of this paper can already be seen from substituting equation A.4 in equation A.3. The number of applications will depend on the sum of the different occupational markets and their respective opportunities. This depends negatively on the number of applications per vacancy in other occupations and will depend positively on the tightness in other occupations. Importantly, the ex ante probability of getting hired in other occupations is interacted with the attractiveness of jobs in those occupations. This in turn depends negatively on the task distance between occupations a and b. Instead of discretely summing over all other possible occupational destination, our empirical equation features a binned distances between occupations as we explain in detail in the subsequent Appendix Section A.5.

Although applications cannot be observed directly, Manning and Petrongolo (2017) link these positively to the probability of getting hired through a simple urn-ball framework of job matching with a constant probability (< 1) of an application being suitable. This allows us to estimate the link between the right hand side of equation A.3 and job finding. Our main estimation reflects this relationship of summing over all possible occupational markets after binning them into four bins ranked according to the similarity in task content.

A.5 Estimation approach

Our estimation approach follows (Goos et al., 2019) and (Manning and Petrongolo, 2017, Appendix D.). We briefly summarize the log-linearization and estimation procedure in this section. We define the job finding probability T_i for job seeker i:

$$ln(T_i) = \alpha_1 \, ln(V_i^{full} + \sum_{j=1}^{4} \gamma_j V_{ij}^{partial}) + \alpha_2 \, ln(U_i^{full} + \sum_{j=1}^{4} \beta_j U_{ij}^{partial}) \tag{A.5}$$

with V_i^{full} being the number of vacancies in jobs with full task overlap and U_i^{full} being the number of other job seekers in jobs with full task overlap for job seeker i. This composes the tightness in jobs that are part of the job seekers' profile. We account for other jobs by binning jobs into four bins indicated by j. These four bins group jobs according to the task similarity by percentiles of the overlap cosine similarity distribution. Consequently, $V_{ij}^{partial}$ indicates the number of vacancies in jobs with partial task overlap and $U_{ij}^{partial}$ gives the number of other job seekers in jobs with partial task overlap. The coefficient α_1 captures the positive search externality from additional vacancies for which job seeker i (at least partially) qualifies in terms of task competencies. Respectively, α_2 captures the negative search externality from additional job seekers searching for jobs for which job seeker i (at least partially) qualifies in terms of task competencies.

In our sample of unemployed job seekers, we encounter a non-monotonic duration dependence, with first increasing and then decreasing hazard rates. Therefore, we estimate a log-logistic survival function:

$$\begin{aligned} S(\tau_i) &= \frac{1}{1 + (\tau_i \, \exp[-(T_i + X_i'\delta)])^{1/\kappa}} \\ &= \frac{1}{1 + (\tau_i \, \exp[-[\alpha_1 \, ln(V_i^{full} + \sum_{j=1}^{4} \gamma_j V_{ij}^{partial}) + \alpha_2 \, ln(U_i^{full} + \sum_{j=1}^{4} \beta_j U_{ij}^{partial})]])^{1/\kappa}} \end{aligned} \tag{A.6}$$

The vector X_i' contains a set of control variables describing relevant job search characteristics of job seeker i. It includes person characteristics such as age, gender, region of birth (Belgium, Europe, Non-Europe), Dutch proficiency (none, limited, good, very good), urbanisation of place of residence (urban, suburban, rural), and labor market relevant characteristics such as education (primary, lower secondary, upper secondary, post-secondary non-tertiary, bachelor, master), channel of registration on the platform (batch, unemployment agent, job seeker himself) and whether jobs seeker is currently registered as a student. We also control for month- and region-fixed effects. Finally, unemployed job seeker i's expected unemployment duration is given by:

$$E[S(\tau_i)] = \int_0^\infty S(\tau_i) \, d\tau_i \tag{A.7}$$

which can then also take into account right-censoring in the observation.

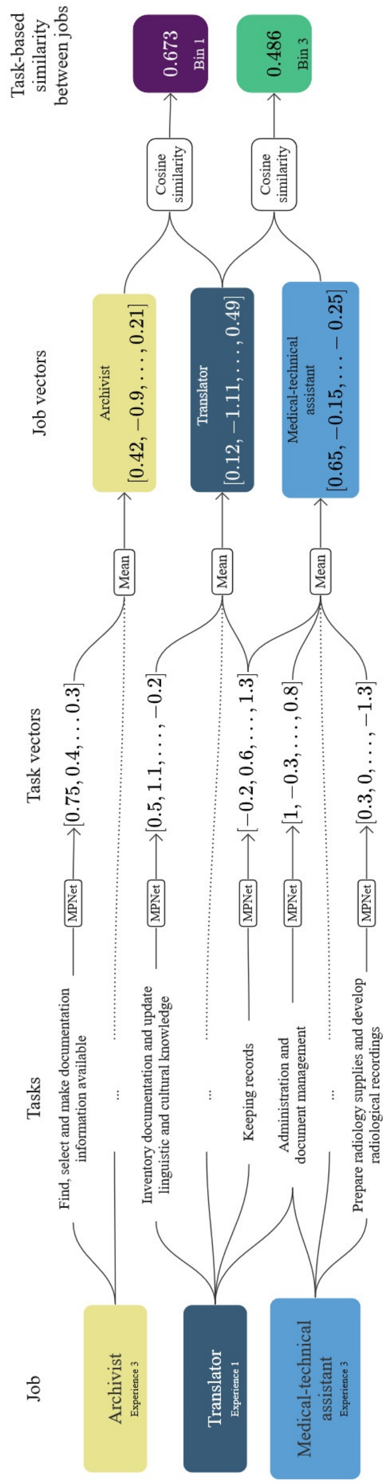
However, if γ_j and β_j are not all zero for $j = [1, \ldots, 4]$, equation A.5 is not linear in parameters α_1 and γ_j or in α_2 and β_j. We follow (Manning and Petrongolo, 2017) and apply a linear approximation to estimate γ_j and β_j. Therefore, we define $V_i \equiv V_i^{full} + V_{i1}^{partial}$ to be the number of vacancies for which i has full task overlap plus the vacancies from the first bin, that means jobs with 5% highest overlap. Equally, we define $U_i \equiv U_i^{full} + U_{i1}^{partial}$ to be the number of job seekers in jobs for which job seeker i possess all the required task competencies and the number of job seekers in other jobs in the first bin for with 5% highest overlap. We then approximate $T_i + X_i'\delta$ by the following first order Taylor approximation using V_i and U_i as normalization point:

$$
\begin{aligned}
T_i + X_i'\delta \approx\ & \alpha_1 \ln(V_i) + \alpha_2 \ln(U_i) + \alpha_1 \frac{1-\gamma_1}{\gamma_1} \frac{V_i^{full}}{V_i} + \alpha_2 \frac{1-\beta_1}{\beta_1} \frac{U_i^{full}}{U_i} + X_i'\delta \\
& + \alpha_1 \sum_{j=1}^{4} \frac{\gamma_j}{\gamma_1} \frac{V_{ij}^{partial}}{V_i} + \alpha_2 \sum_{j=1}^{4} \frac{\beta_j}{\beta_1} \frac{U_{ij}^{partial}}{U_i}
\end{aligned}
\tag{A.8}
$$

where $\alpha_1\, ln(\gamma_1) + \alpha_2\, ln(\beta_1)$ are absorbed by the constant term in X_i and its coefficient. Substituting the right hand side of equation A.8 into equation A.6, the estimated coefficients on $\ln(V_i)$ and $\ln(U_i)$ are point estimates of α_1 and α_2 respectively and, together with estimated coefficients on $\frac{V_i^{full}}{V_i}$ and $\frac{U_i^{full}}{U_i}$, allow us to back out estimates for γ_1 and β_1. Given these point estimates for α_1 and γ_1, one can then compute a value for γ_j from the estimated coefficient on $\frac{V_{ij}^{partial}}{V_i}$ for $j = [1, \ldots, 4]$. Similarly, having an estimate for α_2 and β_1 allows us to calculate a value for β_j from the estimated coefficient on $\frac{U_{ij}^{partial}}{U_i}$ for $j = [1, \ldots, 4]$. Table 3 contains the coefficients for α_1 and α_2 as well as backed out coefficients for γ_j's and β_j's.

A.6 Appendix Figures and Tables

Figure A.1: Diagram of the language algorithm for task similarity



Source: VDAB: Mijn Loopbaan. *Notes:* Illustration of our approach for constructing task-based similarity between three exemplary jobs. Jobs are composed of a list of detailed task descriptions. The full list of task descriptions for the job translator and medical-technical assistant are listed in Table A.1. First, we translate the textual information of the task descriptions into numerical vectors using MPNet fine-tuned for sentence similarity (Song et al., 2020). Second, we average the task vectors within each job and compare in the third step job vectors pairwise using cosine similarity. For estimating the matching elasticities of overlapping job markets, we group these task-based similarities based on the entire distribution of similarities across jobs into four bins.

Table A.1: Example of task descriptions in COMPETENT

Translator, less than 2 years of experience
Administration and document management
Check translations and terminology use (quality, term, revision, correction, ...)
Determine the modalities for translation or interpretation according to the target audience, the context and according to the needs of the client
Drafting and designing the translation of documents, texts, speeches, ...
Inventory documentation and update linguistic and cultural knowledge
Keeping records
Monitor and update glossaries, databases and translation tools
Understand a text or dialogue and reconstruct it in another language for interlocutors or on documents
Medical-technical assistant, more than 5 years of experience
Administration and document management
Check the operation of the devices inform the maintenance service and/or the person responsible in case of problems
Create the medical administrative file of the person with care needs or refer an additional person with care needs to the waiting room, care or examination room, ...
Disinfect, identify, package (print, ...) the medical-technical elements and send them to the relevant service
Follow up the file of the person with medical need for care
Follow up the product stock, determine shortages pass on orders
Inform the medical-technical elements of the receiving doctor and person with care needs with a view to intervention
Plan the appointments according to the needs of the person with care needs, the type of intervention, the urgency and inform the doctor
Prepare the material, the products according to the doctor's instructions and the nature of the intervention
Receive the person with care needs and inform them about the opening hours, the appointment options, the necessary documents
Tidying up and cleaning workspace, material and products after a consultation
Assist the dentist during the procedure (indicating the instruments, ...)
Perform operations for the accounting and budget management of the organization
Prepare radiology supplies and develop and store radiological recordings
Prepare radiology supplies and develop radiological recordings
Sensitize and advise people with care needs about oral and dental hygiene and the maintenance of dental prostheses
Sterilize and package the material and instruments (assemble sets, put in bags, ...)

Source: PES Flanders, Belgium. *Notes:* The table lists the descriptions of task competencies for two exemplary jobs.

GI-NI PROJECT IDENTITY

Project name

Growing Inequality: a novel integration of transformations research — GI-NI

Coordinator

Nederlandse Organisatie Voor Toegepast Natuurwetenschappelijk Onderzoek TNO, Netherlands

Consortium

CNAM – CEET, Centre d`études de l´emploi et du travail (France)
University of Groningen (Netherlands)
Centre for European Policy Studies (Belgium)
University of Adger (Norway)
Centre for Economic and Regional Studies (Hungary)
Utrecht University (Netherlands)
Europa-Universität Flensburg (Germany)
University of the Basque Country (Spain)

Duration

2021 – 2025

Funding Scheme

Grant Agreement no 101004494 — GI-NI — H2020-programme

Website

https://www.gini-research.org

Growing Inequality:
A novel integration of
transformations research

www.gini-research.org