RESEARCH REPORT
STUDY PAPER 279
MARCH 2026

Relative Wage Misperceptions and Early Career Job Search

Asker Nygaard Christensen, Nikolaj Arpe Harmon
and Sonja Settele Daphné Skandalis

THE ROCKWOOL FOUNDATION
RESEARCH

STUDY PAPER 279 MARCH 2026

Relative Wage Misperceptions and Early Career Job Search

Published by:

Address:
The ROCKWOOL Foundation Research Unit
Ny Kongensgade 6
1472 Copenhagen, Denmark

Telephone +45 33 34 48 00
E-mail: kontakt@rff.dk
en.rockwoolfonden.dk/research/

March 2026

Relative Wage Misperceptions and Early Career Job Search

Asker Nygaard Christensen Nikolaj Arpe Harmon
Sonja Settele Daphné Skandalis

December 23, 2025

Abstract

Misperceptions about the relative wages offered in different jobs may distort job search, especially among young workers entering the labor market. We study this using a survey of early-career job seekers in Denmark linked to administrative data. The survey i) elicits job seekers' beliefs about typical wages in three jobs that are relevant to them, ii) contains a randomized information treatment revealing actual typical wages, and iii) elicits job seekers' beliefs about their own potential wages in each job and their planned search behavior. Comparing beliefs about typical wages to administrative data reveals large relative misperceptions. In 80% of cases, the perceived wage gap between two jobs differs from the truth by more than 50%. In two-thirds of cases, job seekers underestimate the true gap, meaning they perceive lower-paying jobs to be overly attractive. Leveraging the information treatment, we show that these misperceptions causally affect search. Receiving information about actual typical wages causes job seekers to update beliefs about their own potential wages, which in turn changes their planned applications: a 1% increase in the perceived wage of one job over another increases the relative likelihood of applying for that job by 4.2%. This affects actual post-survey wages measured in administrative data: since misperceptions mostly inflate the attractiveness of lower-paying jobs, the information treatment shifts most job seekers to ultimately obtain higher-paying jobs. In a simple discrete choice framework, we estimate that removing relative wage misperceptions would reallocate 9.0% of workers to different jobs and increase wages by 1.2%.

JEL-Codes: J31, J62, J64, D83, D84

Contact: Asker Nygaard Christensen, University of Copenhagen, anc@econ.ku.dk. Nikolaj Harmon, University of Copenhagen, nikolaj.harmon@econ.ku.dk. Sonja Settele, University of Cologne, ECONtribute and MPI for Research on Collective Goods Bonn settele@wiso.uni-koeln.de. Daphné Skandalis, University of Copenhagen, daphne.skandalis@econ.ku.dk. **Acknowledgements**: We thank Michele Belot, Teodora Boneva, Ingrid Haegele, Philipp Kircher, Roland Rathelot, Jeremias Klaeui and Michael Siegenthaler. Annasofie Marckstrøm Olesen, Sophie Laura Bune, Dánjal Hátún Augustinussen, Carl Benedikt Huels, and Emma Thevissen provided excellent research assistance. **Funding**: Financial support through the Rockwool Foundation is gratefully acknowledged. Sonja Settele additionally acknowledges funding from the Danish National Research Foundation (CEBI, Grant DNRF134) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866. **Ethics approval:** Ethics approval was received from the IRB of the University of Copenhagen. **Pre-registration:** The data collection was pre-registered at the AEA RCT Registry (`https://www.socialscienceregistry.org/trials/11592`).

1 Introduction

A growing body of research demonstrates that workers' misperceptions about labor market conditions can prolong unemployment. Yet a distinct—but far less studied—consequence of imperfect information is the potential misallocation of workers across jobs. Workers engaged in active job search typically face a range of relevant jobs that differ in potential wages and non-wage amenities. If workers misperceive wage differences, they may sort into positions that poorly match their preferences: some workers may end up in suboptimally low-paying jobs while others may end up in suboptimally high-paying jobs with unattractive non-wage amenities. These distortions may be particularly prevalent and consequential among early career workers who have limited experience with wages and where job choices may have lasting effects on their career trajectories.[1]

To examine the consequences of relative wage misperceptions, we conducted a survey with early career job seekers and graduates from higher education in Denmark, which we analyze in conjunction with individually-linked administrative data on actual job outcomes. Recruiting via university registers and real-time unemployment data, we obtained a final sample of 1,902 respondents, all surveyed around the time of active job search. In the survey, job seekers' were asked about three specific types of jobs relevant to recent graduates with their educational background. For each job, the survey first included an incentivized question about the average wage earned by prior cohorts from the same educational degree as the respondent. Comparing answers to ground truth wage data from administrative records allows us to directly measure job seekers' misperceptions about typical wages in each job. Next, to examine whether misperceptions causally affect search and job outcomes, the survey included a randomized information treatment. Based on administrative data, treated job seekers were shown the actual average wages for prior cohorts in the three jobs. Finally, the survey elicited job seekers' planned applications and beliefs about the terms the job seekers *themselves* expect to face in each of the three jobs, including the potential wage they expected to receive if hired.

[1] A large literature highlights the persistence of negative effects of graduating in a bad labor market (e.g. von Wachter and Bender (2006), Kahn (2010), Cockx and Ghirelli (2016), and Wachter (2020)). Arellano-Bover (2024) documents long-term positive effects from getting a first job at a large firm as opposed to a small firm.

The first part of our analysis provides descriptive evidence about the extent and nature of relative wage misperceptions. Using survey responses about average wages of prior cohorts, we compute job seekers' perceptions of the typical wage gap between each of the three jobs in the survey. We then compare these to the actual gaps in wages among prior cohorts in administrative data. Perceived and actual wage gaps are clearly related (regression slope 0.54; correlation 0.27), suggesting that job seekers are broadly aware of wage differences across jobs. At the same time, substantial relative wage misperceptions exist: In 80% of cases, the perceived wage gap between two jobs differs from the truth by more than 50%. While there is substantial heterogeneity, the typical pattern is that job seekers understate the true gap in wages so that lower-paying jobs appear relatively more attractive. Approximately 65% of wage gaps are underestimated, with 55% underestimated by more than 50% and 36% underestimated by more than 100%—meaning respondents actually perceive the lower-paying job to pay higher wages. This pattern of predominantly underestimated wage gaps holds at the individual level as well: two-thirds of job seekers in our sample are underestimators, in the sense that they underestimate the gap between the highest-paying job and the other two jobs.

We substantiate that these measured misperceptions are practically meaningful in a number of ways. Since we compare beliefs about an objective benchmark—realized average wages of prior cohorts—the observed misperceptions cannot be explained by respondents having private information about their own situation or about labor market conditions. As we verify in the data however, prior cohort wages provide a natural benchmark for the belief-formation of graduates and early-career workers: beliefs about typical wages for prior cohorts are strongly correlated with beliefs about the potential wages job seekers' themselves expect to face. Examining alternative measures and subsamples, we also find no evidence that measured misperceptions are driven by job seekers misunderstanding the three jobs asked about in the survey or by job seekers not viewing these jobs as relevant for themselves. In line with theories of rational inattention, we find some evidence that job seekers are better informed when true wage gaps are larger and thus more consequential. Large misperceptions remain even in these cases however.

In the second part of our analysis, we leverage the randomized information treatment to establish the causal effects of relative wage misperception on job search. After job seekers have reported their beliefs about wages of prior cohorts, treated job seekers are shown the actual wages for prior cohorts allowing them to infer their own misperceptions about the different jobs. Using post-treatment survey questions, we show that this causes job seekers to update beliefs about the wage differences they themselves expect to face and also changes their planned application behavior accordingly. When job seekers are informed that the wage gap between two jobs for prior cohorts is in fact $X\%$ larger than they thought, the perceived gap in wages that the job seekers themselves expect to face increases by $0.3X\%$. In turn, the relatively likelihood of applying to the two jobs changes by $1.2X\%$.

Combining the reduced form treatment effects on wage beliefs and application behavior, we estimate directly how misperceptions about own potential wages affect job search. Formally, this amounts to a 2SLS estimator where the information shock from our treatment serves as the instrument for the perceived gap in potential wages between jobs. We find a large effect: a 1% change in the perceived wage gap between two jobs changes the relative likelihood of applying by 4.2%. The validity of this 2SLS estimate requires an exclusion restriction that our information treatment only affects applications by changing beliefs about potential wages. Importantly, we are able to substantiate this assumption using the survey: across a range of post-treatment questions, we see no evidence that the information treatment affected beliefs about the likelihood of getting the job if applying or about other dimensions of jobs' attractiveness.

Using individually-linked administrative data, we verify that the estimated treatment effects on planned application behavior also translate into actual changes in behavior and job outcomes. For the majority of job seekers, who tend to underestimate true wage gaps, misperceptions make lower-paying jobs appear relatively more attractive. The information treatment should shift such job seekers to apply for and get hired into higher-paying jobs, while the reverse should be true for job seekers who overestimate true gaps. We verify these predictions in actual job outcomes post-survey. Among the two-thirds of job seekers who are underestimators, the information treatment significantly increases the actual wages earned in the first job post-survey. Conversely, for overestimators, we see evidence of a negative effect. As should be expected, these wage effects are driven by the job seekers who are still currently searching at the time of the survey and can therefore adjust their behavior to new information without incurring additional switching or search costs.

Finally, we provide a quantification of the overall costs of relative wage misperceptions among early career workers. To this aim, our survey and reduced form analyses were deliberately set up to identify a simple discrete choice model of job search and hiring. We use this to compute counterfactual application and hiring outcomes in a situation without relative wage misperceptions. We find that existing relative wage misperceptions distort one-in-ten job applications and have substantial implications for job outcomes. Removing relative wage misperceptions would reallocate 9% of workers to different jobs and increase wages by 1.2% on average. Mirroring conclusions in the reduced form analysis, however, we find substantial heterogeneity: for 10% of job seekers, the removal of wage misperceptions would lead to wage *declines* of more than 0.2%, while another 10% would experience wage *increases* above 3.9%. Further exploiting the model structure, we estimate that removing relative wage misperceptions would deliver the same average welfare gains for workers as an across-the-board wage increase of 0.5-0.8% in all jobs.

This paper is connected to several strands of literature. A number of papers find that misperceptions about the rate of job finding or the overall level of reemployment wages have implications for how long workers remain unemployed (e.g. Krueger and Mueller (2016), Spinnewijn (2015), Mueller et al. (2021), Belot et al. (2019),Belot et al. (2022a), Behaghel et al. (2024), Altmann et al. (2018), Ben Dhia et al. (2022), Altmann et al. (2022), Harmon et al. (2024), Altmann et al. (2025)). The focus of our analysis is fundamentally different from these previous papers as we examine misperceptions about relative wages across different jobs. Our results show that such relative wage misperceptions generate costs beyond excessive unemployment duration because they distort the allocation of workers to jobs.

Our results are especially connected to recent studies of misperceptions about wage dispersion and their effects on job transitions and wages. In some models, these misperceptions reduce worker mobility, increase firm monopsony power and push wages below competitive levels (Manning (2003)). Recent empirical results are mixed. Jäger et al.

(2024) find that employed workers systematically underestimate outside wages relative to their current wage, and that this substantially lowers job mobility. Caldwell et al. (2025) instead find that workers are relatively well-informed about cross-firm wage differences but that high switching costs restrict mobility and make outside wage information less relevant. By studying early career job seekers—most of whom do not currently have a job—we focus on an important population who face negligible switching costs and have no current wage to benchmark against. Consistent with Caldwell et al. (2025), we find that workers are broadly aware of relative wage differences across jobs in this setting; beliefs about wage gaps across jobs are clearly correlated with actual gaps, although far from perfectly so. As in Jäger et al. (2024) however, we find that misperceptions are prevalent and impactful enough to distort job search and mobility. Our results thus confirm relative wage misperceptions as a potential source of market power for firms.

Third, we relate to a literature studying how increases in offered wages affect workers' propensity to apply for a job, i.e., the wage elasticity of job applications. Prior studies analyze variation in wages in online job vacancies and find positive but small elasticities—typically below 1 (Holzer et al., 1991; Dal Bó et al., 2013; Banfi and Villena-Roldán, 2019; Marinescu and Wolthoff, 2020; Dube et al., 2020; Belot et al., 2022b; Azar et al., 2022).[2] If generalizable, these results suggest very high monopsony power for firms. An explanation for these low elasticities may be that job seekers interpret wage premiums as signals that jobs are harder to get or that they offer lower non-wage amenities. This inference may be especially common when people apply online to individual job postings, as they have little information beyond the ad itself. Since we measure beliefs, we are able to examine such belief-updating directly. In the context of our information treatment, beliefs about the probability of being hired or about non-wage amenities are inelastic to changes in wage beliefs. Accordingly, translating our 2SLS estimates to a wage elasticity of applications suggests a markedly larger elasticity of around 3.

Finally, our paper is connected to the literature studying the determinants of college major choices (e.g., Betts (1996); Arcidiacono et al. (2012); Wiswall and Zafar (2015b,a);

[2]The estimated firm-level application elasticity in Azar et al. (2022) varies widely across specifications, being around 2.4 in their preferred specification.

Conlon and Patel (2025)). We complement this literature by focusing on the choice of career path once education is completed which may also have persistent consequences on the rest of the career (e.g, von Wachter and Bender (2006), Arellano-Bover (2024)). While causal effects of expected wages in different majors seem to have limited effects on major choice (Hastings et al., 2016; Ballarino et al., 2022; Bonilla-Mejía et al., 2019; Kerr et al., 2020), our findings suggest that career choices at the job entry stage are more elastic to job-specific wage expectations. Methodologically, the link between our information treatment and discrete choice model is also particularly closely related to the study of major choice in Wiswall and Zafar (2015a), although our specific model builds instead on Harmon et al. (2025).

The rest of the paper is organized as follows. Section 2 discusses the survey design and data. Section 3 provides novel descriptive evidence about misperceptions of wage differences across jobs. Section 4 establishes the causal effects of these misperceptions on job search. Section 5 quantifies aggregate implications using a discrete choice model. Section 6 concludes.

2 Data and survey design

To be able to measure relative wage misperceptions, we fielded a survey with early career job seekers in Denmark in the summer of 2023. Below, we first describe our setting and the linked administrative data that we use. We then describe the survey data in detail. At the end, we discuss the reliability of our survey measures and present a range of validity checks.

2.1 Empirical setting

Our analysis focuses on graduates from higher education in Denmark. We focus on graduates early in their career since we expect relative wage misperception to be particularly prevalent and consequential when workers have limited labor market experience and may face persistent effects from early job choices.

Our focus on Denmark allows us to leverage rich administrative registers in both the design, sampling and analysis of the survey. A few other features of the Danish setting are important as well. While most jobs in Denmark are covered by collective bargaining agreements, wage setting is highly decentralized, thus generating substantial wage dispersion across jobs (Dahl et al., 2013). Moreover, higher education degrees in Denmark typically qualify candidates for a number of distinct careers, meaning graduates face a choice between a set of different jobs and careers, often offering very different terms and wages. Whether workers correctly perceive these wage differences across jobs is an empirical question that we tackle in this paper. On the one hand, workers have limited experience with wages early in their career and — as is the case in many countries — direct information about wages is rarely included in job postings. On the other hand, there are many sources of wage information that active job seekers could rely on in the absence of information costs or behavioral biases.[3]

Finally, an important feature of the Danish setting is that graduates are eligible to claim public unemployment benefits immediately upon graduation. As we returnt tobelow, high-frequency claims data thus helps us to sample and survey workers right at the start of their post-graduation job search.

2.2 Administrative data

We use administrative data for two main purposes in this project. First, we use it in the design of our survey to identify which types of jobs are relevant for graduates from different degrees, and to construct ground truth measures of the typical wages that these jobs have paid in the past. This is based on the transitions of higher-education graduates into first jobs which we observe using data on individual labor market outcomes (BFL) and educations (UDDA) for individuals who graduated in 2010-2018.

Second, we link our survey sample to administrative data to obtain background characteristics and post-survey labor market outcomes. Specifically, we link the Danish registers on individual labor market outcomes (BFL), educations (UDDA), government transfers (DREAM) up to December 2024.

[3]For example, in addition to wage information from their own social and professional connections, many Danish worker unions publicize information about typical wages in various categories of jobs.

2.3 Definition of survey job types

A strength of our research design is that we analyze graduates not from a single university or degree but from a *wide range* of educations. Since different jobs are relevant for graduates from different educations, however, this requires us to think carefully about the specific jobs we ask respondents about when eliciting beliefs about wage differences.

To avoid survey fatigue, we ask each job seeker about three specific job types which we design to achieve the following objectives: First, the job types must be relevant and capture a large share of the respondent's effective labor market. Second, the three job types should differ in terms of wages so that choosing among them represents a meaningful trade-off. Finally, in order to use administrative data as an objective benchmark for jobs' typical wage levels, the job types must be defined in a way that can be mapped one-to-one to administrative data.

To achieve these objectives we vary the jobs respondents are asked about depending on the respondents education (measured using the 6-digit Danish ISCED codes) and define the three job types by grouping jobs based on some combination of occupation, industry, firm sector and firm size. We implement a data-driven procedure to select the appropriate job groupings using administrative data on the first jobs of graduates from the same education between 2010-2018: For each education, we consider all possible ways of grouping jobs by firm sector and size (public/private, above or below 50 employees) occupations codes (using either 3, 4, or 6 digit Danish ISCO codes), and/or industry codes (using either 1- or 2-digit Danish NACE codes).[4] Among all such possible groupings, we then select those that satisfy two conditions: i) the 3 most common job types must cover more than 40% of the transitions into first jobs for prior graduates and ii) the 3rd most common job type must cover at least 5% of transitions. These conditions ensure that the three most common job types altogether cover a substantial share of relevant jobs and that none of them are too narrow. If multiple groupings satisfy the two conditions, we choose the grouping that best predicts entry-level wages.[5]

[4]Since the concept of firm size is not well defined for public sector jobs, we only allow groupings on firm size when also grouping into private sector jobs.

[5]For each possible grouping into three jobs, we regress log wages on dummies for the three jobs and calculate the mean squared error of the regression predictions (using 5-fold cross validation). We then

After hand-checking the resulting job types for each education and making marginal adjustments, this procedure gives us three job types based on basic categories that we can intuitively describe to each survey respondents.[6] For respondents with a Master in Economics, for example, the selected grouping is based on sector and 1-digit industries, with the three most common job types being "Private sector jobs, in banking, financial, or insurance activities", "Public sector jobs, in public administration, defense or social security activities" and "Private sector jobs, in professional, scientific or technical activities". Respondents with a Master in Economics were thus asked about these three job types throughout the survey.

For very specialized education programs, where all graduates take similar jobs, the procedure above does not deliver any job types because no detailed grouping satisfies conditions i) and ii). This reflects the fact that graduates in these programs de facto make their career-defining decisions before entering the labor market and thus are not affected by the potential wage misperceptions we study in this paper. Accordingly we exclude these educations from our analysis. To avoid asking only abroad broad sectors, we also exclude educations where only coarse grouping into public vs. private sector jobs satisfies conditions i) and ii). Finally, we exclude education programs where we observe very few graduates in past administrative data, since we cannot calculate precise measures of typical wages for those.[7] These restrictions lead us to focus on 88 higher-education programs, covering 76% of higher-education graduates in 2010-2018 (see Appendix Table A2). We refer to these as the "selected" educations.[8]

2.4 Sampling procedure

To examine the implications of relative wage misperceptions for job search, it is important to survey workers immediately around the time when they actively search. The extent of misperceptions may be overstated otherwise if workers only seek out detailed wage

choose the grouping with the lowest mean squared error.

[6]Appendix section C.3.1 describes the few manual changes we made to labels of occupations and industries to ensure clarity.

[7]We exclude educations where fewer than 50 people transitioned to a job type in 2010-2018.

[8]Note that individuals from non-selected educations could participate in the survey but were ineligible for the randomized information treatment and are excluded from our analysis.

information when they need it.

Using two sampling procedures, our survey targeted individuals who are either looking for their first job just after graduation, or who enter unemployed job search within a few years following their graduation. First, we collaborated with the Danish Agency for Labor Market and Recruitment (STAR) to contact all individuals below age 40 who initiated a claim for unemployment benefits during the summer of 2023. Since most education programs end in the summer and many graduates immediately claim benefits, sampling during these months allows us to reach a large fraction of new graduates.[9]

Second, we collaborated with the University of Copenhagen (UCPH) to contact all UCPH students about to finish their Master's degree in the summer of 2023. We invited them into the survey in mid-June 2023, i.e. right before typical Master graduation dates. This second sampling procedure allows us to also survey graduates who enter the labor market without going through an unemployment spell.

Sampled individuals were invited to the survey in three waves via the so-called "e-Boks"— the official channel through which all government agencies in Denmark communicate with citizens.[10] The invitation contained an individualized link to the online survey and included a monetary incentive to participate.[11]

Appendix Table A1 provides descriptive statistics for our sample of invitees and participants. In total 22,138 people from our selected educations were invited to participate in our survey (Column (2)) and 1,902 completed the survey and are included in our study sample (Column (1)). Using administrative data to compare the observables of all invited individuals to the study sample, Appendix Table A1 show them to be similar in most dimensions. Individuals in the study sample tend to have higher educational attainment (64% have completed a Master's degree) than the targeted population (45% with a Master's degree)—a common phenomenon in online surveys (Haaland et al., 2023). Moreover, survey participants are somewhat more likely to have graduated less than a year before the

[9] Økonomi og indenrigsministeriet (2018) report that about half of Danish graduates receive unemployment benefits within 6 months of graduation.

[10] The first wave in mid-June included all individuals sampled from UCPH, as well as new unemployment benefit claimants from the past three weeks. The second and third wave were conducted three and eight weeks later and each included only benefit claimants from the preceeding three weeks.

[11] Among everyone who finished the survey, we raffled 20 gift cards worth DKK1,000 (approx. USD160) each.

survey (66%, compared to 59% in the targeted population). A similar picture emerges if we compare our study sample to administrative data on *all* young job seekers, defined as graduate from the selected educatiosn who are below 40 and have been unemployed and/or graduated from a degree in 2023 (Column (3) of Appendix Table A1).[12] Accordingly, we see little reason to expect that selection into the survey distorts our conclusions. As we show in Table 3 and Appendix Table A5, our conclusions are also unchanged if we reweigh the study sample to correct for non-participation.

2.5 Survey questions

This section describes the key elements of the survey. Additional details, including the full survey instructions, are provided in Appendix C.

Belief about prior-cohort average wages The first part of the survey aims to measure how informed respondents are about typical wage differences across their possible career paths. Chiefly, we elicit respondents' belief about the entry-level *average wage received by prior cohorts* in the three different job types. For each job type, we ask respondents to consider people who graduated from the same education program as themselves over the years 2010-2018 and who were hired into a full time job of the given type. We then ask them to report the average pretax monthly wages over the first year of employment for these individuals and jobs. To further minimize misreporting errors from misunderstandings about the wage concept we are asking about, we show respondents the average wage received by prior cohorts with *any* higher education diploma and in *any* job type. Obviously, this benchmark contains no information about wage differences across jobs for any specific education groups but helps respondents in understanding the wage concept (e.g. monthly vs. hourly/yearly wages) and, more generally, gives them a sense of the scale on which wages are measured, thereby reducing potential measurement error in beliefs (Ansolabehere et al., 2013). To incentivize effort and truthful reporting, we provide a monetary incentive

[12] Again our sample of respondents tend to have obtained higher degrees (64% versus 36% got a Master's degree), graduated more recently (66% graduated less than one year before the survey versus 53%), and been exposed to unemployment more (90% versus 70% received UI in 2023). This is in line with expectations given our sampling procedure

to get close to the true value of the average wages computed in administrative data: the 3 closest answers received a 1,000 DKK gift-card.

In a similar way, we also elicit respondents beliefs about two other statistics for prior cohorts: the average wage earned 5 years after starting in each job type, and the rate of successful applications sent to each job type. Along with a range of socioeconomic questions at the start of the survey, these beliefs elicited before the information treatment serve as predetermined covariates in various parts of our analysis (cf. Table 1).

Randomized information about prior-cohort average wages Next, the survey includes a randomly assigned information treatment about average prior-cohort wages in each of the three job types. This information is illustrated by a figure listing the true average prior-cohort wages in the three job types alongside the respondent's own answers. The control group is shown a corresponding figure only containing their own answers (see Appendix Figures C1- C2).

Planned search behavior and beliefs about own potential wages After the information treatment, we ask respondents to think of three specific jobs that they view as representative of each of the three job types asked about in the survey. We then elicit their planned search behavior and beliefs about what they themselves expect to earn in these jobs as described below.

To measure search behavior, we ask respondents to consider a situation where they have found job postings for each of three jobs and are able to apply to at most one of them. We then ask them to report the likelihood that they would end up applying to each job posting or to none of them. Answers are reported as probabilities, recorded using three sliders that automatically adjust to make the four probabilities sum to 100%.

Eliciting self-assessed probabilities instead of a single deterministic choice offers two advantages here (Wiswall and Zafar, 2018). First, it increases the information content of the data because respondents' answers express the strength of their preferences for applying to different jobs rather than just their highest ranked option. Second, it implies that the respondents' answers can be mapped directly to a standard discrete choice framework with decision uncertainty. We use this to compute counterfactuals in Section 5. As an alternative measure of job attractiveness, respondents are also presented with a scenario in which they simultaneously received mutually exclusive job offers from all three jobs and are then asked to report the likelihood of accepting each offer.

Job seekers' application decisions should depend on the potential wage they expect to face in each job type. To elicit beliefs about such *own potential wages*, we next ask respondents to report what starting monthly pretax wage they expect they would earn if hired into each of the three jobs.

Finally, we ask a range of questions about other job characteristics the job seeker expected to face in each of the three jobs. These include work hours, earnings 5 years in the future, the probability of receiving an offer if applying, and how well they expected to perform and get along with colleagues (both on 6-item Likert scales). As a summary measure of the perceived attractiveness of non-wage amenities in the three jobs, respondents are also presented with a scenario in which they simultaneously received mutually exclusive job offers from all three jobs but where the offers are atypical in that they all involved the same wage. As above, job seekers are then asked to report the likelihood of accepting each job offer. As we expand on in Section 4.3, we use these additional questions to assess belief spillovers from the information treatment.

Winsorization and outliers As always, typed numerical survey responses are prone to generating outliers. For the preferred specifications, we winsorize all such variables at the 2.5th and 97.5th percentile prior to our analyses. This includes survey-reported wages, hours and application success probabilities.

2.6 Reliability and validation of survey data

Our analyses rest crucially on our survey questions being meaningful for respondents and eliciting reliable measures of respondents' beliefs and behavior. We discuss this below and provide a range of validation exercises.

Understanding of job types A crucial question for our survey design is whether respondents properly understood the three job types they were asked about in the survey. We validate this in two ways. First, we directly asked respondents how well they understood the job types: 75% of respondents reported that their understanding of job types was good or very good (Appendix Figure A1). Second, we tested their understanding at the end of the survey by asking them to place a more concrete example job into one of the three types: 76% of the respondents answered correctly (Appendix Figure A2).[13] As we show later, results are also robust to excluding respondents showing poor understanding in these questions (Table 3 and Appendix Table A5).

Job types' relevance and wage differences Another premise for our analysis is that the three survey job types are relevant to respondents and represent a trade-off in terms of wages the wage levels they offer. In Table 2 we provide detailed descriptives regarding the survey job types based on both administrative data and survey responses. Columns correspond to different variables and data sources, while panels correspond to different ways of grouping the job types across respondents.

In Panel A, we group job types according to the share of prior cohorts who started their first job within each type. As expected, the job types cover a substantial share of starting jobs: the most common job type on average covers 26% of starting jobs for prior cohorts, the second most common 15% and the least common one 10% (Column (1)). In Panel B, we instead group jobs according to each respondents' stated likelihood of applying to the jobs in our application question. Answers confirm that the job types capture relevant employment options considered by job seekers. The reported likelihood of applying for their most popular job is 62% on average, suggesting that most job seekers have a clear leaning towards one of the job types (Column (2)). The likelihood of applying for their second and third most popular job types are 22% and 7%, however, so job seekers still have substantial uncertainty about their choice across the three jobs. The generally high reported likelihoods of applying also underscore that all three job types were perceived as relevant by respondents.[14] Finally, in Panel C we group job types according to the average

[13] For the details on the creation of the examples see Appendix Section C.3.2

[14] Recall that our job application question included the option of not applying to any of the three job

wage paid to prior cohorts and see substantial differences. For the average responedent, typical wages in the highest-paying job are 10% higher than in the lowest paying one (Column (3)).

Prior cohort wages versus own potential wages Our survey elicits respondents' beliefs about two distinct wage concepts for each of the three job types: average wages for prior cohorts (prior to the information treatment) and respondents' own potential wages (after the information treatment). We focus on prior cohort wages because they constitute an objective benchmark of typical wages in the different jobs which we can compute from administrative data. Contrasting administrative data with elicited beliefs about prior cohorts therefore allows us to construct an individual-level measure of wage misperceptions which is not confounded by job seekers having correct, private information about their own labor market opportunities. [15] The administrative data on prior cohorts also allows us to provide objective and correct wage information in our treatment.

Importantly, because we focus on early career graduates, we expect prior cohort wages to be a very salient benchmark for the wage differences that job seekers themselves expect to face across the different jobs. As we expand on in Sections 3.2 and 4, this is borne out by the data; beliefs about prior cohort and own potential wages are strongly correlated in the cross-section, and receiving information about prior cohort wages causes job seekers to update beliefs about their own potential wages.

Planned applications versus real transitions Whenever survey respondents report their planned behavior, a critical question is how likely it is to reflect their actual future behavior. With the linked administrative, we can validate our survey measure of planned job applications against data on respondents' actual job transitions post-survey. We present these validations in Appendix Figure A3 and Appendix Table A3 and find a strong association. Comparing job seekers from the same educational background, a 1 percentage point

types. For the average respondent, the reported likelihood of choosing this option was only 8 %.

[15]Measuring misperceptions using elicited beliefs about own potential wages raises a fundamental identification problem: the actual wage a job seeker *would* earn in a given job is only observed if the job seeker in fact ends up in this job. Trying to infer these unobserved actual wages from data on other workers will lead to erroneous measures of mispreceptions if job seekers have correct private information about how their wage prospects differ from other workers.

increase in the reported likelihood of applying to some job types is associated with a 0.25 percentage point increase in the likelihood of being hired into this job type later. Conditional on being hired into one of the three job types covered by the survey, this association increases further to 0.66. In Section 4.4 later, we show that the estimated effects of our information treatment on planned applications also translate to changes in actual hiring outcomes in administrative data.

Covariates and timing of survey responses Table 1 shows descriptive statistics for all pretreatment covariates elicited in the survey, both for the full study sample and for the treatment and control group separately. Comparing treatment and control groups, treatment randomization appears to have been successful; across the 30 pre-determined covariates, we find the statistically significant treatment-control differences for two variables at the 10% level. As shown later, all results are also robust to the inclusion of pretreatment controls. A key objective of our survey design was to measure wage beliefs around the time of active job search. Table 1 confirms that we were successful in reaching respondents at this time. 65% of respondents repor actively searching at the time of the survey and another 24% report having just secured a new job or being about to start job search. In total 89% of respondents were thus either actively searching, had recently finished, or were about to start searching. As we show later, all results are also robust if we restrict the data to only include respondents who report actively searching at the time of the survey.

3 Measuring relative wage misperceptions

In this section, we provide descriptive evidence on the extent and nature of wage misperceptions. We do this by contrasting workers' beliefs about the wages of prior cohorts in different jobs with administrative data on the true wages earned by these cohorts.

3.1 Misperception about wage levels in individual jobs

While the focus of paper is on misperceptions about relative wages across jobs, we begin by briefly considering general misperceptions about the level of wages. The analysis sample consists of 1,902 job seekers indexed by i, each of which answered questions about three different jobs indexed by j. For each individual-by-job, i, j, our survey data contains the beliefs about the average wage earned by prior cohorts, $\widetilde{V_{i,j}}$ which we compare to the true value $V_{i,j}$ computed from administrative data. In addition to the convention that tilde ($\tilde{}$) denotes beliefs, we use lowercase letters to denote logged variables throughout the paper (e.g. $\widetilde{v_{i,j}}, v_{i,j}$).

For each job seeker and associated job, Figure 1 compares respondents' beliefs about the average wages earned by prior cohorts to the true values. Panel (a) shows a binscatter plotting the perceived log wage ($\widetilde{v_{i,j}}$) against the truth ($v_{i,j}$). There is a clear positive relationship but it is not one-to-one; the corresponding regression slope is 0.65. This is consistent with recent evidence in Caldwell et al. (2025): people's beliefs are highly but not perfectly correlated with true wages. Second, Panel (b) shows the full distribution of wage misperceptions, measured as the difference between the true and perceived wage in percent of the true wage, i.e., $\left(V_{i,j} - \widetilde{V_{i,j}}\right)/V_{i,j}$. Misperceptions about wages are substantial and highly heterogeneous. For some individuals and jobs, wages are overestimated ($V_{i,j} - \widetilde{V_{i,j}} < 0$), while for other they are underestimated ($V_{i,j} - \widetilde{V_{i,j}} > 0$). Wage overestimation dominates overall however. In 60% of cases wages are overestimated and the mean overestimation rate is 3.3 %. This is in line with findings from the literature on misperceptions about general wage levels. Altmann et al. (2025) find that unemployed Danish job seekers overestimate re-employment wage of comparable workers by about 2% on average.

3.2 Relative wage misperceptions

Next, we turn to the key focus of our paper, misperceptions about wage *differences* across jobs. When job seekers make decisions about *which* jobs to target in their search, it is these *relative* wage misperceptions that should matter. At the same time, results about wage level misperceptions (such as those in Figure 1) say little about relative misperceptions. If a worker overestimates wages in all jobs to the same extent for example, misperceptions about wage levels can be large but there are no misperceptions about relative wages.[16]

[16]Conversely, if a given worker overestimates the wage in low paying jobs and underestimates wages in high paying jobs, misperceptions about wage gaps can be large even when misperceptions about wage

To measure relative wage misperceptions, we focus on misperceptions about wage gaps between pairs of jobs. For each job seeker, we form all possible pairwise comparisons between the three jobs they have been asked about, resulting in dyadic data where observations are individuals-by-job pairs, i, j, j'. Letting Δ denote differences in variables across jobs, our survey data allows to compute job seekers' perceived (log) wage gap between jobs j and j' for prior cohorts, either in levels, $\Delta\widetilde{V_{i,j,j'}} = \widetilde{V_{i,j}} - \widetilde{V_{i,j'}}$, or in logs, $\Delta\widetilde{v_{i,j,j'}} = \widetilde{v_{i,j}} - \widetilde{v_{i,j'}}$. We compare these to the true gaps from administrative data, $\Delta V_{i,j,j'} = V_{i,j} - V_{i,j'}$ and $\Delta v_{i,j,j'} = v_{i,j} - v_{i,j'}$.

Figure 2 compares respondents' beliefs about prior cohort wage gaps across jobs to the true prior cohort gaps. Panel (a) shows a binscatter plotting perceived gaps in log wages ($\Delta\widetilde{v_{i,j,j'}}$) against true gaps ($\Delta v_{i,j,j'}$). The binscatter is based on respondents-by-job pairs (i, j, j') ordered so that $\Delta\widetilde{v_{i,j,j'}}$ is positive. Again, we see that beliefs are strongly but not perfectly correlated with the truth; the corresponding regression slope is 0.52.

Panel (b) characterizes the full distribution of relative wage misperceptions. While we focus on log wage gaps both in Panel (a) and the regression analyses presented later, our preferred measure when describing the distribution of misperceptions is the *underestimation rate*, i.e., the difference between the true and perceived gap in percent of the true gap:

$$\frac{\Delta V_{i,j,j'} - \Delta\widetilde{V_{i,j,j'}}}{\Delta V_{i,j,j'}} \tag{1}$$

The advantage of this measure is that it succinctly captures several different qualitative features of misperceptions: An underestimation rate of 0 implies correct beliefs. An underestimation rate between 0 and 100% means that the size of the true gap is being understated. An underestimation rate of 100% implies that the jobs are perceived to pay exactly the same, while an underestimation rate above 100% implies getting the relative wage ranking wrong. Finally, a negative underestimation rate means that the size of the wage gap is in fact being *over*estimated. Panel (b) shows the distribution of the underestimation rate across all individuals and job pairs.[17]

levels are modest.

[17]Appendix Figure A4 instead shows the raw distribution for the underestimation of the log wage gap, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$, Motivated by theory, this is the misperception measure emphasized in the regression

In green, we present as a benchmark the CDF that we would observe under perfect information. In blue, the actual CDF highlights that in about 65% of cases, the wage gap is underestimated. Further, we see that 55% of wage gaps are underestimated by more than 50%, while 36% are underestimated by strictly more than 100%—meaning that respondents actually rank the two jobs wrong. We also see that 12% of wage gaps are underestimated by exactly 100%. This corresponds to situations where respondents attributed the exact same wage to two job types even though the jobs in fact paid different wages. Finally, we also see that some wage gaps are instead overestimated, sometimes by a large amount. 25% of wage gaps are overestimated by more than 50% and 20% by more than 100%. Overall, in less than 20% of cases, beliefs fall within 50% of the true wage gap.

In Table 3, we show additional descriptive statistics regarding relative wage misperceptions using different measures and different subsamples of our data. The first row corresponds to the full sample and Columns (1)-(4) simply summarizes key statistics regarding the wage gap underestimation rate. In Columns (5) and (6), we consider a different measure: the underestimation of the log wage gap, $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$. Motivated by theory, this is the main misperception measure used in the regression analyses of Section 4 later.[18] Unsurprisingly, we also find large misperceptions using this measure; its mean absolute value is 0.13 log points implying that on average job seekers beliefs about wage gaps are off by 13%. Appendix Figure A4 shows the distribution of this misperception measure across the sample of all job pairs.

In Column (7) we show that the pattern of wage gaps being mostly underestimated holds also if we look across individuals instead of job pairs. Classifying each job seeker as either an *overestimator* or *underestimator* depending on whether they overestimate or underestimate the gap between the highest paying job and the average of the two other jobs, we find that 67% of job seekers in our sample are underestimators.

Heterogeneity and robustness To substantiate that the measured misperceptions are meaningful in practice, the additional rows of Table 3, shows how the distribution of relative

analysis of Section 4.

[18]As we return in Section 5, our reduced form specifications involving log wage gaps can be derived from a simple discrete choice job search model.

wage misperceptions is affected by various sample restrictions and changes in variable definitions.

A possible explanation for the substantial misperception we see is that some true wage differences may be relatively small; if tracking wage differences is somehow costly for job seekers, theories of rational inattention suggest that they may optimally decide to remain poorly uninformed.[19] To test this, we restrict our sample to only consider job pairs where the true wage gaps for prior cohorts are above 5% or 10% respectively. Consistent with rational inattention, we find evidence that the relative misperceptions are indeed smaller for job pairs with larger gaps. We continue to see very substantial misperceptions even in these samples however. Even restricted to job pairs where the true gap is above 10%, we find that 51% are underestimated by more than half.

The remaining rows dispel a number of other concerns. Measured misperceptions are not driven by rare jobs or by jobs that job seekers do not view as relevant; we see similar levels of misperceptions if we restrict to the two job where prior cohorts were most frequtly hired, or if we restrict to job types where respondents have a high reported likelihood of applying in the control group.[20] Applying a simple procedure to correct obvious reporting errors suggests that such errors do not explain the measured misperceptions.[21] Restricting only to job seekers actively searching at the time of answering the survey, we see that misperceptions are also not driven by possible inattention among respondents who have already completed their job search or among graduates who have yet to start searching actively. Finally, we see no evidence that measured misperceptions are affected by survey non-particpation or by respondents not understanding the job types in the survey.

A final concern is that relative misperceptions about prior cohort wages could be large simply because they are unrelated to job seekers' search decisions and the wages they themselves expect face. Using our survey data on own potential wages for the control group Figure 3 plots beliefs about the gap in log wages that job seekers *themselves* expect to face across two jobs, $\Delta\widetilde{w_{i,j,j'}}$, against their belief about the log gap in prior cohort wages

[19]For jobs where the true difference in typical wages is small, we might worry that measured misperceptions stem simply from rounding in responses.

[20]Since the likelihood of applying is elicited after the information treatment, we limit to the control group when imposing sample restrictions based on application likelihoods.

[21]Appendix Section D.1 explains the procedure we use to correct errors

for these jobs, $\Delta\widetilde{v_{i,j,j'}}$. The plot confirms prior cohort wages as a relevant benchmark for job seekers; there is a clear positive relationship with an estimated regression slope of 0.63. As we return to in Section 4 below, receiving information about wages of prior cohorts also leads job seekers to update their beliefs about the potential wages they will face.

Summary of descriptive results Summing up, we find that young job seekers have substantial misperceptions about wage differences across jobs that they consider in their job search. Most commonly, misperceptions take the form of underestimating true wage gaps, although overestimation of wage gaps occurs about a third of the time. An implication of this is that relative wage misperceptions cause lower-paying jobs to appear overly attractive for most job seekers but that the opposite also occurs. Over the next sections, we examine the extent to which these misperceptions distort search behavior and job outcomes.

4 Effects of wage gap misperceptions on job search

In this section, we examine to what extent relative wage misperceptions causally affect job search behavior. We do this by leveraging exogenous variation generated by our information treatment: treated job seekers were shown actual wages for prior cohorts in the different jobs which should generate variation in perceptions about what they themselves would expect to receive in each job.

We proceed in three steps. First we show nonparametric treatment-control comparisons to establish the existence of a causal effect. Second, we use a 2SLS framework to collapse these into an overall elasticity that captures the responsiveness of job search decisions to perceived relative wages. Throughout these analyses, we use the same individual-by-job pair survey data as in Section 3.2. Finally, we use individual-level administrative data to validate the effects of the treatment on post-survey hiring outcomes.

4.1 Treatment-control comparisons

Our treatment provided information about actual wages for prior cohorts in the different jobs, and therefore about the actual *wage gaps* in prior cohort between each job pair (j, j').

The expected effect therefore depends on job seekers' baseline misperception about this wage gap. We measure baseline misperceptions of the wage gap between job j and j', as the difference between the actual log wage gap and the perceived log wage gap for prior cohorts (elicited *before* the information treatment), $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$. If this misperception measure is positive, the information treatment reveals to the job seeker that they underestimated the wage gap between j and j' for prior cohorts. In response, we expect the job seeker to upward-adjust the potential wage they expect to receive in job j relative to job j'. Assuming that beliefs about wage gaps affect job search, we also expect them to become more likely to apply for job j relative to job j'.

Column (1) of Table 4 confirms these predictions using simple treatment-control comparisons. Reordering the data so that all included pairs have $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}} \geq 0$, we regress our outcomes of interest on a treatment dummy, T_i.[22] Since treatment was assigned at the individual level, we use cluster-robust inference clustered on individuals throughout all analyses. In Panel A, the outcome is the beliefs about the gap in (log) potential wages that the respondent expects to face, $\Delta \widetilde{w_{i,j,j'}}$ (elicited *after* the treatment). We see a significant effect: informing job seekers that they underestimated prior cohort wages in job j relative to job j' increases the potential wage they themselves expect to face in j relative to j' by 4.2 % on average. In Panel B the outcome is the gap in the (log) reported likelihood of applying to the two jobs which we denote by $\Delta \pi_{i,j,j'} = \pi_{i,j} - \pi_{i,j'}$. We see again a significant effect: informing job seekers that they underestimated the prior cohort wages in job j relative to job j increases their relative likelihood of applying to job j by 19.5% on average. As robustness checks, Columns (2)-(4) repeats the analyses after controlling for all predetermined survey responses and/or restricting only to respondents who reporting being actively searching at the time of the survey. Results are virtually identical to those in Column (1). Finally, as shown in Appendix Table A4 we find similar positive treatment effects if we use the likelihood of accepting job offers as the outcome instead of the likelihood of applying.

[22]Note that this sample definition does not mean that we exclude any unique job pairs from the analysis. For any individual, and given pair of jobs A and B, either the wage in A is underestimated relative to B, or the wage in B is underestimated relative to A (no respondents are exactly correct about wage gaps in our survey).

Heterogeneity by baseline misperceptions In addition to the treatment effect depending on the *sign* of $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$, we of course also expect it to depend on its *magnitude*; we expect a bigger effect if the treatment reveals that the job seeker underestimated the wage gap for prior cohorts by a large amount. Starting from the sample of all possible individual-by-job pairs, Figure 4 collects observation in seven equal width-bins based on the value of $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$ and estimate treatment effects within each bin using simple treatment-control comparisons.[23] Panel A of Figure 4 shows effects on the perceived gap in potential wages, while Panel B shows effects on the relative application likelihood. While splitting into subsamples makes results for the application likelihood notably less precise, estimates in both panels of Figure 4 show the expected pattern. For individuals and job comparisons where the information treatment revealed larger underestimation of the prior cohort wage gap, we estimate larger upward-adjustments of the perceived gap in potential wages, as well as larger increases in the relative likelihood of applying. Mechanically, the pattern also extends to negative values of $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$; informing job seekers that they *overestimated* the wage gap for prior cohort leads to downwards adjustments. [24]

4.2 The effect of misperceptions on search: 2SLS estimates

The results above confirm that our information treatment shifted relative wage perceptions and job application behavior as expected. To collapse these results into an overall estimate of *how much* wage perceptions affect job applications, we next turn to a parametric 2SLS framework.

First stage and reduced form To setup and relate our 2SLS framework to the results above, we first consider the corresponding first stage and reduced form. The results in Figure 4 indicate that the effect of the information treatment is approximately linear in $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$. This suggests the following two parsimonious regression specifications

[23] In the appendix we show results for various alternative choices of bins in Figures A5-A7

[24] By definition, if our treatment reveals to a job seeker that they underestimated the wage gap between job A and B by some amount, it of course also reveals that they overestimated the wage gap between B and A by this same amount. This is mirrored in the fact that our data set of all possible pairwise job comparisons mechanically has $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}} = -\big(\Delta v_{i,j',j} - \Delta\widetilde{v_{i,j',j}}\big)$. Accordingly, if the chosen bins in Figure 4 were symmetric around 0, the magnitude of the estimated treatment effects on either side of zero would in fact be numerically the same (see Appendix Figures A6 and A7).

which will correspond to the first stage and reduced form of our analysis:

$$\Delta\widetilde{w_{i,j,j'}} = \gamma_1 \, T_i \times (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) + \gamma_2 \, (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) + e_{i,j,j'} \tag{2}$$

$$\Delta\pi_{i,j,j'} = \theta_1 \, T_i \times (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) + \theta_2 \, (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) + u_{i,j,j'} \tag{3}$$

The coefficients on the interaction terms, γ_1 and θ_1, capture the (linear) effects of the information treatment and have a simple interpretation: if the information treatment reveals to a job seeker that they underestimated a particular prior cohort wage gap by $X\%$, equations (2) and (3) imply that the job seeker's perceived gap in potential wages increases by $\gamma_1 X\%$ and that their relative likelihood of applying changes by $\theta_1 X\%$ percent. Since we control for the underestimation of the prior cohort wage gap, $(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})$, identification of γ_1 and θ_1 comes from treatment-control comparisons *within* groups of individuals and job pairs with different levels of $(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})$ (as in Figure 4). Random assignment therefore ensures that γ_1 and θ_1 have a causal interpretation. Finally, note that equations (2) and (4) do not include constant terms (γ_0, θ_0) or an uninteracted treatment dummy (T_i). This is without loss of generality because all included variables are mean zero by construction, both among the treatment and control group.[25]

Column (1) of Table 5 presents OLS estimates of (2) (Panel A) and (3) (Panel B). Panel A shows that the estimate for γ_1 is 0.3: if informed that they underestimated the wage gap by $X\%$ for prior cohorts, job seekers adjust their perceived gap in potential wages by $0.3\,X\%$. This is in line with typical estimates from the literature on information treatments and belief updating.[26] In Panel B the estimate for θ_1 is 1.2: if informed that they underestimated the wage gap between job j and j' by $X\%$ for prior cohorts, job seekers respond by increasing their relative likelihood of applying for j by $1.2\,X\%$.

Columns (2)-(4) demonstrate the robustness of the estimates to controlling for pre-treatment covariates interacted with the underestimation of the prior cohort wage gap,

[25]As is easily verified, introducing a constant term and/or an uninteracted treatment dummy therefore generates exact zero estimates for the constant term and the coefficient on the treatment dummy, while leaving existing coefficients completely unchanged.

[26]Causal parameters akin to our γ_1 are commonly referred to as learning rates. Haaland et al. (2023) (Table) provides an overview of empirical learning rates found in different contexts. Studying beliefs about the overall wage level specifically for Danish job seekers, Altmann et al. (2025) find learning rates between 12% and 34%.

and/or restricting only to the sample of currently active job seekers.

Second stage and 2SLS estimates Having produced estimates of how much the information treatment shifts both relative wage perceptions and application decisions, we finally combine these to arrive at our key parameter of interest: the elasticity of the application decisions with respect to perceptions about relative potential wages. Formally, this corresponds to treating (2) and (3) as the first stage and reduced form in a 2SLS framework. The excluded instrument is the interacted treatment dummy, $T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$, and the second stage equation of interest is

$$\Delta\pi_{i,j,j'} = \beta_1 \, \Delta\widetilde{w_{i,j,j'}} + \beta_2 \, (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) + \varepsilon_{i,j,j'} \tag{4}$$

The coefficient β_1 corresponds to our key causal elasticity of interest; if the perceived wage gap between job j and j' increases by 1 percent, the relative likelihood of applying to j increases by β_1 percent. Consistent 2SLS estimation of β_1 requires the instrument, $T_i \times (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})$, to be uncorrelated with the error term, $\varepsilon_{i,j,j'}$. As above, random assignment implies that such a correlation cannot exist because the treatment and control group has different unobservables to begin with. The 2SLS estimator however also requires the stronger exclusion restriction that the instrument *only* affects application behavior through its effect on perceived wages. The information treatment cannot affect any other decision-relevant beliefs. As we return to in Section 4.3 below, our survey designs allow us to provide direct tests of this exclusion restriction and we find the assumption to be well-supported by the data.

Assuming that the exclusion restriction holds, Panel C of Table 5 shows 2SLS estimates of the key elasticity of interest. In the baseline specification in Column (1), increasing the perceived wage gap between j and j' by 1 percent increases the relative likelihood of applying for j by 4.2% percent. This is a substantial effect and - as we flesh out in Section 5 below - implies that wage misperceptions can generate important distortions in job search.

As a different useful point of comparison here, we can also compare to the literature on imperfect competition in the labor market and firm-specific labor supply curves. A key parameter in this literature is the own wage elasticity of job applications, i.e the percentage increase in likelihood of applying to a given job, if this job increases its offered wage by 1 percent. This is slightly a different parameter what we estimate here; Panel C of Table 5 shows the effect of changes in (perceived) wage gap between jobs on the *relative* likelihood of applying to the two jobs. [27] Using the tight link between our 2SLS framework and a standard discrete choice application model, however, it is straightforward to convert our parameter estimate into an own-wage elasticity of applications (see Appendix Section B.4 for details). At the bottom of Table 5 we apply this conversion and find an own wage elasticity of applications of about 3 — substantially higher than existing estimates using variation in posted wages typically on online job platforms (Holzer et al., 1991; Dal Bó et al., 2013; Banfi and Villena-Roldán, 2019; Marinescu and Wolthoff, 2020; Dube et al., 2020; Belot et al., 2022b; Azar et al., 2022). As we return to below, a likely explanation for this is that our information treatment isolates the effect of changes in perceived wages only, while existing studies capture the combined effect of job seekers observing higher wages and using these to make inferences about other job characteristics.

As above, Columns (2)-(4) of Table 5 show robustness to the inclusion of pre-treatment covarates and/or restricting to active job seekers. Appendix Table A5 show that results are also robust to considering only job seekers with a good understanding of the three job types and reweighing to correct for survey non-participation. Finally, Column (1) of Appendix Table A6 shows robustness to relaxing linearity of controls.

4.3 Testing the exclusion restriction: Belief spillovers

As noted above, a potential concern with our 2SLS estimates is belief spillovers from our wage information treatment. In general, upon learning that a particular job offers a higher wage, a job seeker could also perceive that the job is harder to get or that it offers a particular set of non-wage amenities. If such belief-spillover also affect job search decisions,

[27] To most clearly see the difference, consider a case where job j increases its offered wage while wages are unchanged in job j' and all other jobs. The likelihood of applying to job j increases in this case and the own wage elasticity measures the extent of this increase. If focusing instead on the *relative* likelihood of applying to j vs. j' however, there is an additional effect because in response to the wage increase, the likelihood of applying to job j' also decreases.

this would violate our 2SLS exclusion restriction.

Ex ante, we expect belief spillovers to be less likely in our case than in many other commonly studied settings. In response to a treatment that involves (or mimics) firms changing their actual posted wage for example, it would be rational for job seekers to expect higher posted wages to attract more applicants thus making the job harder to get. Similarly, they might also expect the wage change to go hand-in-hand with firms' making systematic changes to offered non-wage amenities; a higher wage might for example be introduced to compensate for longer or more strenuous work hours. In contrast, since our information treatment only informs a small number of job seekers about realized past wages, our treatment should not induce aggregate changes in the number of applicants and should be less likely to convey information about systematic changes in firms' offered terms of employment.

None of this of course rules that our wage information treatment affected beliefs about how hard it is to get hired or about non-wage amenities so it remains an empirical question.[28] Accordingly, we designed our survey to include questions that directly allow us to estimate and test for the existence of belief spillovers from our information treatment.

First, to examine potential spillovers in beliefs about how hard each job is to get, we asked respondents to report their perceived probability of getting offered each job if applying for it. Panel A of Table 6 reestimates the reduced form of our 2SLS framework using as outcome the log gap in these perceived success probabilities. We see no evidence that the information treatment affected perceived differencs in the probability of getting hired if applying. In addition to the treatment effect not being statistically significant, point estimates go in the opposite direction of what standard theories would predict.

Second, as a test for whether the information treatment changed job attractiveness by changing perceived non-wage amenities, we use answers in the "equal wage"-follow up to our job offer acceptance question. As described in Section 2.5, our basic job offer acceptance question asked respondents to consider a situation in which they simultaneously

[28]In a Bayesian learner framework for example, it is straightforward to write down both information structures in which wage signals do not lead to spillovers on beliefs about other characteristics and information structures in which they do. This is true regardless of whether job seekers perceive wages and other job characteristics to be correlated overall.

received mutually exclusive job offers from each of the three jobs, and then asked them to report their likelihood of accepting each offer. In the follow up, respondents were asked to report corresponding acceptance likelihoods in a counterfactual version of this situation. The counterfactual situation was described as differing *only* in that the offered starting wage in the three jobs were now equal.[29] All other aspects of the jobs were explicitly described as being the same as in the original situation. If the information treatment changes the perceived attractiveness of jobs' non-wage amenities, it should affect the reported acceptance likelihoods here. As a proxy for differences in non-pecuniary attractiveness, Panel B of Table 6 reestimates our reduced form using the (log) relative likelihood of job acceptance in this question as the outcome. We see no evidence that the information treatment affected the relative likelihood of accepting the different jobs in this 'equal wage' benchmark, suggesting that perceptions of non-wage amenities were unaffected by the information treatment.

Finally, we also asked respondents directly about four other decision-relevant job characteristics: the monthly wage they would expect to earn in five years if starting in each of the jobs today, the actual number of weekly hours they expect to work in each job, how well they expect to perform in each of the five jobs (6-item Likert scale) and how well they expect to get along with their colleagues in each job (6-item Likert scale). In Table 7, we use gaps across jobs in these dimensions as the outcome in our reduced form. We only see a significant effect for the expected wage in five years (Panel A). This again corroborates that while the information treatment affects wage perceptions, it does not affect perceived non-wage amenities.

To the extent we are able to test it, Tables 6 and 7 supports the exclusion restriction from our 2SLS analysis. As shown in Appendix Table A6, our main 2SLS estimates are also robust to using gaps in other perceived job characteristics as controls, including the reported offer acceptance likelihood under equal wages.

As noted above, this lack of belief spillovers may also explain why our estimates correspond to larger own wage elasticity of applications than in many existing studies; if wage

[29]Specifically, the wage of each job was specified to be equal to the middle offered wage that the job seeker expected to face across the three jobs.

increases are associated with perceived lower application success probabilities or worse non-wage amenities in these other setting, this will partially offset the positive effect of higher wages on applications. For completeness, Appendix Table A7 shows descriptive evidence on the correlation in beliefs about potential wages and other job characteristics.

4.4 Treatment effects on post-survey wages

The estimates above are based on measures of planned job search behavior reported in the survey. To validate whether the estimated effects also translate to actual changes in behavior post-survey, we finally use our individually-linked administrative data to examine treatment effects on post-survey wages.

As above, the predicted effect of the treatment depends on initial misperceptions. For job seekers who underestimate true wage gaps, misperceptions make lower-paying jobs appear relatively attractive. The estimated effects on job application behavior imply that the information treatment should cause these job seekers to apply for and earn higher wages post-survey. The converse should be true for job seekers who tend to overestimate true wage gaps.

In Table 8 we estimate treatment effects on starting wags in job seekers' first new job post-survey. Note that the analysis data differ from the sections above in that it now contains a single observation per job seeker. Moreover, since not all job seekers find a new job within our sample frame, the number of job seekers included in the analysis drops to 1,170.[30]

As a benchmark, Column (1) ignores heterogeneity by initial misperceptions and simply regresses log monthly wages in the new job on a treatment dummy. In line with most job seekers underestimating wage gaps, the overall average treatment effect is estimated to be positive, however it is not significantly different from zero.

[30]Our available administrative data runs until December 2024. Job seekers' new jobs are measured as the first post-survey job spell at a firm that the job seeker did not work at prior to the survey. To be counted, we require job spells to last more than one month and be the workers highest paid job. To avoid measurement errors from jobs the de facto started later in a month, we exclude the first month of employment but average over all remaining months of the employment when computing monthly wages in the new jobs. As we show in Appendix Table A8 we see no effect of the treatment on the likelihood of finding a new job within our sample period.

In Column (3) we instead test the predicted heterogeneous effects by initial misperceptions. As in Section 3, we classify each job seeker as either an *overestimator* or *underestimator* depending on whether they overestimate or underestimate the gap between the highest paying job and the average of the two other jobs. We then add a dummy for being an overestimator as well as an interaction between treatment and this overestimator dummy. With this, the coefficient on the treatment dummy measures the treatment effect for the group of underestimators. It reveals a statistically significant positive effect as predicted: receiving the information treatment causes underestimators to have 3.8 percent higher monthly wages on average in their post-survey job.

Turning to overestimators, the estimated coefficient on the interaction term captures the differential treatment effect for this group. Consistent with the treatment going in the opposite direction for overestimators, the estimated coefficient negative and statistically significant, albeit only at the 10% level. The implied average treatment effect for overestimators is shown at the bottom of the table: receiving the information treatment is estimated to lower monthly wages by 2.4 percent in the post-survey job, although this effects is not statistically significant.

Columns (2) and (4) of Table 8 examines the robustness of these estimates to the inclusion of pre-treatment controls. As expected given random assignment, estimates change little, although the coefficient on the interaction term fails to be significant when controls are added.

As in previous analyses, Columns (5) to (8) repeats the analyses using only data on respondents that report being actively searching at the time of the survey. Ex ante, we expect larger treatment effects for this subsample since we are now analyzing post-survey hiring outcomes. As discussed in Section 2.6 respondents that report not actively searching largely correspond to individuals who should have recently accepted a new job at the time of the survey. If job switching or additional search is costly, post-survey job outcomes for this group should respond less to the information treatment than for individuals who are still actively searching. This pattern is borne out by the data as all estimates increase markedly in magnitude when considering only active job seekers. Reaffirming the heterogeneous treatment effects for underestimators vs. overestimators, the estimated coefficient on the interaction term is also significant at the 5% level throughout in this sample.

Finally, the fact that Table 8 only includes job seekers who successfully found a job implies the usual concerns that the estimated wage effects could reflect composition changes. As we show in Appendix Table A8 however, we see no evidence that the information treatment affects job finding post-survey. This is consistent with our information treatment primarily changing beliefs about relative wages across jobs. These are first-order important for decisions about *which* jobs to target but not for job finding rates.[31]

5 Overall costs of relative wage misperceptions

Having documented extensive relative wage misperceptions and established their causal effects on job search behavior, we close our analysis by gauging the overall costs they impose on workers. With this aim in mind, our survey and reduced form analysis was deliberately designed to identify a simple discrete choice model of job search. Combining our reduced form estimates with this additional structure, we compute a counterfactual, showing what job search and hiring outcomes would look like in the absence of relative wage misperceptions. Below we outline the model framework and the key assumptions involved in the counterfactuals. Appendix Section B gives full details and derivations.

5.1 Model outline

Matching the setup in our survey, we consider a job seeker i who at a point in time has the option of applying for one of three different jobs indexed by $j = 1, 2, 3$. Alternatively, she can decide to apply for none of them. We assume that the continuation value of being hired into job j is $Y_{i,j}$ and that the continuation value of continuing to search is U_i. Accordingly, the surplus value of job j over continued search can be defined as $S_{i,j} = Y_{i,j} - U_i$. We will assume that this surplus value depends on the wage the worker will earn, $W_{i,j}$, and a set

[31] In standard models of job search, job finding rates respond primarily to beliefs about the overall level of wages since these affect the relative attractiveness of unemployment vs. employment (see e.g. Altmann et al. (2025)). Given that beliefs about wage levels are much closer to correct on average (cf. Figure 1, our information treatment should have a much more limited systematic impact on these beliefs, both overall and for the subgroups we consider here.

of non-wage amenities summarized by the scalar $Z_{i,j}$ as follows:[32]

$$S_{i,j} = \Psi_i W_{i,j}^{\delta} Z_{i,j} \tag{5}$$

Here Ψ_i is an individual-specific parameter allowing for heterogeneity in the general returns to employment, while δ is a parameter governing the role of wages in utility.

Optimal application behavior can be characterized by considering the surplus value of applying relative to the value of continued search. We let $A_{i,j}$ denote the corresponding surplus value of applying to job j. This will simply equal the likelihood of getting hired times the surplus value from being hired. Letting $P_{i,j}$ denote the probability of being hired at job j if applying, we have:

$$A_{i,j} = P_{i,j} S_{i,j} \tag{6}$$

When making decisions about where to apply, however, we assume that job seekers do not necessarily base their decision on the true surplus value of applying, $A_{i,j}$, but on their perceived surplus value, $\widetilde{A_{i,j}}$. We allow the perceived surplus to differ from the true value in two ways:

First, to allow for misperceptions, we assume that when evaluating the surplus of the job, the job seeker perceives the potential wage they would earn to be $\widetilde{W_{i,j}}$, the probability of being hired to be $\widetilde{P_{i,j}}$ and the non-wage job characteristics to be $\widetilde{Z_{i,j}}$.

Second, to generate standard discrete choice decision uncertainty we assume that the perceived surplus, $\widetilde{S_{i,j}}$, is subject to an idiosyncratic taste shocks, $\xi_{i,j}$, capturing idiosyncrasies in the specific surplus of job j and moreover that the perceived value of applying to any job is subject to a taste shock, $\xi_{i,0}$, capturing idiosyncrasies in the perceived utility costs of sending an application. The role of the latter is simply to generate decision uncertainty also around the relative value of deciding not to apply anywhere.

[32] Treating $Z_{i,j}$ as scalar is not restrictive. For example, if we were to instead assume that jobs are characterized by a K-dimensional vector $\left(Z_{i,j}^1, Z_{i,j}^2, ..., Z_{i,j}^K\right)$ and that surplus is given by $\Psi_i W_{i,j}^{\delta} \Pi_k \left(Z_{i,j}^k\right)^{\delta_k}$, we would still arrive at equation (5) by simply defining $Z_{i,j} = \Pi_k \left(Z_{i,j}^k\right)^{\delta_k}$.

Putting this together the perceived surplus value of applying to job j satisfies:

$$\widetilde{A_{i,j}} = \widetilde{P_{i,j}}\widetilde{S_{i,j}} - \xi_{i,0} \tag{7}$$

$$\widetilde{S_{i,j}} = \Psi_i \widetilde{W_{i,j}}^{\delta} \widetilde{Z_{i,j}} \xi_{i,j} \tag{8}$$

Finally, imposing a standard extreme value distribution on the distribution of the taste shocks, optimal behavior for job seekers implies a standard expression for the likelihood of applying to job j:[33]

$$P\left(\widetilde{A_{i,j}} = \max_{j'} \widetilde{A_{i,j'}}\right) = \frac{\exp((\psi_i + \widetilde{p_{i,j}} + \delta \widetilde{w_{i,j}} + \widetilde{z_{i,j}})/\sigma)}{1 + \sum_{j'=1,2,3} \exp((\psi_i + \widetilde{p_{i,j'}} + \delta \widetilde{w_{i,j'}} + \widetilde{z_{i,j'}})/\sigma)} \tag{9}$$

Here, σ is a scale parameter for the taste shocks, and we again use lower case letters to refer to logged values.

Equation 9 motivates our survey design and reduced form regression analyses. Prior to actually making an application decision, the left hand side of Equation (9) shows the expected likelihood that job seeker i will choose to apply to job j. This is exactly what job seekers report in our main survey question on application probabilities. The right hand side of (9) shows how this likelihood depends on perceived characteristics of the different jobs, including particularly the potential wage they expect to face in the job, $\widetilde{w_{i,j}}$, which job seekers also report in our survey. Accordingly, the second stage regression of our 2SLS framework can be derived from (4) (see Appendix Section B for details). This implies a direct link between our reduced form 2SLS estimates and the stuctural parameters of the model. Below, we use this to compute a model counterfactual based on the 2SLS estimates.

5.2 Counterfactuals and additional assumptions

To measure the overall consequences of relative wage misperceptions, we first compute counterfactual outcomes for every individual in our control group in the absence of relative wage misperceptions. We then compare these to their actual observed outcomes under current misperceptions. Besides our reduced form estimates, the exact additional assump-

[33]Specifically, we assume that the taste shocks are i.i.d. Frechet-distributed with shape parameter σ^{-1}

tions we need as inputs to compute counterfactuals depends on the range of outcomes we consider. We discuss these in turn below. Appendix B present formal statements of the assumptions and associated derivations.

Counterfactual application behavior and applied-for wages Since search behavior depend on the potential wages workers themselves expect to face, computing any counterfactual requires us to first take a stand on what relative misperceptions about these potential wages are like at baseline. Given the content of our survey, we adopt the simple assumption that our measured relative misperceptions about wage gaps for prior cohorts are the same as the relative misperceptions about the potential wage gaps job seekers expect to face across jobs. This benchmark assumption is sufficient to compute counterfactual application behavior without relative wage misperceptions, including the average wage in applied-for jobs.

5.2.1 Counterfactual employment outcomes and wages

In addition to application behavior, we are interested in counterfactual hiring and wage outcomes. To compute these, we need to impose additional structure on the probability of getting hired if applying. Since the focus of our counterfactual is on wage misperceptions, a simple benchmark is to abstract from misperceptions about hiring probabilities and use job seekers self-reported application success probabilities as the truth in all cases. Importantly, this assumption also abstracts from general equilibrium effects, where changes in aggregate application behavior may change the individual likelihood of getting hired in a job.[34]

Additionally, we need to take a stand on what happens to job seekers for whom the first application decision does not result in employment. Mirroring standard job search models, we assume that these workers continue searching with some individual-specific flow utility and time discount rate, b_i and $\rho_i > 0$, while sending additional applications at some Poisson rate, $\lambda_i > 0$.

[34] Accounting satisfactorily for such effects would require us to take a stand on how job creation responds to aggregate application patterns. In isolation, if job seekers start to send more application to a particular job type, hiring likelihoods will drop. In response to the increase in applicants however firms will have an incentive to create additional jobs, thereby offsetting the drop. Since our data contain no information about job creation, we view fixed hiring likelihoods as the most transparent benchmark.

5.2.2 Counterfactual welfare gains

Besides changes in applications, hiring and wages, we can also use the model framework to assess welfare. Specifically, we benchmark the welfare gains from removing relative wage misperceptions against an across-the-board wage increase in all jobs. This amounts to asking how much higher all wages would have to be in order to deliver the same welfare gains as removing relative wage misperceptions.

To perform this welfare counterfactual requires two additional assumptions. First, we need to take a stand on misperceptions about non-wage amenities. Again, since such misperceptions are not the focus of our analysis, we adopt the benchmark that workers in fact correctly perceive non-wage amenities. Second, we need to impose a parameter restriction either on the parameter governing the role of wages in utility, δ, or on the scale parameter of the taste shocks, σ. This reflects the usual challenge that choice behavior does not separately identify these parameters.[35] We present results from three different parameter restrictions here: i) imposing that the surplus value of a job is linear in the wage, $\delta = 1$, ii) normalizing the scale of taste shocks to unity as is often done in applications, $\sigma = 1$, ii), and iii) calibrating the scale of the taste shocks to $\sigma = 1.66$ based on the cross-sectional relationship between perceived application success probabilities and the likelihood of applying.[36]

5.3 Model counterfactual, results

Table 9 shows the results of from our model counterfactual. The columns correspond to different outcome variables and show how these would change in the absence of relative wage misperceptions. Since our model counterfactual reveals the full distribution of these changes across job seekers, the table shows both the mean change, the 10th and the 90th percentile when these are relevant. Standard errors in parenthesis were obtained by bootstrapping individuals and recomputing both the 2SLS estimates and the model

[35] From our reduced form equations, the identified parameter governing application behavior is $\beta_1 = \frac{1}{\sigma}\delta$. Scaling up σ and δ by the same amount thus leaves application behavior unchanged.

[36] Across two jobs j and j', Equation (9) implies that if the perceived gap in the likelihood of being hired increases by one percent, the relative likelihood of applying should change by σ^{-1} percent. As shown in Appendix Table A6, adding log gap in the likelihood of being hired as an additional control in our 2SLS equation of interest yields a coefficient of $0.602 \approx 1.66^{-1}$, corresponding to $\sigma = 1.66$.

counterfactuals.

The first four columns examine changes in application behavior. Removing relative wage misperceptions would cause 10.4 % of applications to go to a different job on average. In other words, about one-in-ten applications are distorted by misperceptions. In line with the fact that wage gaps are underestimated in most cases, misperceptions tend to distort applications away from high-paying jobs; for the average job seeker, removing misperceptions leads to an increase in the average applied-for wage of 1.4 %. Given the variation in the size and direction misperceptions however, this average effect reflects a high degree of heterogeneity. For more than 10% of job seekers, removing relative wage misperceptions would in fact *decrease* the average applied-for wage by more than 0.2%, while another 10% of job seekers would see increases in excess of 4.6%.

The remaining columns of the table consider changes in actual hiring outcomes and welfare under gradually stronger assumptions. The middle columns examine hiring outcomes, accounting for that fact that removing relative misperceptions might lead job seekers to apply for jobs with higher or lower application success probabilities. As one might expect, effects on hiring outcomes are slightly more muted because of this, however, the difference is slight. Removing relative wage misperceptions would reallocate 9.2% of workers to different jobs. Wages would increase by 1.2% on average but again with highly heterogeneous effects: 10% of job seekers would see decreases of more than 0.2%, while another 10% would see increases in excess of 3.9%. Finally the last three columns examine welfare effects under different parameter restrictions. Depending on the restriction, we estimate that the average welfare gain from removing relative wage misperceptions is equivalent to a 0.5-0.8 % increase in all wages.

Overall, we conclude that relative wage misperceptions impose substantial costs in our setting by distorting the allocation of workers to jobs.

6 Conclusion

This paper examines relative wage misperceptions and their consequences for job search among early career job seekers in Denmark. Job seekers have large misperceptions about wage differences across jobs. While there is substantial heterogeneity, most job seekers tend to underestimate the true wage gap between jobs such that lower-paying jobs look relatively more attractive. These misperceptions have large effects on job search decisions and hiring outcomes, on average shifting workers towards suboptimally lower-paying jobs.

The findings confirm that relative wage misperceptions impose substantial costs by distorting job search and the allocation of workers to jobs. They also confirm that wage misperceptions may be a source of market power for employers, at least among early career workers; if most workers systematically perceive wage differences across employers to be smaller than they are, this dampens wage competition between employers.

In terms of future work, the findings also raise several additional questions. First, given the substantial heterogeneity in wage misperceptions that we observe, it is natural to ask how differences in wage misperceptions may contribute to labor market inequality, both overall and across specific subgroups.

Second, while our analysis focuses on misperceptions about differences in starting wages across jobs, it is natural to ask whether misperceptions about other job characteristics also distort job search. This includes misperceptions about differences across jobs in both opportunity for career-advancement, non-wage amenities or the likelihood of getting hired if applying.

Finally, as suggested by our simple information treatment, labor market outcomes can be improved by interventions that correct relative wage beliefs. Understanding and developing policy initiatives that achieve this effectively thus seems like an important topic for future work.

Bibliography

Altmann, Steffen, Anita Glenny, Robert Mahlstedt, and Alexander Sebald, "The Direct and Indirect Effects of Online Job Search Advice," IZA Discussion Papers 15830, Institute of Labor Economics (IZA) 2022.

_ , Armin Falk, Simon Jäger, and Florian Zimmermann, "Learning about job search: A field experiment with job seekers in Germany," *Journal of Public Economics*, August 2018, *164*, 33–49.

_ , Robert Mahlstedt, Malte Rattenborg, Alexander Sebald, Sonja Settele, and Johannes Wohlfart, "Wage Expectations and Job Search," 2025. .

Ansolabehere, Stephen, Marc Meredith, and Erik Snowberg, "Asking about numbers: Why and how," *Political Analysis*, 2013, *21* (1), 48–69.

Arcidiacono, Peter, V. Joseph Hotz, and Songman Kang, "Modeling College Major Choices Using Elicited Measures of Expectations and Counterfactuals," *Journal of Econometrics*, January 2012, *166* (1), 3–16.

Arellano-Bover, Jaime, "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size," *Journal of Labor Economics*, April 2024, *42* (2), 549–589. Publisher: The University of Chicago Press.

Azar, José A., Steven T. Berry, and Ioana Marinescu, "Estimating Labor Market Power," August 2022. .

Ballarino, Gabriele, Antonio Filippin, Giovanni Abbiati, Gianluca Argentin, Carlo Barone, and Antonio Schizzerotto, "The effects of an information campaign beyond university enrolment: A large-scale field experiment on the choices of high school students," *Economics of Education Review*, 2022, *91*, 102308.

Banfi, Stefano and Benjamín Villena-Roldán, "Do High-Wage Jobs Attract More Applicants? Directed Search Evidence from the Online Labor Market," *Journal of Labor Economics*, 2019.

Behaghel, Luc, Sofia Dromundo, Marc Gurgand, Yagan Hazard, and Thomas Zuber, "The Potential of Recommender Systems for Directing Job Search: A Large-Scale Experiment," IZA Discussion Papers 16781, Institute of Labor Economics (IZA) February 2024.

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

_ , _ , and _ , "Do the Long-Term Unemployed Benefit from Automated Occupational Advice during Online Job Search?," IZA Discussion Papers 15452, Institute of Labor Economics (IZA) July 2022.

_ , _ , and _ , "How Wage Announcements Affect Job Search—A Field Experiment," *American Economic Journal: Macroeconomics*, October 2022, *14* (4), 1–67.

Betts, Julian R., "What Do Students Know about Wages? Evidence from a Survey of Undergraduates," *The Journal of Human Resources*, 1996, *31* (1), 27–56.

Bó, Ernesto Dal, Frederico Finan, and Martín A. Rossi, "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*," *The Quarterly Journal of Economics*, August 2013, *128* (3), 1169–1218.

Bonilla-Mejía, Leonardo, Nicolas L Bottan, and Andrés Ham, "Information policies and higher education choices experimental evidence from Colombia," *Journal of Behavioral and Experimental Economics*, 2019, *83*, 101468.

Caldwell, Sydnee, Ingrid Haegele, and Jörg Heining, "Firm Pay and Worker Search," February 2025.

Cockx, Bart and Corinna Ghirelli, "Scars of recessions in a rigid labor market," *Labour Economics*, August 2016, *41*, 162–176.

Conlon, John J and Dev Patel, "What Jobs Come to Mind? Stereotypes about Fields of Study," 2025. .

Dahl, Christian, Daniel le Maire, and Jakob Munch, "Wage Dispersion and Decentralization of Wage Bargaining," *Journal of Labor Economics*, 2013, *31* (3), 501 – 533. Publisher: University of Chicago Press.

Dhia, Aïcha Ben, Bruno Crépon, Esther Mbih, Louise Paul-Delvaux, Bertille Picard, and Vincent Pons, "Can a Website Bring Unemployment Down? Experimental Evidence from France," Working Paper 29914, National Bureau of Economic Research April 2022.

Dube, Arindrajit, Jeff Jacobs, Suresh Naidu, and Siddharth Suri, "Monopsony in Online Labor Markets," *American Economic Review: Insights*, March 2020, *2* (1), 33–46.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart, "Designing Information Provision Experiments," *Journal of Economic Literature*, March 2023, *61* (1), 3–40.

Harmon, Nikolaj A, Marie Pietraszek, and Jonas Maibom, "Why do coworker networks affect job search outcomes?," Study Paper No. 262, The ROCKWOOL Foundation 2025.

_ **, Robert Mahlstedt, and Mette Rasmussen**, "Job Search, Overoptimism and Statistical Profiling: Can Information Provision Improve Job Search Outcomes?," 2024.

Hastings, Justine S, Christopher A Neilson, Anely Ramirez, and Seth D Zimmerman, "(Un) informed college and major choice: Evidence from linked survey and administrative data," *Economics of Education Review*, 2016, *51*, 136–151.

Holzer, H. J., L. F. Katz, and A. B. Krueger, "Job Queues and Wages," *The Quarterly Journal of Economics*, August 1991, *106* (3), 739–768.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer, "Worker Beliefs About Outside Options," *The Quarterly Journal of Economics*, August 2024, *193* (3), 1505–1556.

Kahn, Lisa B., "The long-term labor market consequences of graduating from college in a bad economy," *Labour Economics*, April 2010, *17* (2), 303–316.

Kerr, Sari Pekkala, Tuomas Pekkarinen, Matti Sarvimäki, and Roope Uusitalo, "Post-secondary education and information on labor market prospects: A randomized field experiment," *Labour Economics*, 2020, *66*, 101888.

Krueger, Alan B. and Andreas I. Mueller, "A Contribution to the Empirics of Reservation Wages," *American Economic Journal: Economic Policy*, February 2016, *8* (1), 142–179.

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

Marinescu, Ioana and Ronald Wolthoff, "Opening the Black Box of the Matching Function: The Power of Words," *Journal of Labor Economics*, April 2020, *38* (2), 535–568.

Mueller, Andreas I., Johannes Spinnewijn, and Giorgio Topa, "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias," *American Economic Review*, January 2021, *111* (1), 324–363.

Spinnewijn, Johannes, "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, February 2015, *13* (1), 130–167.

von Wachter, Till, "The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources," *Journal of Economic Perspectives*, November 2020, *34* (4), 168–194.

von Wachter, Till and Stefan Bender, "In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers," *American Economic Review*, December 2006, *96* (5), 1679–1705.

Wiswall, M. and B. Zafar, "Determinants of College Major Choice: Identification using an Information Experiment," *The Review of Economic Studies*, April 2015, *82* (2), 791–824.

Wiswall, Matthew and Basit Zafar, "How Do College Students Respond to Public Information about Earnings?," *Journal of Human Capital*, June 2015, *9* (2), 117–169.

_ **and** _ , "Preference for the Workplace, Investment in Human Capital, and Gender*," *The Quarterly Journal of Economics*, February 2018, *133* (1), 457–507.

Økonomi og indenrigsministeriet, "Dagpenge til nyuddannede," Økonomisk Analyse nr. 33, Økonomi og indenrigsministeriet 2018.

7 Tables and Figures

Table 1: Study sample individuals, covariate balance and descriptive statistics

	(1) All	(2) Control group	(3) Treatment group	(4) p-value (2) = (3)
Age	28.07	28.15	27.98	0.326
Female	0.62	0.61	0.63	0.420
Months between survey and expected/actual graduation	16.61	17.58	15.60	0.195
Employed long-term	0.11	0.12	0.11	0.270
Employed short-term, without job offer	0.05	0.06	0.04	0.117
Employed short-term, with job offer	0.03	0.04	0.02	0.122
Unemployed, with job offer	0.25	0.25	0.25	0.972
Unemployed, without job offer	0.56	0.53	0.58	0.060*
Currently studying	0.07	0.07	0.07	0.957
Actively searching for a job	0.65	0.63	0.67	0.130
Currently searching or about to start new job	0.87	0.86	0.88	0.112
Currently searching, about to start new job or planning to search in the next 3 months	0.89	0.88	0.89	0.270
Good understanding of:				
Most common job	0.82	0.81	0.82	0.564
Second most common job	0.73	0.73	0.73	0.940
Least common job	0.72	0.73	0.71	0.388
Close to someone:				
In most common job	0.59	0.60	0.58	0.394
In second most common job	0.49	0.48	0.49	0.666
In least common job	0.51	0.53	0.49	0.092*
Perceived avg. wage of prior cohorts (1000s):				
In most common job	32.27	32.13	32.42	0.259
In second most common job	31.21	31.16	31.26	0.717
In least common job	32.55	32.37	32.73	0.180
Perceived avg. wage of prior cohorts after 5 years (1000s):				
After starting at most common job	39.51	39.51	39.52	0.968
At second most common job	38.15	38.10	38.21	0.784
At least common job	39.89	39.91	39.88	0.943
Perceived avg. probability to receive offer, if applying:				
In most common job	0.17	0.17	0.16	0.263
In second most common job	0.15	0.15	0.14	0.481
In least common job	0.14	0.15	0.14	0.639
Part of STAR sample	0.91	0.91	0.91	0.971
Individuals	1902	976	926	

Notes: The first three columns of the table show sample means for all pre-treatment survey variables, respectively for the full study sample (Column (1)), only for the control group (Columns (2)) and only for the treatment group (Column (3)). Column (4) shows p-values from a test of equal means between the treatment and control groups. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table 2: Descriptive statistics, survey job types

	(1) Share of first jobs, prior cohorts	(2) Stated likelihood of applying	(3) Avg. wage, prior cohorts	(4) Perceived avg. wage, prior cohorts	(5) Perceived own potential wage
Data source:	Admin	Survey	Admin	Survey	Survey
Panel A: *Grouped by share of first jobs in prior cohorts*					
Highest	25.64%	37.80%	30,883	32,269	33,312
Middle	14.81%	24.11%	30,735	31,208	32,385
Lowest	09.62%	28.75%	31,674	32,548	33,503
Panel B: *Grouped by stated likelihood of applying*					
Highest	18.66%	61.94%	31,308	33,057	34,298
Middle	15.93%	21.73%	31,199	32,095	33,101
Lowest	15.48%	07.00%	30,786	30,873	31,801
Panel C: *Grouped by avg. wage of prior cohorts*					
Highest	14.75%	31.46%	33,148	32,763	34,091
Middle	17.93%	31.04%	31,176	31,901	33,075
Lowest	17.39%	28.16%	28,968	31,360	32,033

Notes: The table shows study sample means for various characteristics of the three job types asked about in the survey. Columns correspond to different variables, either from administrative data or the survey. Panels correspond to different ways of grouping the three jobs into highest, middle and lowest. In Panel A jobs are grouped in terms of how large a share of entry level jobs they cover in past cohorts. In Panel B, jobs are grouped in terms of the respondent's stated likelihood of applying to them. In Panel C, jobs are grouped by their average wage for prior cohorts.

Table 3: Relative wage misperceptions, summary

	Wage gap underestimation rate between job pairs $\frac{\Delta V_{i,j,j'} - \Delta \widetilde{V_{i,j,j'}}}{\Delta V_{i,j,j'}}$				Underestimation of log wage gap between job pairs $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$		Individuals who are underestimators	Sample size:	
	$> 100\%$	$> 50\%$	$> 0\%$	$< -50\%$	Mean abs. value	Std. Dev.	Share	Individuals	Job pairs
Full sample	36.1%	55.2%	66.7%	25.1%	0.129	0.167	66.6%	1902	5706
Actual log wage gap > 0.05	30.0%	51.1%	68.6%	20.2%	0.142	0.164	68.8%	1583	2456
Actual log wage gap > 0.1	20.9%	45.7%	69.2%	17.3%	0.146	0.166	70.3%	1046	1471
Two most frequent job types	36.2%	55.4%	66.5%	25.3%	0.123	0.162	66.6%	1902	3804
Likelihood of applying $> 15\%$ for both jobs, control group	31.3%	54.9%	66.3%	24.8%	0.114	0.151	65.7%	623	1392
Full sample, correcting reporting errors	36.2%	55.1%	66.8%	24.9%	0.126	0.162	66.5%	1901	5703
Currently active job seekers	35.9%	55.3%	65.9%	26.5%	0.129	0.166	66.1%	1237	3711
Reweighted for non-participation	35.9%	55.3%	67.0%	24.4%	0.130	0.170	65.6%	1826	5478
Good understanding of all job types	34.5%	54.2%	66.7%	24.5%	0.125	0.163	66.3%	1019	3057
Correct answer to job type validation question	35.0%	52.4%	63.8%	27.6%	0.123	0.159	65.8%	1135	3405

Notes: The table summarizes the distribution of relative wage misperceptions. Rows correspond to different subsamples. The first four columns shows the share of individual-by-job pairs (i, j, j') for which the wage gap underestimation rate exceeds 100%, 50%, 0% or falls below -50% respectively. The next two columns report the mean absolute value and standard deviation of the underestimation of the log wage gap. The third last column instead uses individual-level data (i) and shows what fraction of individuals are underestimators in the sense that they underestimate the typical wage in the highest paying job relative to the average of the two other jobs. The last two columns show the number of individuals and unique individual-by-job pairs in each subsample. The shown subsamples are as follows: *Actual log wage gap* restricts to job pairs where the actual log gap in prior cohort wages is above some cutoff. *Two most frequent job types* restricts to the two jobs for each education that cover the largest share of starting jobs for past cohorts. *Likelihood of applying>15% for both jobs* only considers the control group and restricts the sample based on the respondent's reported likelihood of applying for the job types. *Correcting reporting errors* uses a simple procedure to correct likely reporting errors (Appendix Section D.1 provides details). *Currently active job seekers* are respondents who report being actively searching at the time of the survey. *Reweighted for non-participation* reweighs the study sample to match the invited population using propensity score reweighting. *Good understanding of all job types* restricts to respondents who reported having a "Good" or "Very good" understanding of what the survey job types mean. *Correct answer to job type validation question* restricts to respondents who were shown the job type validation question and answered it correctly (see Appendix Section C.3.2).

Table 4: Overall treatment effect

	(1)	(2)	(3)	(4)
A): Perceived gap in log own potential wages $\Delta\widetilde{w_{i,j,j'}}$				
Treatment, T_i	0.042***	0.042***	0.043***	0.042***
	(0.005)	(0.005)	(0.006)	(0.006)
B): Gap in log likelihood of applying $\Delta\pi_{i,j,j'}$				
Treatment, T_i	0.195***	0.196***	0.199**	0.192**
	(0.073)	(0.071)	(0.089)	(0.088)
Pre-treatment controls		✓		✓
Currently active job seekers			✓	✓
Individuals	1902	1902	1237	1237
N	5706	5706	3711	3711

Notes: The table shows OLS estimates from a regression including a constant term, the treatment dummy and possibly controls. The base data are individual-by-job pairs (i, j, j') in the study sample, ordered so that the wage gap for prior cohorts is underestimated in all cases, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}} > 0$. Columns (3) and (4) restrict attention only to individual reporting that they are actively searching at the time of the survey. Columns (2) and (4) include as controls all pre-treatment survey variables from Table 1. Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table 5: Effect of relative misperceptions on search, 2SLS

	(1)	(2)	(3)	(4)
A) First stage: Perceived gap in log own potential wages $\Delta\widetilde{w_{i,j,j'}}$				
$T_i \times (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})$	0.294***	0.298***	0.302***	0.297***
	(0.032)	(0.031)	(0.040)	(0.039)
B) Reduced form: Gap in log likelihood of applying $\Delta\pi_{i,j,j'}$				
$T_i \times (\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})$	1.235***	1.275***	1.151*	1.039*
	(0.468)	(0.450)	(0.596)	(0.577)
C) Second stage: Gap in log likelihood of applying $\Delta\pi_{i,j,j'}$				
$\Delta\widetilde{w_{i,j,j'}}$	4.197***	4.276***	3.813**	3.500*
	(1.539)	(1.452)	(1.909)	(1.862)
Controls:				
$\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$	✓	✓	✓	✓
Pre-treatment controls		✓		✓
Currently active job seekers			✓	✓
First stage F-stat	85.48	90.92	56.72	57.44
Own wage elasticity	2.93	2.98	2.66	2.44
Individuals	1902	1902	1237	1237
N	11412	11412	7422	7422

Notes: The table shows results from the 2SLS framework. The base data are all individual-by-job pairs (i, j, j') in the study sample. Panel A and B show OLS estimates, while Panel C shows 2SLS estimates using $T_i \times \left(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right)$ as the excluded instrument. Besides the reported regressors, all specifications include the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$, as a control. Columns (3) and (4) restrict attention only to individuals reporting that they are actively searching at the time of the survey. Columns (2) and (4) include as additional controls all pre-treatment survey variables from Table 1 interacted with the underestimation of the log wage gap for prior cohorts. The implied own wage elasticity of applications to a job is computed at the bottom via a standard discrete choice framework (see Appendix Section B.4 for details). Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table 6: Belief spillovers and the exclusion restriction

	(1)	(2)	(3)	(4)
A): Perceived gap in log probability of offer, if applying				
$T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$	0.361	0.371	0.389	0.440
	(0.256)	(0.254)	(0.331)	(0.325)
Individuals	1859	1859	1204	1204
N	10762	10762	6936	6936
B): Perceived gap in non-pecuniary attractiveness of job				
$T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$	0.355	0.457	0.758	0.825
	(0.461)	(0.456)	(0.588)	(0.579)
Individuals	1902	1902	1237	1237
N	11412	11412	7422	7422
Controls:				
$\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$	✓	✓	✓	✓
Pre-treatment controls		✓		✓
Currently active job seekers			✓	✓

Notes: The table shows OLS estimates. The base data are all individual-by-job pairs (i, j, j') in the study sample. The outcome variable in Panel A is the job seekers perceived gap in the (log) probability of getting an offer if applying to the jobs. The outcome variable in Panel B is a proxy for the difference in the non-pecuniary attractiveness of the jobs which is defined as the gap in the log reported likelihood of accepting an offer from the jobs in the situation where the jobs offer the exact same starting wage. Changes in observation totals across panels stem from missing reports and zeros in the raw outcome variables. Besides the reported regressor, all specifications include the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$ as a control. Columns (3) and (4) restrict attention only to individuals reporting that they are actively searching at the time of the survey. Columns (2) and (4) include as additional controls all pre-treatment survey variables from Table 1 interacted with the underestimation of the log wage gap for prior cohorts. Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table 7: Belief spillovers on additional job characteristics

	(1)	(2)	(3)	(4)
A): Perceived gap in log wage after 5 years				
$T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$	0.234***	0.239***	0.270***	0.265***
	(0.038)	(0.038)	(0.047)	(0.046)
Individuals	1887	1887	1228	1228
N	11314	11314	7364	7364
B): Perceived gap in log weekly hours worked				
$T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$	0.017	0.025	0.029	0.027
	(0.026)	(0.026)	(0.032)	(0.031)
Individuals	1855	1855	1208	1208
N	10998	10998	7160	7160
C): Gap in how well they think they would get on with their colleagues				
$T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$	0.041	-0.016	0.374	0.284
	(0.237)	(0.238)	(0.300)	(0.295)
Individuals	1902	1902	1237	1237
N	11412	11412	7422	7422
D): Gap in how well they think they would perform in the job type				
$T_i \times (\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}})$	0.139	0.128	0.318	0.291
	(0.329)	(0.322)	(0.416)	(0.403)
Individuals	1902	1902	1237	1237
N	11412	11412	7422	7422
Controls:				
$\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$	✓	✓	✓	✓
Pre-treatment controls		✓		✓
Currently active job seekers			✓	✓

Notes: The table shows OLS estimates. The base data are all individual-by-job pairs (i, j, j') in the study sample. In Panels A-B, outcome variables are gaps in logged values of perceived job characteristics. In Panel C-D, outcome variables are the gap in how well the job seeker think they would get on with colleagues in the jobs and in how well the job seeker think they would perform in the jobs (6-item Likert scales). Changes in observation totals across panels stem from missing reports and zeros in the raw outcome variables. Besides the reported regressor, all specifications include the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$ as a control. Columns (3) and (4) restrict attention only to individuals reporting that they are actively searching at the time of the survey. Columns (2) and (4) include as additional controls all pre-treatment survey variables from Table 1 interacted with the underestimation of the log wage gap for prior cohorts. Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effect of information treatment on post-survey wages

	Full sample				Currently active job seekers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment, T_i	0.017	0.019	0.038**	0.034**	0.026	0.022	0.059***	0.053**
	(0.015)	(0.014)	(0.018)	(0.016)	(0.019)	(0.018)	(0.023)	(0.021)
$T_i \times$Overestimator			-0.062*	-0.047			-0.096**	-0.091**
			(0.032)	(0.030)			(0.040)	(0.039)
Overestimator			0.042*	0.036*			0.050*	0.050*
			(0.022)	(0.020)			(0.028)	(0.026)
Treatment Effect			-0.024	-0.013			-0.037	-0.038
for Overestimators			(0.027)	(0.025)			(0.033)	(0.033)
Controls		✓		✓		✓		✓
N	1,170	1,170	1,170	1,170	802	802	802	802

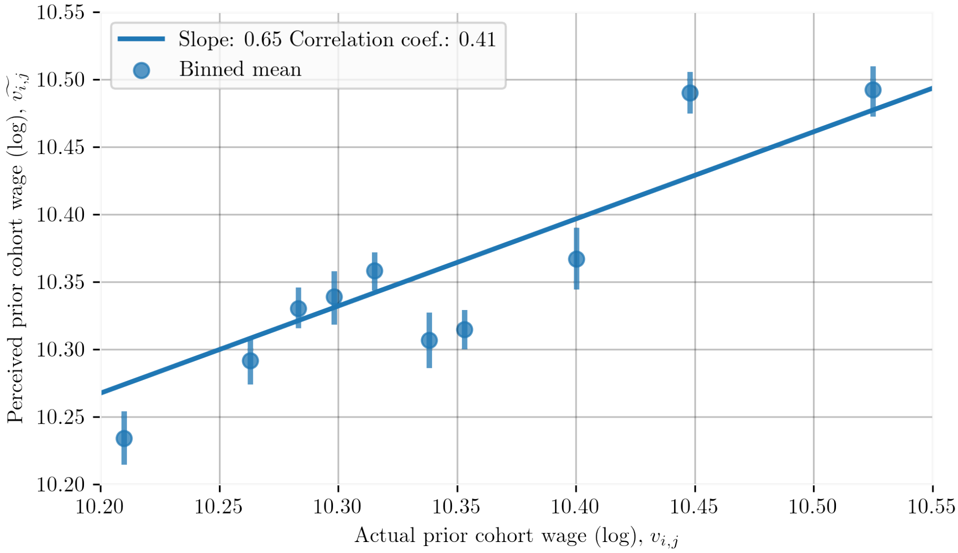
Notes: The table shows OLS estimates from linked administrative data. The base data are individuals (i) in the study sample who are observed finding a new job post-survey. The outcome variable is the log monthly wage in the new job. Columns (1) and (5) contains the most parsimonious regression specifications, including only a constant term and treatment dummy. Other columns additionally add a dummy for the job seeker being an overestimator along with the corresponding interaction with the treatment dummy, and/or add all pre-treatment survey variables from Table 1 as controls. Columns (5)-(8) restrict attention only to individuals reporting that they are actively searching at the time of the survey. In Columns (3), (4), (7) and (8), the implied treatment effect estimate for overestimators is reported in the bottom row along with its standard error. Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table 9: Model Counterfactual: Effects of removing relative wage misperceptions

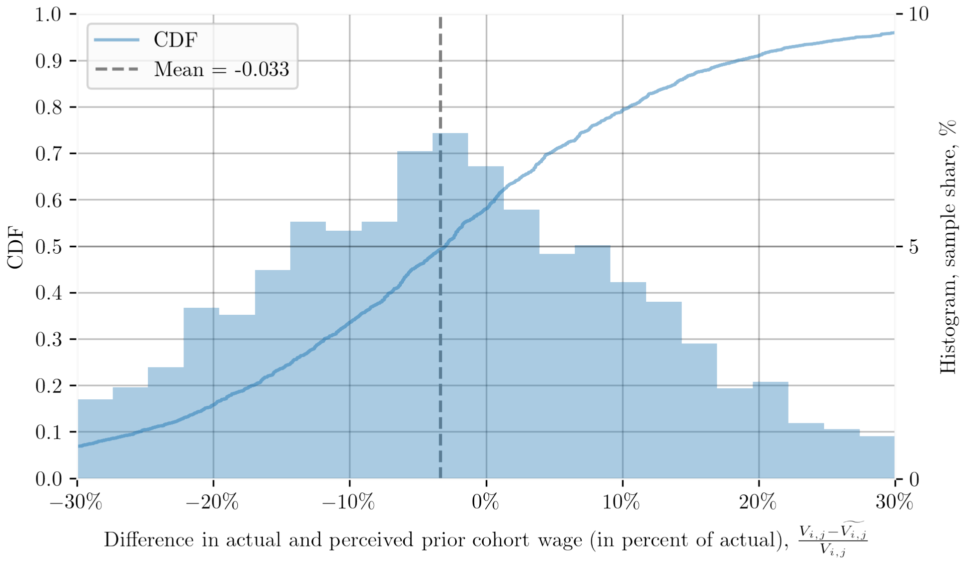
Application behavior:				Hiring outcomes:				Welfare:		
Share of applications reallocated:	Growth, applied-for wage: *Mean*	*10th pctile*	*90th pctile*	Share of workers reallocated:	Growth, actual wage *Mean*	*10th pctile*	*90th pctile*	Mean gain (wage growth equiv.) $\sigma = 0.24$ $\delta \equiv 1.00$	$\sigma \equiv 1.00$ $\delta = 4.20$	$\sigma \equiv 1.66$ $\delta = 6.97$
10.4***	1.41***	-0.29*	4.61***	9.22***	1.24***	-0.25**	3.90***	0.83	0.67*	0.56**
(3.67)	(0.51)	(0.17)	(1.67)	(3.27)	(0.46)	(0.12)	(1.48)	(0.52)	(0.36)	(0.28)

Notes: The table shows the predicted effects of removing relative wage misperceptions on various outcomes based on a discrete-choice model of job search. Using the reduced form estimates, counterfactual model outcomes for each individual in the control group are computed and then compared to actual outcomes. Details of the model and counterfactual calcuations are given in Section 5 and Appendix B. Standard errors in parenthesis are obtained by a bootstrap procedure that resamples individuals and recomputes both 2SLS estimates and model counterfactuals.

Figure 1: Misperceptions about wages in individual jobs



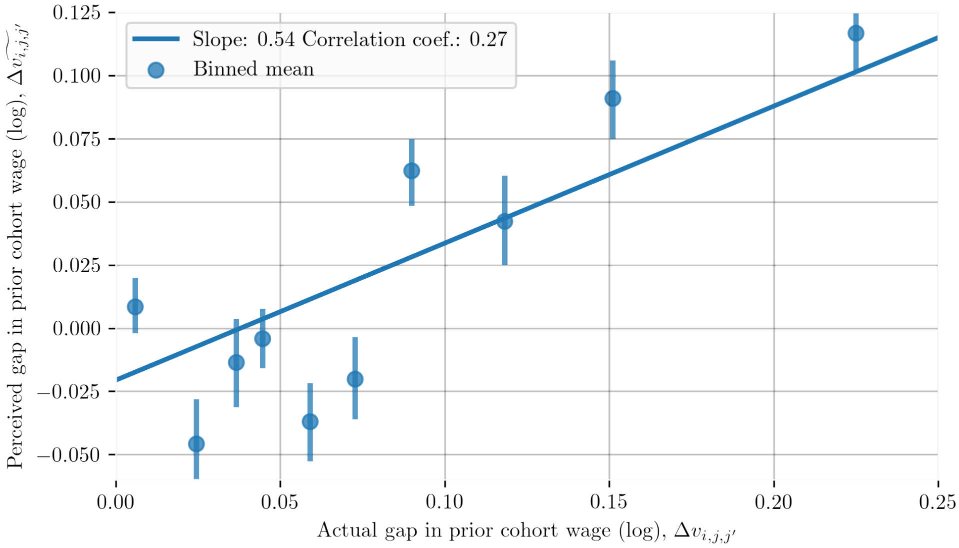
(a) Binscatter, perceived vs. actual prior-cohort wages



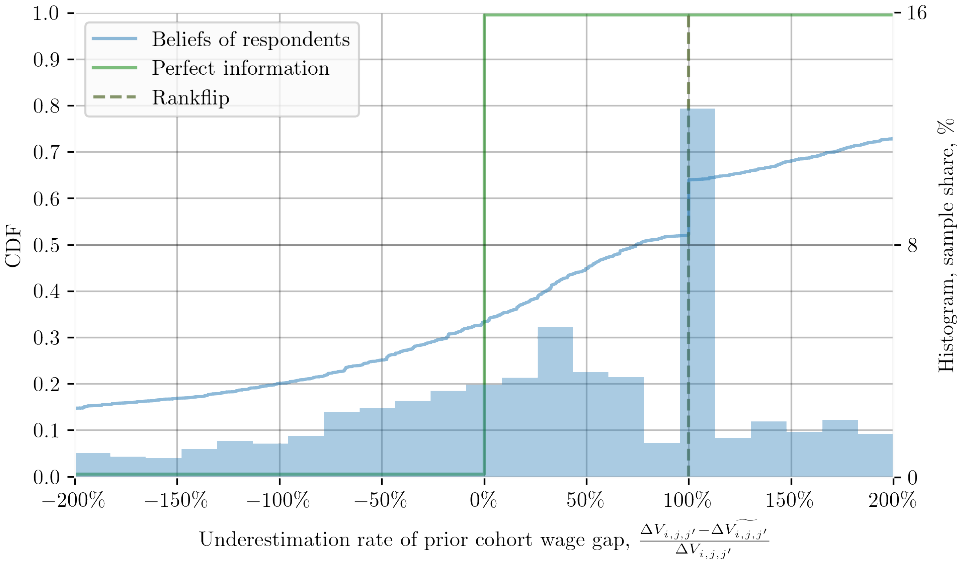
(b) Distribution of misperceptions about individual jobs

Note: The figures are based on all individual-by-jobs (i, j) in the study sample. In Panel (a), the bin scatter splits the data in ten equal-sized bins according to the actual prior cohort wage and then plots within-bin means against each other. Error bars show 95% confidence intervals for the within-bin mean of the perceived (log) prior cohort wage, clustering on individuals. The line shows the fit of an OLS regression. The legend reports the regression line slope as well as the pairwise correlation coefficient. Panel (b) shows a histogram and empirical CDF for the difference in the actual vs. perceived prior cohort wage (in percent of the actual prior cohort wage).

Figure 2: Relative wage misperceptions across job pairs



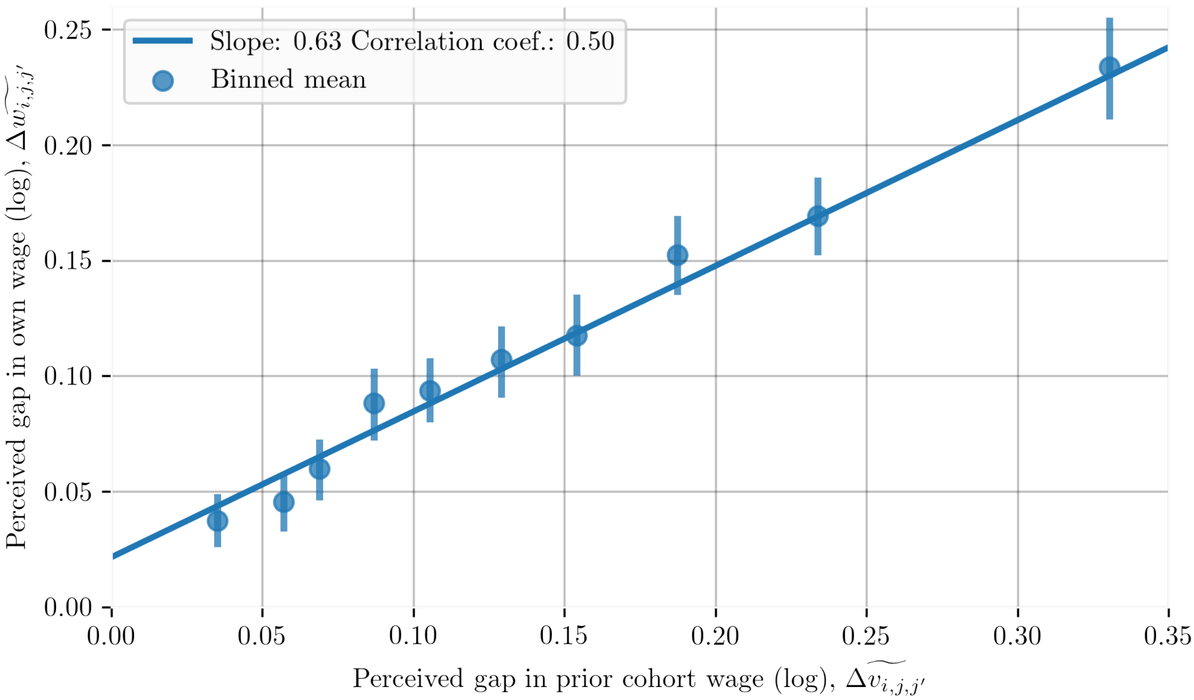
(a) Bin scatter perceived vs actual prior-cohort wage gaps



(b) Distribution of misperceptions about wage gaps

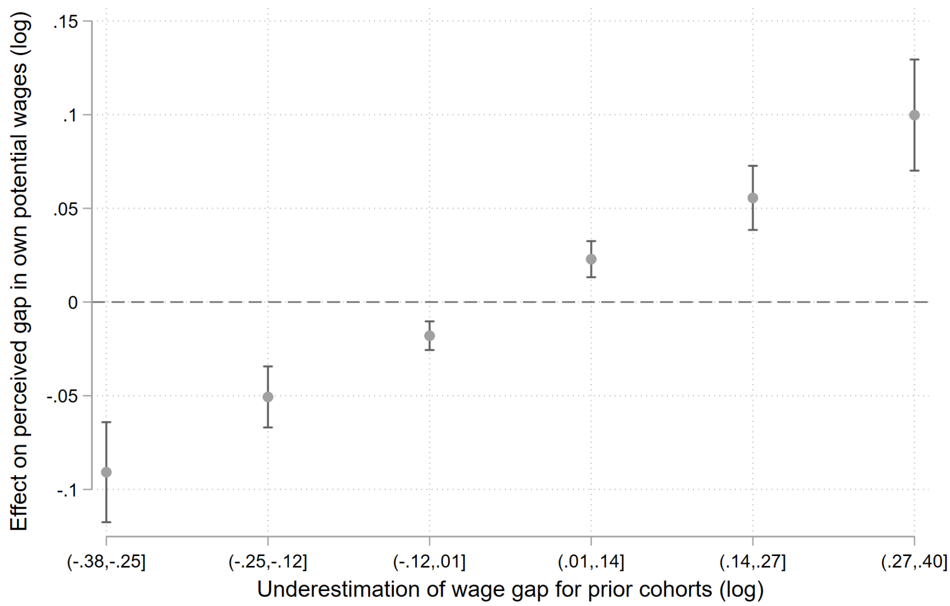
Note: The figure is based on individual-by-job pairs (i, j, j') in the study sample, ordered so that the wage gap for prior cohorts is positive in all cases, $\Delta v_{i,j,j'} > 0$. In Panel (a), the bin scatter splits the data in ten equal-sized bins according to the actual prior cohort wage gap and then plots within-bin means against each other. Error bars show 95% confidence intervals for the within-bin mean of the perceived (log) gap in prior cohort wages, clustering on individuals. The line shows the fit of an OLS regression. The legend reports the regression line slope as well as the pairwise correlation coefficient. Panel (b) shows a histogram and empirical CDF for the underestimation rate of the prior cohort wage gap. The observed mass point at 1 reflects that job seekers report two jobs paying the exact same wage in 11.7% of cases.

Figure 3: Perceived own potential wages vs. perceived prior cohort wages

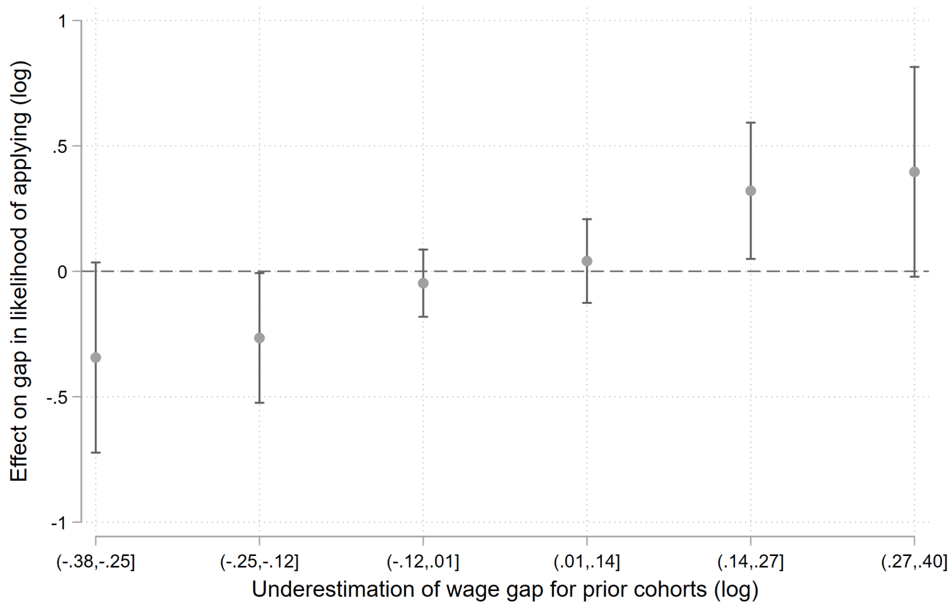


Note: The figure is based on individual-by-job pairs (i, j, j') in the study sample, ordered so that the wage gap for prior cohorts is positive in all cases, $\Delta v_{i,j,j'} > 0$. The bin scatter splits the data in ten equal-sized bins according to the perceived prior cohort wage gap and then plots within-bin means against each other. Error bars show 95% confidence intervals for the within-bin mean of the perceived (log) gap in own potential wages, clustering on individuals. The line shows the fit of an OLS regression. The legend reports the regression line slope as well as the pairwise correlation coefficient.

Figure 4: Heterogeneous effects by size of initial misperceptions



(a) Treatment effect on $\Delta\widetilde{w_{i,j,j'}}$



(b) Treatment effect on $\Delta\pi_{i,j,j'}$

Note: The figure is based on all individual-by-job pairs (i, j, j') in the study sample. The data is split into six equal-width bins according to the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$. The figures plot the effect of the information treatment within each bin, estimated via a linear regression that includes a constant and the treatment dummy. In Panel A, the outcome variable is the perceived gap in (log) own potential wages between the jobs, $\Delta\widetilde{w_{i,j,j'}}$. In Panel B, the outcome variable is the (log) relative likelihood of applying to the two jobs, $\Delta\pi_{i,j,j'}$. Error bars show 95% confidence intervals using clustering on individuals.

Appendices

A Additional tables and figures

Table A1: Descriptives statistics, invited sample, study sample and all young job seekers

	Sample (1) completed survey	Sample (2) invited to survey	Population (3) of young job seekers
Demographic characteristics			
Age	27.97	28.15	28.44
Woman	0.62	0.62	0.59
Danish citizen	0.92	0.91	0.88
Education level			
Less than a Bachelor's degree	0.07	0.11	0.14
Bachelor's degree	0.30	0.42	0.49
Master's degree	0.63	0.46	0.37
Education field			
Teaching and learning	0.09	0.11	0.07
Social sciences	0.18	0.13	0.12
Business economics, administration and law	0.14	0.20	0.25
Science	0.07	0.04	0.03
Information and communication technology (ICT)	0.09	0.07	0.07
Engineering, technology and industrial production	0.08	0.06	0.07
Building and civil engineering	0.03	0.04	0.04
Social and health	0.15	0.22	0.20
Status in 2023			
Graduated less than 1 year ago	0.66	0.44	0.27
Graduated 1-2 years ago	0.06	0.08	0.07
Graduated 2-5 years ago	0.11	0.14	0.14
Graduated less than 1 year after survey	0.04	0.22	0.36
Work experience below 1 year	0.70	0.67	0.62
Work experience of 1-2 years	0.10	0.12	0.13
Work experience of 2-5 years	0.14	0.15	0.17
Work experience above 5 years	0.05	0.06	0.09
Received UI in 2023	0.90	0.89	0.47
Cumulated weeks of UI in 2023	13.39	13.33	7.04
Monthly wage in last job (1,000 DKK)	34.35	34.46	34.47
Lives in Copenhagen	0.27	0.28	0.28
Individuals	1,902	22,461	89,978

Notes: Using a range of variables from administrative data, the table compares the study sample to both the invited sample and the corresponding population of all young job seekers. Column (1) shows variable means for the study sample. Column (2) shows variable means for all individuals invited to the survey. Column (3) shows variable means for all young job seekers from the selected educations, defined as individuals less than 40 years old who either graduated from their degree or were unemployed in 2023.

Table A2: Descriptive statistics, graduates from higher education 2010-2018

	All (%)	**Selected educations (%)**
Education level		
Less than a bachelor's degree	12.3	12.2
Bachelor's degrees	49.5	47.1
Master's degree	35.0	36.4
Ph.d. and research programmes	3.3	4.3
Education field		
Teaching and learning	6.9	8.2
Human sciences	6.9	4.3
Arts	3.0	2.7
Social sciences	9.7	10.6
Business economics, administration and law	22.2	23.7
Science	4.4	2.9
Information and communication technology (ICT)	4.4	5.6
Engineering, technology and industrial production	7.2	7.7
Building and civil engineering	3.7	4.2
Social and health	25.2	24.7
Services	1.6	1.7
Education region		
Copenhagen	43.3	41.7
Zealand	8.6	10.2
Southern Denmark	14.2	13.4
Central Jutland	23.6	24.3
North Jutland	10.2	10.4
Educations	287	88
Individuals	467,267	357,178

Notes: The table shows the composition of all higher education graduates in Denmark 2010-2018, as well as the composition only for graduates from the selected educations included in the study. The first panel shows the shares with different education levels (based on Danish ISCED groupings), the second panel shows the shares in different education fields (based Danish ISCED groupings), the third panel shows the shares in each geographical region of Denmark.

Table A3: Likelihood of transitioning into a job type post-survey vs. vs. planned likelihood of applying for it in the survey, regressions

	(1)	(2)	(3)	(4)
Reported likelihood of applying	0.276***	0.247***	0.779***	0.660***
	(0.018)	(0.022)	(0.039)	(0.050)
Constant	0.029***		0.097***	
	(0.005)		(0.012)	
N	5,706	5,697	1,917	1,881
Education-by-job type fixed effects	No	Yes	No	Yes
Restricted to individuals who found a job in one of the three job types	No	No	Yes	Yes

Notes: The table shows OLS estimates from linked administrative data. The base data are all individual-by-jobs (i, j) in the study sample. The outcome variable is a dummy for whether the individual found a new job of the corresponding type post-survey. The regressor of interest is the reported likelihood of applying to the corresponding job type in the survey. Columns (2) and (4) add fixed effects defined at the level of the respondents' educational background-by-job type (j). Columns (3) and (4) drop individuals who did not find a new job in any of the three survey job types. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table A4: Overall treatment effect on reported job acceptance

	(1)	(2)	(3)	(4)
A): Gap in log likelihood of accepting job offer				
Treatment, T_i	0.222***	0.230***	0.255***	0.255***
	(0.075)	(0.074)	(0.093)	(0.092)
Pre-treatment controls		✓		✓
Currently active job seekers			✓	✓
Individuals	1902	1902	1237	1237
N	5706	5706	3711	3711

Notes: The table shows OLS estimates from a regression including a constant term, the treatment dummy and possibly controls. The base data are individual-by-job pairs (i, j, j') in the study sample, ordered so that the wage gap for prior cohorts is underestimated in all cases, $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}} > 0$. The outcome variable is the gap in the reported log likelihood of accepting an offer from the jobs. Columns (3) and (4) restrict attention only to individual reporting that they are actively searching at the time of the survey. Columns (2) and (4) include as controls all pre-treatment survey variables from Table 1. Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table A5: Effect of relative misperceptions on search, additional robustness

	Reweighted for non-participation		Received and passed validation test		Good understanding of all job types	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\widetilde{w_{i,j,j'}}$	3.762*	3.675**	5.446**	5.384**	2.838	3.636**
	(2.025)	(1.856)	(2.368)	(2.188)	(1.972)	(1.818)
Controls:						
$\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$	✓	✓	✓	✓	✓	✓
Pre-treatment controls		✓		✓		✓
First stage F-stat	48.91	70.42	36.31	38.57	52.17	58.10
Own wage elasticity	2.62	2.56	3.80	3.75	1.97	2.53
Individuals	1826	1826	1019	1019	1135	1135
N	10956	10956	6114	6114	6810	6810

Notes: The table shows additional 2SLS estimates. The base data are all individual-by-job pairs (i, j, j') in the study sample. The outcome variable is the (log) relative likelihood of applying to the two jobs, $\Delta\pi_{i,j,j'}$. The excluded instrument is $T_i \times \left(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right)$. Besides the reported regressors, all specifications include the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$ as a control. Columns (1) and (2) reweighs the study sample to match the invited population using propensity score reweighting. Columns (3) and (4) restricts attention to individuals who were shown the validation test regarding understanding of the survey job types and answered it correctly. Columns (5) and (6) restricts attention to individuals who reported having a good understanding of all the survey job types. Columns (2), (4) and (6) include as additional controls all pre-treatment survey variables from Table 1 interacted with the underestimation of the log wage gap for prior cohorts. The implied own wage elasticity of applications to a job is computed at the bottom via a standard discrete choice framework (see Appendix Section B.4 for details). Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table A6: Effect of relative wage perceptions on search, non-linear controls and controls for differences in other perceived job characteristics

	Gap in log likelihood of applying $\Delta\pi_{i,j,j'}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta\widetilde{w_{i,j,j'}}$	4.285***	4.531***	3.365***	4.052**	4.130***	3.927***	3.445***
	(1.542)	(1.584)	(1.147)	(1.597)	(1.505)	(1.500)	(1.206)
$\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$	-2.705***	-2.229***	-0.800*	-1.804***	-2.047***	-2.161***	-0.710*
	(0.651)	(0.606)	(0.418)	(0.550)	(0.588)	(0.577)	(0.401)
$(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})^2$	-0.000**						
	(0.000)						
$(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}})^3$	4.783						
	(4.110)						
Δ Log probability of offer, if applying		0.602***					0.243***
		(0.034)					(0.026)
Δ Non-pecuniary attractiveness of job			0.690***				0.591***
			(0.016)				(0.017)
Δ Log hours				2.492***			1.233***
				(0.614)			(0.423)
Δ Colleagues					0.475***		0.080***
					(0.033)		(0.024)
Δ Performance						0.570***	0.110***
						(0.024)	(0.019)
First stage F-stat	85.15	75.13	84.80	76.72	85.32	84.91	71.30
Own wage elasticity	2.99	3.15	2.35	2.82	2.88	2.74	2.39
Individuals	1902	1859	1902	1855	1902	1902	1824
N	11412	10762	11412	10998	11412	11412	10500

Notes: The table shows 2SLS estimates with additional controls. The base data are all individual-by-job pairs (i, j, j') in the study sample. The outcome variable is the (log) relative likelihood of applying to the two jobs, $\Delta\pi_{i,j,j'}$. The excluded instrument is $T_i \times \big(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\big)$. Column (1) controls non-linearly for the underestimation of the log wage gap for prior cohorts. Column (2) controls for the perceived gap in the (log) probability of receiving a job offer if applying to the jobs. Column (3) controls for the perceived gap in non-pecuniary attractiveness, defined as the gap in the log reported likelihood of accepting an offer from the jobs in the situation where the jobs offer the exact same starting wage. Column (4) controls for the perceived gap log weekly hours of the jobs. Columns (5) and (6) control for the gap in how well the respondent thinks they would get along with colleagues or perform in the job (6-item Likert scales). Columns (7) combines all job characteristic controls. The implied own wage elasticity of applications to a job is computed at the bottom via a standard discrete choice framework (see Appendix Section B.4 for details). Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table A7: Relationship between jobs' perceived starting wages and other perceived characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\widetilde{w_{i,j,t+5}}$	Hours≤ 35	Hours $\in (35, 40)$	Hours≥ 40	Log hours	Log probability of offer, if applying	Colleagues	Performance
A): Full sample								
Log perceived wage, $\widetilde{w_{i,j}}$	0.837***	-0.363***	-0.099**	0.461***	0.148***	0.385**	0.311***	0.448***
	(0.020)	(0.034)	(0.039)	(0.039)	(0.012)	(0.157)	(0.115)	(0.099)
Constant	✓	✓	✓	✓	✓	✓	✓	✓
Log perceived wage, $\widetilde{w_{i,j}}$	0.818***	-0.333***	-0.557***	0.890***	0.227***	0.101	0.437***	0.737***
	(0.031)	(0.042)	(0.068)	(0.066)	(0.017)	(0.143)	(0.143)	(0.174)
Person FE	✓	✓	✓	✓	✓	✓	✓	✓
Individuals	1887	1902	1902	1902	1855	1859	1902	1902
N	5659	5706	5706	5706	5532	5479	5706	5706
B): Control sample								
Log perceived wage, $\widetilde{w_{i,j}}$	0.865***	-0.378***	-0.108**	0.485***	0.153***	0.145	0.379**	0.363***
	(0.024)	(0.048)	(0.053)	(0.052)	(0.016)	(0.215)	(0.160)	(0.137)
Constant	✓	✓	✓	✓	✓	✓	✓	✓
Log perceived wage, $\widetilde{w_{i,j}}$	0.837***	-0.388***	-0.605***	0.993***	0.250***	0.168	0.558***	0.761***
	(0.044)	(0.058)	(0.090)	(0.089)	(0.023)	(0.197)	(0.183)	(0.233)
Person FE	✓	✓	✓	✓	✓	✓	✓	✓
Individuals	966	976	976	976	946	949	976	976
N	2897	2928	2928	2928	2823	2797	2928	2928

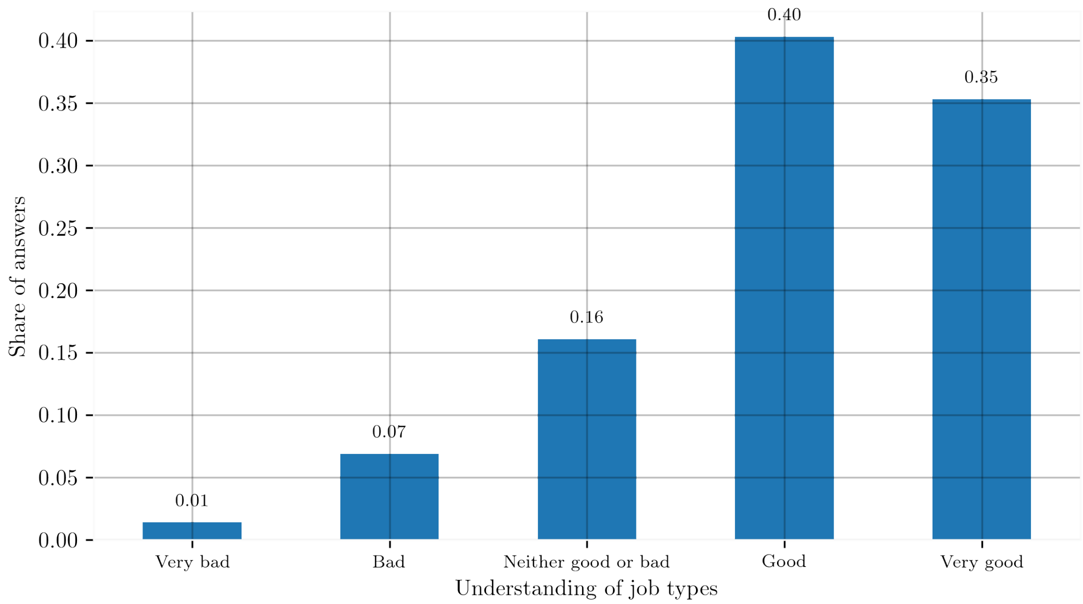
Notes: The table shows OLS estimaes regressing various job characteristics on the log perceived starting wage. The base data are all individual-by-jobs (i, j) in the study sample. Panel A shows results for this full sample, while Panel B restricts attention to the control group. Within each of these panels, the first subpanel shows a specification with just a constant term, while the second subpanel includes individual fixed effects. In Column (1) the outcome variables is the perceived log monthly wage 5 years after starting in the job. In Columns (2), (3) and (4) the outcome variables are dummies for perceived weekly hours being below 35, being 35-40 or being above 40 respectively. In Column (5) the outcome is perceived log weeky hours. In Column (6) the outcome is the perceived log probability of receiving an offer, if applying. In Column (7) the outcome is how well the job seeker expects to get along with their colleagues in the job (6-item Likert scale). In Column (8) the outcome is how well the job seeker expects to perform in the job (6-item Likert scale). Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table A8: Effect of information treatment on the likelihood of finding a new job

	Full sample				Currently active job seekers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment, T_i	0.028	0.015	0.034	0.019	0.030	0.028	0.038	0.030
	(0.022)	(0.021)	(0.027)	(0.026)	(0.027)	(0.027)	(0.033)	(0.033)
$T_i \times$ Overestimator			-0.020	-0.016			-0.025	-0.009
			(0.048)	(0.045)			(0.057)	(0.056)
Overestimator			0.047	0.033			0.064	0.059
			(0.034)	(0.032)			(0.041)	(0.041)
Treatment Effect			0.014	0.004			0.013	0.022
for Overestimators			(0.039)	(0.037)			(0.047)	(0.046)
Controls		✓		✓		✓		✓
N	1,902	1,902	1,902	1,902	1,237	1,237	1,237	1,237

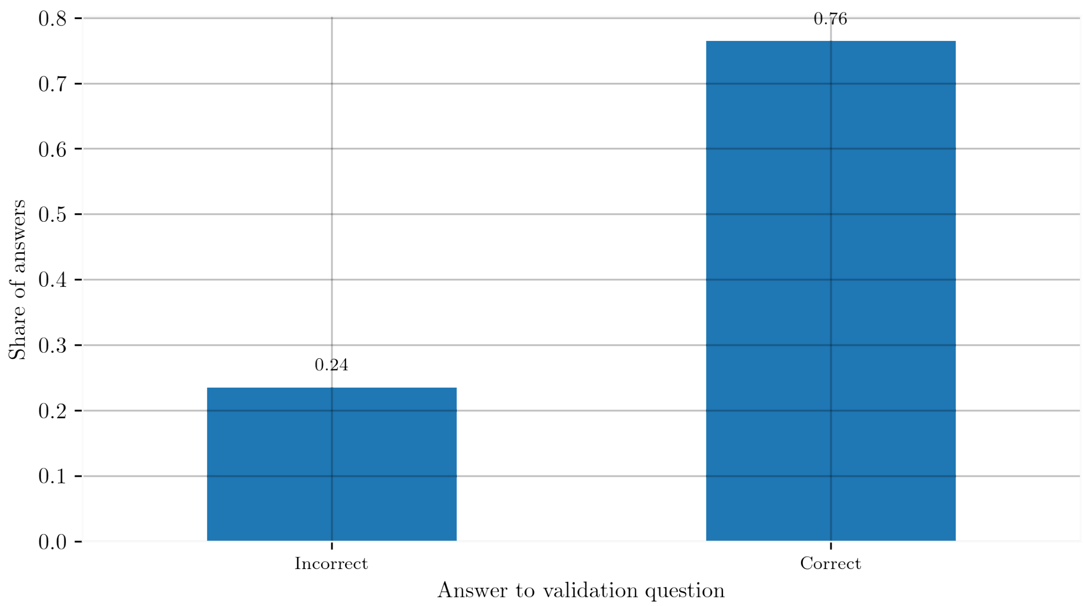
Notes: The table shows OLS estimates from linked administrative data. The base data are all individuals (i) in the study. The outcome variable is a dummy for whether the individual is observed finding a new job post-survey. Columns (1) and (5) contains the most parsimonious regression specifications, including only a constant term and treatment dummy. Other columns additionally add a dummy for the job seeker being an overestimator along with the corresponding interaction with the treatment dummy, and/or add all pre-treatment survey variables from Table 1 as controls. Columns (5)-(8) restrict attention only to individuals reporting that they are actively searching at the time of the survey. In Columns (3), (4), (7) and (8), the implied treatment effect estimate for overestimators is reported in the bottom row along with its standard error. Standard errors in parenthesis are clustered at the level of the individual. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Figure A1: Answers to validation question; 'How good is your understanding of what each of these job types means?'



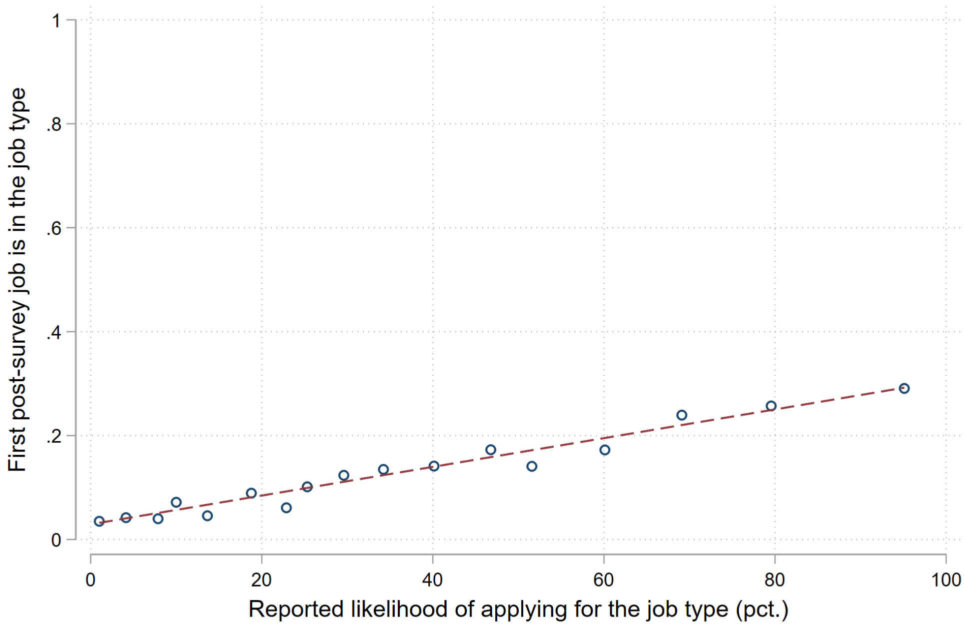
Note: This figure shows a histogram of the answers to the question 'How good is your understanding of what each of these job types means?'.

Figure A2: Answers to validation question; placing an example job in the right job type



Note: A subset of respondents were asked to place a more detailed example job into the correct survey job type (see Section C.3.2). The histogram shows the shares of respondent getting this correct vs. incorrect.

Figure A3: Likelihood of transitioning into a job type post-survey vs. planned likelihood of applying for it in the survey, binscatter

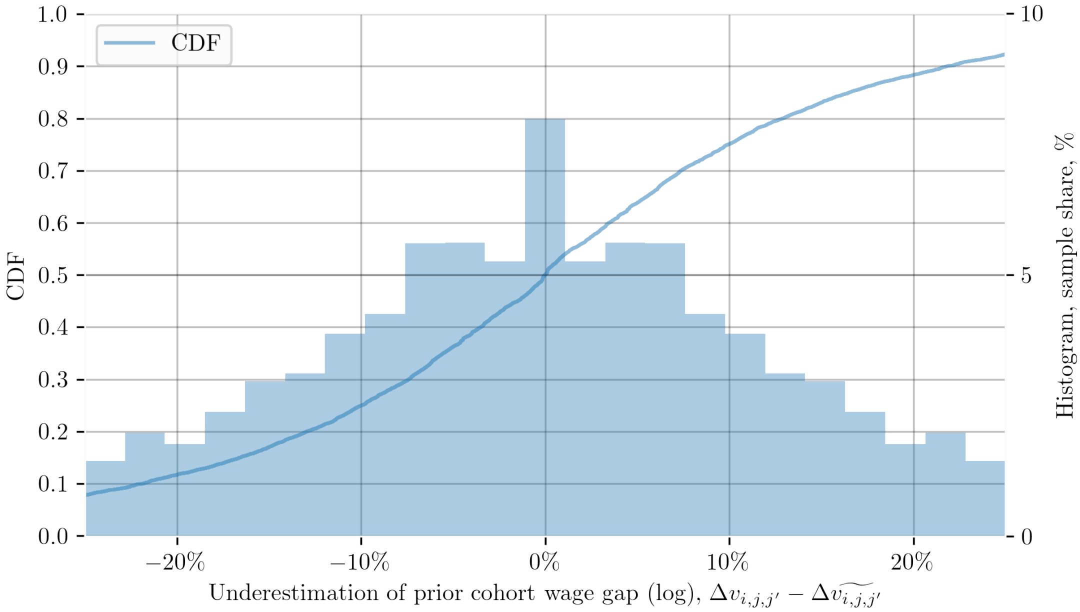


(a) All job seekers

(b) Conditional on finding a job within one of the three job types post-survey

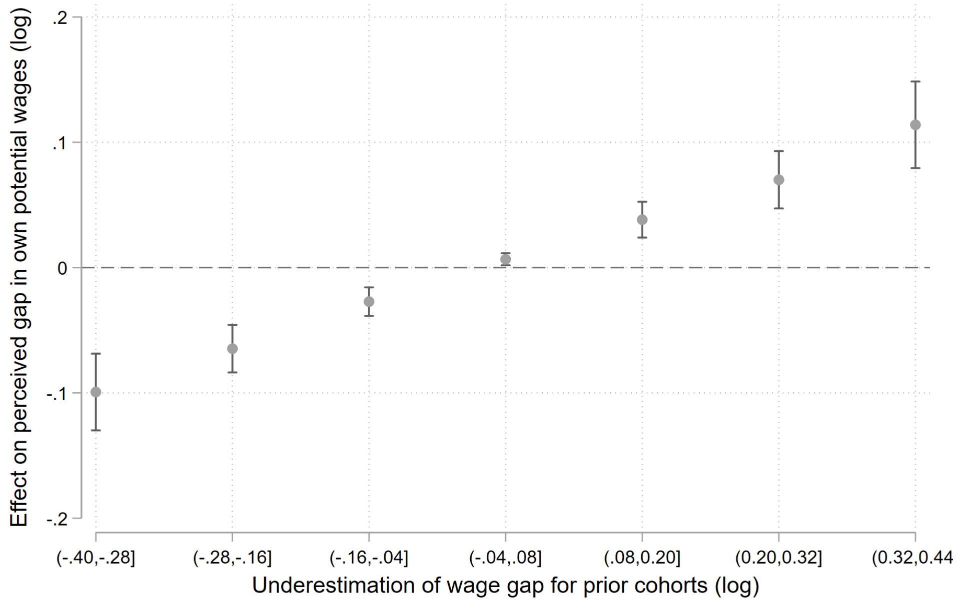
Note: The figures plots the likelihood of finding a new job post-survey in a given job type against the reported likelihood of applying for this job type in the survey. The base data are all individual-by-jobs (i, j) in the study sample. Panel A examines this full sample, while Panel B restricts attention to individuals who are in fact observed finding a new job within one of the three survey job types. The circle plot splits the data into bins according to the reported likelihood of applying for the job type in question and then plots the within-bin share actually finding a new job within this type against the mean reported likelihood of applying for the job type. The line plot shows the fit of a corresponding OLS regression (see Table A3 for corresponding OLS estimates).

Figure A4: Distribution of misperceptions about wage gaps; log wage gaps

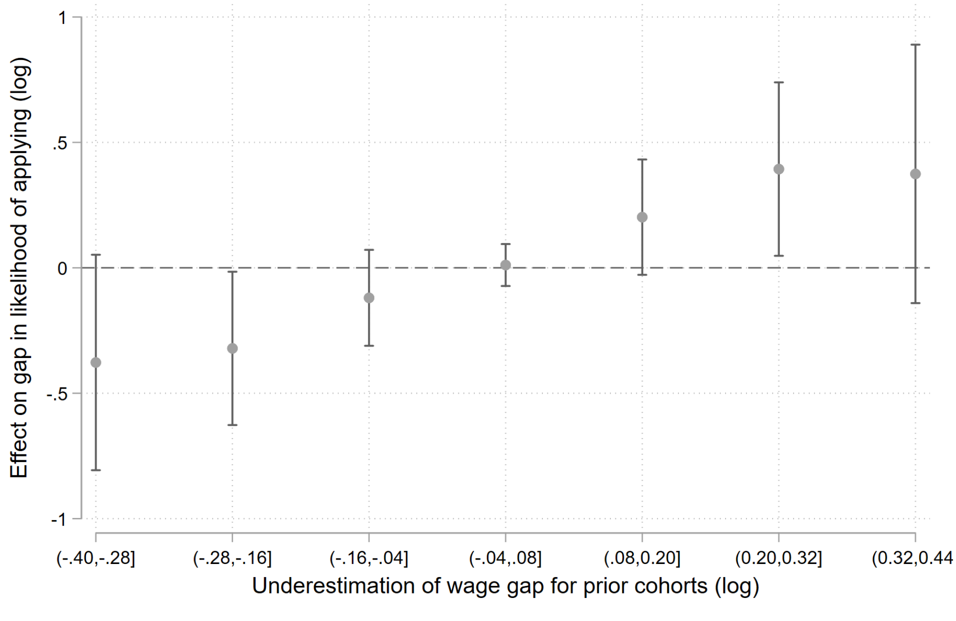


Note: The figure is based on all individual-by-job pairs (i, j, j') in the study sample. The figure shows a histogram and empirical CDF for the underestimation of the prior cohort wage gap in logs, $\Delta v_{i,j,j'} - \Delta \widetilde{v_{i,j,j'}}$.

Figure A5: Heterogeneous effects by size of initial misperceptions, alternative bins 1



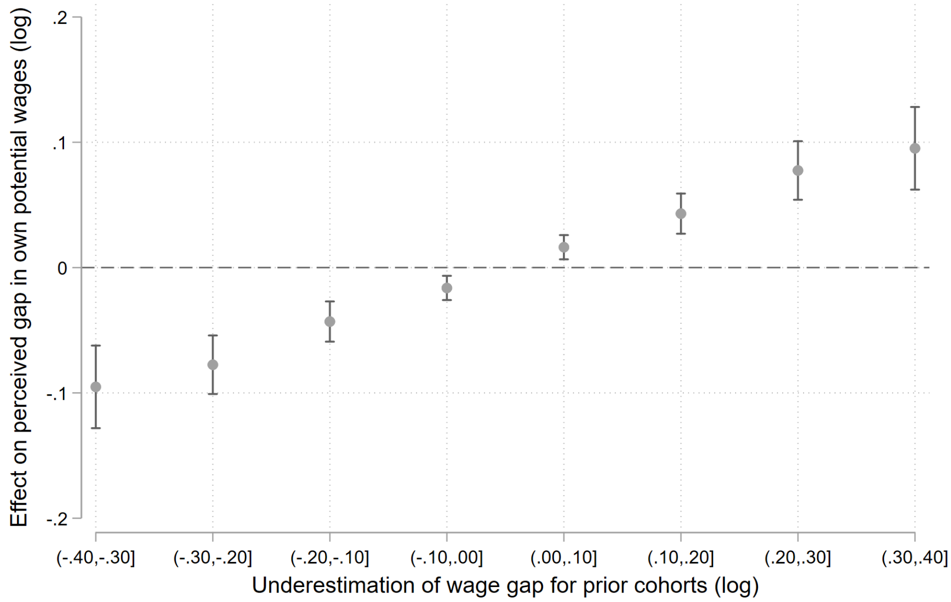
(a) Treatment effect on $\Delta\widetilde{w_{i,j,j'}}$



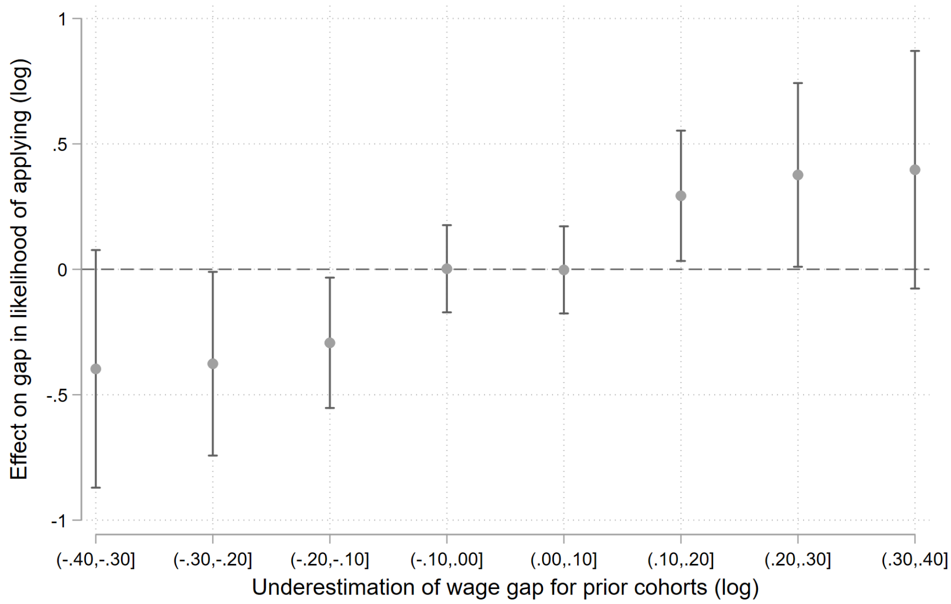
(b) Treatment effect on $\Delta\pi_{i,j,j'}$

Note: These figures examines the sensitivity of Figure 4 to alternative choices of bins. The figure is based on all individual-by-job pairs (i, j, j') in the study sample. The data is split into six equal-width bins according to the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$. The figures plot the effect of the information treatment within each bin, estimated via a linear regression that includes a constant and the treatment dummy. In Panel A, the outcome variable is the perceived gap in (log) own potential wages between the jobs, $\Delta\widetilde{w_{i,j,j'}}$. In Panel B, the outcome variable is the (log) relative likelihood of applying to the two jobs, $\Delta\pi_{i,j,j'}$. Error bars show 95% confidence intervals using clustering on individuals.

Figure A6: Heterogeneous effects by size of initial misperceptions, alternative bins 2



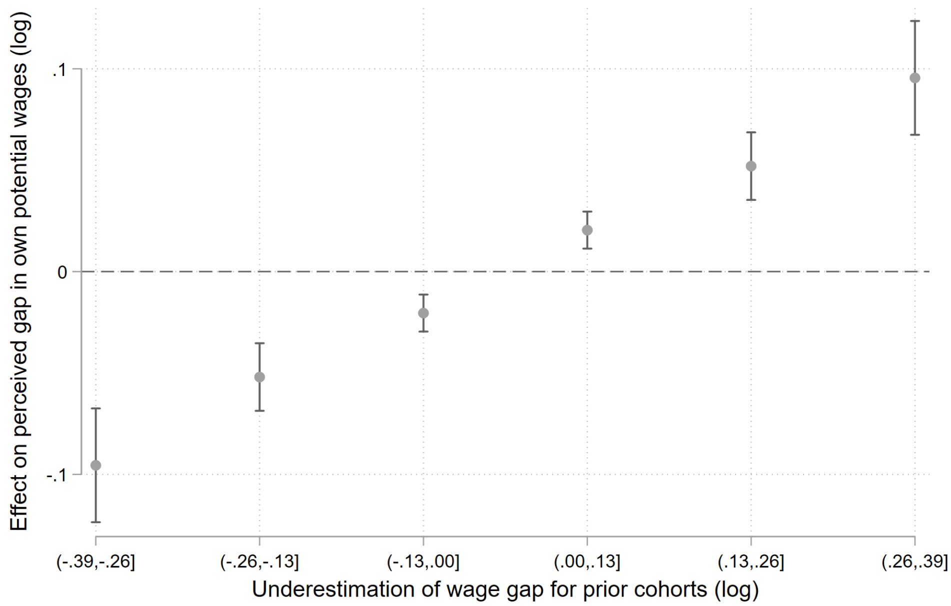
(a) Treatment effect on $\Delta\widetilde{w_{i,j,j'}}$



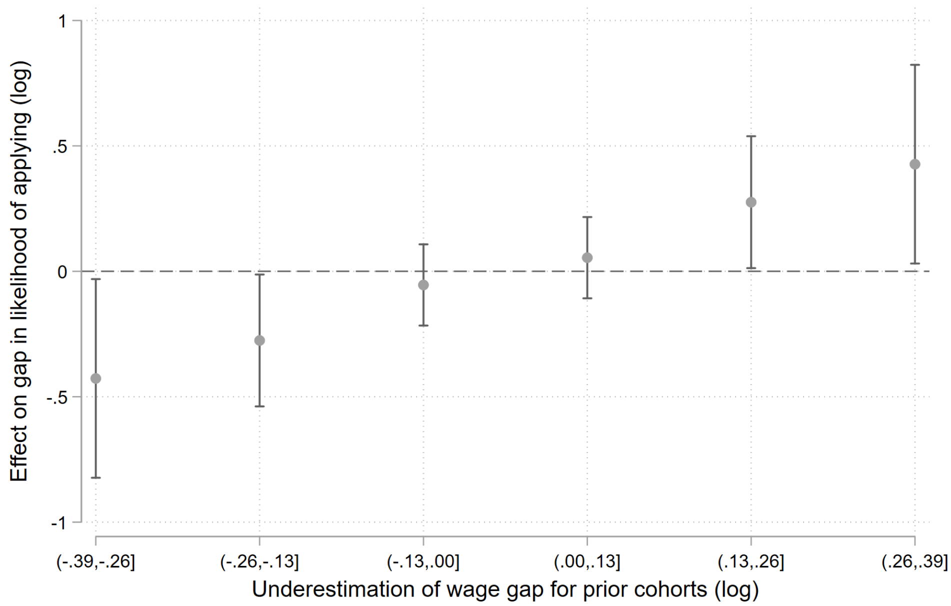
(b) Treatment effect on $\Delta\pi_{i,j,j'}$

Note: These figures examines the sensitivity of Figure 4 to alternative choices of bins. The figure is based on all individual-by-job pairs (i, j, j') in the study sample. The data is split into six equal-width bins according to the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$. The figures plot the effect of the information treatment within each bin, estimated via a linear regression that includes a constant and the treatment dummy. In Panel A, the outcome variable is the perceived gap in (log) own potential wages between the jobs, $\Delta\widetilde{w_{i,j,j'}}$. In Panel B, the outcome variable is the (log) relative likelihood of applying to the two jobs, $\Delta\pi_{i,j,j'}$. Error bars show 95% confidence intervals using clustering on individuals.

Figure A7: Heterogeneous effects by size of initial misperceptions, alternative bins 3



(a) Treatment effect on $\Delta\widetilde{w_{i,j,j'}}$



(b) Treatment effect on $\Delta\pi_{i,j,j'}$

Note: These figures examines the sensitivity of Figure 4 to alternative choices of bins. The figure is based on all individual-by-job comparisons (i, j, j') in the study sample. The data is split into six equal-width bins according to the underestimation of the log wage gap for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$. The figures plot the effect of the information treatment within each bin, estimated via a linear regression that includes a constant and the treatment dummy. In Panel A, the outcome variable is the perceived gap in (log) own potential wages between the jobs, $\Delta\widetilde{w_{i,j,j'}}$. In Panel B, the outcome variable is the (log) relative likelihood of applying to the two jobs, $\Delta\pi_{i,j,j'}$. Error bars show 95% confidence intervals using clustering on individuals.

B Discrete choice job search framework

This section contains additional details and derivations regarding the discrete job search framework used in the main text.

B.1 Model setup, recap

The full model setup and notation is described in Section 5 of the main text. We briefly recap the key elements here for completeness. At a point in time, job seeker i has the opportunity to apply for one of three different jobs indexed by $j = 1, 2, 3$. The actual surplus value of applying to job j relative to continued search are assumed to be

$$A_{i,j} = P_{i,j}\big(\Psi_i W_{i,j}^{\delta} Z_{i,j}\big) \tag{B.1}$$

while the job seeker's perceived surplus value of applying is assumed to be

$$\widetilde{A_{i,j}} = \widetilde{P_{i,j}}\left(\Psi_i \widetilde{W_{i,j}}^{\delta} \widetilde{Z_{i,j}} \xi_{i,j}\right) - \xi_{i,0} \tag{B.2}$$

In these expressions, $W_{i,j}$ and $\widetilde{W_{i,j}}$ are the actual and perceived wages for worker i in job j, $Z_{i,j}$ and $\widetilde{Z_{i,j}}$ are the actual and perceived non-wage amenity values and $P_{i,j}$, and $\widetilde{P_{i,j}}$ are the actual and perceived likelihood of being hired if applying. The parameters δ and Ψ_i are preference parameters governing the role of wages in utility and the (individual-specific) general value of employment over continued search. Finally, $\xi_{i,j}$ is an idiosyncratic shock to i's perceived surplus value of job j, while $\xi_{i,0}$ is an idiosyncratic shock to the perceived costs of sending an application. The shocks are assumed to be i.i.d. (over all i and over $j = 0, 1, 2, 3$) and to follow a Frechet distribution with shape parameter σ^{-1}.

In what follows, we adopt the convenient notation that $A_{i,0}$ and $\widetilde{A_{i,0}}$ denotes the actual and perceived surplus value of choosing not to apply anywhere. Since not applying anywhere guarantees that the worker continues searching, the surpluses $A_{i,0}$ and $\widetilde{A_{i,0}}$ are both mechanically 0.

B.2 Choice probabilities and the value of an application

Let j_i^* denote the actual application choice that the worker makes, with $j_i^* = 0$ corresponding to the choice to not apply anywhere. Optimizing behavior implies that the worker chooses the option that offers the highest perceived surplus:

$$j_i^* = \text{argmax}_j \widetilde{A_{i,j}}$$

Since the optimal choice is unaffected by monotone transformations of the objective, we can rewrite this as:

$$j_i^* = \text{argmax}_j \quad \log\left(\widetilde{A_{i,j}} + \xi_{i,0}\right) \tag{B.3}$$

To see why this is useful, note that for $j = 1, 2, 3$ we have (using lowercase to denote logged values as in the main text)

$$\log\left(\widetilde{A_{i,j}} + \xi_{i,0}\right) = \psi_i + \widetilde{p_{i,j}} + \delta\widetilde{w_{i,j}} + \widetilde{z_{i,j}} + \log \xi_{i,j} \tag{B.4}$$

while for $j = 0$

$$\log\left(\widetilde{A_{i,0}} + \xi_{i,0}\right) = \log \xi_{i,0} \tag{B.5}$$

Under the assumed Frechet distribution, $\log \xi_{i,j}$ follows a Type I extreme value distribution with scale parameter σ. Equations (B.3), (B.4) and (B.5) therefore correspond to a standard discrete choice logit structure. As useful notation here, we let $\Pi_{i,j}^* = \Pr(j_i^* = j)$ be the probability of choosing alternative j in this problem. For $j = 1, 2, 3$ we thus have the following standard form for the choice probabilities:

$$\Pi_{i,j}^* = \frac{\exp((\psi_i + \widetilde{p_{i,j}} + \delta\widetilde{w_{i,j}} + \widetilde{z_{i,j}})/\sigma)}{1 + \sum_{j'=1,2,3} \exp((\psi_i + \widetilde{p_{i,j'}} + \delta\widetilde{w_{i,j'}} + \widetilde{z_{i,j'}})/\sigma)} \tag{B.6}$$

This is equation (9) from the main text. For $j = 0$ we have

$$\Pi^*_{i,0} = \frac{1}{1 + \sum\limits_{j'=1,2,3} \exp((\psi_i + \widetilde{p_{i,j'}} + \delta\widetilde{w_{i,j'}} + \widetilde{z_{i,j'}})/\sigma)} \tag{B.7}$$

For much of our analysis, we will be focusing on the likelihood that a given application goes to a particular job, i.e., the likelihood of applying to a given job j conditional on applying *somewhere*. We denote this by $\overline{\Pi}^*_{i,j}$:

$$\overline{\Pi}^*_{i,j} = \frac{\Pi^*_{i,j}}{1 - \Pi^*_{i,0}}$$

Finally, for the purpose of the welfare counterfactual later, we let Ω_i be the total value of making an application. By definition the total value of applying for job j is $A_{i,j} + U_i$ so we have:

$$\Omega_i = \sum_{j=1,2,3} \overline{\Pi}^*_{i,j} \times (A_{i,j} + U_i) \tag{B.8}$$

B.3 Link to data, 2SLS analysis and exclusion restriction

The model framework above links naturally to the survey data we use in our reduced form analysis. Elicited beliefs about the potential wage each job in the survey would offer the respondent, is a direct measure of $\widetilde{W_{i,j}}$. Similarly, the reported probabilities of applying to each job in the survey, corresponds to the choice probabilities in equation (B.6) and (B.7). As convenient notation, we let $\Pi_{i,j}$ (without an asterisk) denote the reported probabilities in the data.[37]

Taking logs and differencing (B.6) across two jobs j and j' now yields the equation (with Δ denoting differences across pairs of jobs, as in the main text):

$$\Delta\pi_{i,j,j'} = \beta_1 \Delta\widetilde{w_{i,j,j'}} + \Delta u_{i,j,j'} \tag{B.9}$$

[37] We do this to distinguish the data from the endogenous objects in the model. The distinction becomes relevant when we turn to consider counterfactuals further below.

Here we have defined the random variable (error term) $\Delta u_{i,j,j'}$ as

$$\Delta u_{i,j,j'} = \frac{1}{\sigma}\left(\Delta\widetilde{p_{i,j,j'}} + \Delta\widetilde{z_{i,j,j'}}\right)$$

Additionally, we have defined the key reduced form parameter β_1 as

$$\beta_1 = \frac{\delta}{\sigma} \tag{B.10}$$

As emphasized in the main text, (B.9) implies that β_1 has the reduced form interpretation of being the causal effect of the wage gap between two jobs on the relative likelihood of applying to them.[38] Specifically, β_1 is the elasticity of the relative application likelihood between two jobs ($\Pi_{i,j}/\Pi_{i,j'}$) with respect to the percieved relative wage offers ($\widetilde{W}_{i,j}/\widetilde{W}_{i,j'}$). As shown in (B.10), however, β_1, also has a direct, formal mapping to key primitive parameters in the model framework. We use this further below to compute model counterfactuals from the reduced form estimates.

Equation (B.9) and the model framework can also be used to formally discuss the role played by the random information treatment and instrumental variable assumptions in our 2SLS results. The error term in (B.9), $\Delta u_{i,j,j'}$, reflects beliefs about all non-pecuniary differences that affect the attractiveness of applying to job j versus j', specifically the perceived difference in application success probability, $\Delta\widetilde{p_{i,j,j'}}$ and in the non-wage amenities, $\Delta\widetilde{z_{i,j,j'}}$. In general, we expect job seekers with particular beliefs about wage differences, to likely also have particular beliefs about differences in these other dimensions. This leads to a classic identification problem because the regressor in (B.9) is correlated with the error term.

We address this identification problem by leveraging our randomized treatment assignment, T_i, in a 2SLS framework. As covered in Section 4 of the main text, for a given level of the wage gap underestimation for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$, our information treatment generates variation in perceived own potential wages, $\Delta\widetilde{w_{i,j,j'}}$. This means that it can be used to construct a relevant instrument. Moreover, due to random assignment, treated

[38]To be precise here, under the model, β_1 is a causal effect in the sense that it measures the effect of changes in perceived wages, keeping everything else constant.

and untreated job seekers cannot have systematically different beliefs prior to treatment. Under the additional exclusion restriction that the information treatment also does not change beliefs about non-wage amenities or the application success probability, treatment will be independent of non-pecuniary beliefs, both overall and for a given level of the wage gap underestimation for prior cohorts, $\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}$:

$$T_i \perp \left(\Delta\widetilde{p_{i,j,j'}}, \Delta\widetilde{z_{i,j,j'}}\right)\Big|(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) \tag{B.11}$$

As covered in Section 4.3, we use a range of additional questions in our survey to test this exclusion restriction directly and find it to be well-supported by the data. If we finally impose a standard 'linearity-of-controls' functional form assumption, we arrive at the second stage equation and instrument validity condition used in our 2SLS analysis:[39]

$$\Delta\pi_{i,j,j'} = \beta_1\, \Delta\widetilde{w_{i,j,j'}} + \beta_2\,(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}) + \varepsilon_{i,j,j'} \tag{B.12}$$

$$E[\varepsilon_{i,j,j'}|T_i, \Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}] = 0 \tag{B.13}$$

B.4 Own-wage Elasticity of Applications

As discussed above, the parameter of interest in our 2SLS analysis, β_1, measures the effect of the perceived *wage gap* between two jobs on the *relative likelihood* of applying to them. Using the discrete choice framework above, however, we can convert this into an own wage elasticity of applications to a given job, i.e., a measure of how much more likely job seekers

[39]The usual functional form assumption imposes that in the absence of treatment, the conditional expectation of the error term is linear in the conditioning variable:

$$E\left[\Delta u_{i,j,j'}|T_i = 0, \left(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right)\right] = \beta_2\left(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right)$$

With this, the second stage equation then arises simply by appropriately defining its error term:

$$\varepsilon_{i,j,j'} = \Delta u_{i,j,j'} - E\left[\Delta u_{i,j,j'}|\left(\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right), T_i = 0\right]$$

The instrument validity condition follows from (B.11) because $\Delta u_{i,j,j'} = \frac{1}{\sigma}\left(\Delta\widetilde{p_{i,j,j'}} + \Delta\widetilde{z_{i,j,j'}}\right)$:

$$\begin{aligned}
&E\left[\varepsilon_{i,j,j'}|T_i, \Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right] = \\
&E\left[\Delta u_{i,j,j'}|T_i, \Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right] - E\left[\Delta u_{i,j,j'}|T_i = 0, \Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right] = \\
&E\left[\Delta u_{i,j,j'}|\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right] - E\left[\Delta u_{i,j,j'}|\Delta v_{i,j,j'} - \Delta\widetilde{v_{i,j,j'}}\right] = 0
\end{aligned}$$

are to apply for a given job when this job unilaterally increases its wage. This is a key parameter in the literature on job search and employer market power.

For person i and job j the own wage elasticity is simply defined as:

$$\mathcal{E}_{i,j} = \frac{\partial \ln \Pi^*_{i,j}}{\partial \widetilde{w_{i,j}}}$$

Standard derivations using the choice probability formula (B.6) yields:

$$\mathcal{E}_{i,j} = (1-\Pi_{i,j})\beta_1 \tag{B.14}$$

To arrive at an aggregate elasticity over the sample, we plug in our estimate for β_1 as well as the observed choice probability $\Pi_{i,j}$, and then average over the N job seekers in our data and the three jobs they each face:

$$\widehat{\mathcal{E}}^a = \frac{1}{N}\sum_i \left(\frac{1}{3} \sum_{j=1,2,3} \hat{\beta}_1(1-\Pi_{i,j}) \right) \tag{B.15}$$

B.5 Counterfactuals without wage misperceptions

In Section 5 of the main text, we combine our reduced form estimates and the discrete choice framework to assess the overall costs of relative wage misperceptions. Specifically, for our control group of untreated individuals ($T_i = 0$), we compare actual current outcomes to a counterfactual in which all relative wage misperceptions are removed. The subsections below go through the key derivations and additional assumptions underlying this counterfactual.

For the purpose of the describing the counterfactual, we first modify and introduce additional notation relative to the preceding sections. Let $\boldsymbol{X} = (X_1, X_2, X_3)$ and $\widetilde{\boldsymbol{X}} = (\widetilde{X_1}, \widetilde{X_2}, \widetilde{X_3})$ be arbitrary vectors of wages for the three jobs a given worker faces. Chiefly, we are interested in examining how application choices would be different under different beliefs; we thus let $\Pi^*_{i,j}\left(\widetilde{\boldsymbol{X}}\right)$ and $\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{X}}\right)$ denote application choice probabilities if the worker perceives the wages to be $\widetilde{\boldsymbol{X}}$. We are also interested in examining employment outcomes; we similarly let $h_{i,j}\left(\widetilde{\boldsymbol{X}}\right)$ denote the job finding hazard rate of job seeker i

into job job j. Finally, we consider counterfactuals for the utility value of making an application. We let $\Omega_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}})$ denote the utility value of making an application if the wages are actually $\boldsymbol{X}$ but the worker perceives the wages to be $\widetilde{\boldsymbol{X}}$. In what follows, we additionally use $U_i\Big(\boldsymbol{X}, \widetilde{\boldsymbol{X}}\Big)$ and $A_{i,j}\Big(\boldsymbol{X}, \widetilde{\boldsymbol{X}}\Big)$ for the corresponding continuation value of unemployment and surplus value of applying to a given job.

In terms of the specific counterfactual comparison we are interested in, we let $\boldsymbol{W_i} = (W_{i,1}, W_{i,2}, W_{i,3})$ be a vector of actual wages in the three jobs for person i and $\widetilde{\boldsymbol{W}_i^0} = (\widetilde{W_{i,1}^0}, \widetilde{W_{i,2}^0}, \widetilde{W_{i,3}^0})$ be their vector of currently perceived wages. We are then interested in computing outcomes under a counterfactual set of wage beliefs $\widetilde{\boldsymbol{W}_i^1} = (\widetilde{W_{i,1}^1}, \widetilde{W_{i,2}^1}, \widetilde{W_{i,3}^1})$ in which there is no relative wage misperception. This corresponds to $\widetilde{\boldsymbol{W}_i^1}$ satisfying:

$$\widetilde{\boldsymbol{W}_i^1} = \iota_i \boldsymbol{W}_i \tag{B.16}$$

for some $\iota_i > 0$ which we leave unrestricted. Note that the constant ι_i here can be interpreted as the extend of overestimation/underestimation of the general level of wages. As will become clear below, our counterfactual calculation is invariant to ι_i, reflecting that our counterfactual, deliberately only captures the effects of removing *relative* wage misperceptions, not misperceptions about the overall level of wages.

B.5.1 Counterfactual application behavior and applied-for wages

The first counterfactual of interest regards how application behavior would change if job seeker i did not have any misperceptions about relative wages. Specifically, we want to compare the likelihood that an application currently goes to a particular job, $\overline{\Pi}^*_{i,j}\Big(\widetilde{\boldsymbol{W}_i^0}\Big)$ to the counterfactual likelihood without relative wage misperceptions $\overline{\Pi}^*_{i,j}\Big(\widetilde{\boldsymbol{W}_i^1}\Big)$. Given that we are focusing on individuals in the control group, of course, $\overline{\Pi}^*_{i,j}\Big(\widetilde{\boldsymbol{W}_i^0}\Big)$ can be directly computed from the survey responses about likelihood of applying to the different jobs:

$$\overline{\Pi}^*_{i,j}\Big(\widetilde{\boldsymbol{W}_i^0}\Big) = \frac{\Pi_{i,j}}{1 - \Pi_{i,0}}$$

Further note that *if* we knew the current perceived and actual wages, $\widetilde{\boldsymbol{W}_i^0}$ and $\boldsymbol{W_i}$, the only thing necessary to compute $\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}^1_i}\right)$ is the reduced form parameter β_1 since (B.6) and (B.16) imply:

$$\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}^1_i}\right) = \frac{\left(\frac{W_{i,j}}{\widetilde{W^0_{i,j}}}\right)^{\beta_1}\frac{\Pi_{i,j}}{\Pi_{i,0}}}{\sum_{j'=1,2,3}\left(\frac{W_{i,j}}{\widetilde{W^0_{i,j'}}}\right)^{\beta_1}\frac{\Pi_{i,j'}}{\Pi_{i,0}}}$$

For the control group of course we in fact observe $\widetilde{W^0_{i,j}}$ directly from the survey questions:

$$\widetilde{W^0_{i,j}} = \widetilde{W_{i,j}}$$

To compute the counterfactual however, we still need to take a stand, on how actual potential wage differences differ from perceptions to get a handle on $\boldsymbol{W_i}$. Given the content of our survey, our benchmark will be to simply equate misperceptions about own potential (log) wage gaps to the measured misperceptions about wage gaps for prior cohorts from the survey, i.e. we assume:

$$\Delta v_{i,j,j'} - \widetilde{\Delta v_{i,j,j'}} = \Delta w_{i,j,j'} - \widetilde{\Delta w^0_{i,j,j'}} \tag{B.17}$$

Since this additionally allows us to infer $\boldsymbol{W}_i$, imposing B.17 is enough for us to compute the counterfactual application likelihood $\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}^1_i}\right)$ using our estimate for β_1 and the equation further above.

Having computed the counterfactual application probabilities, our first measure of interest is the share of i's applications that would go to different jobs in the absence of relative wage misperceptions. We summarize this by computing the predicted total change in the shares of applications going to the different jobs:[40]

$$\frac{1}{2}\sum_{j=1,2,3}\left|\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}^1_i}\right) - \overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}^0_i}\right)\right|$$

[40] Invoking a law of large numbers, this measure formally measures the share of applications that will have gone to a different job, after i have made a large number of applications. A different summary measure would be to compute the probability that a given single application goes to a different job in the absence of relative misperceptions. Due to decision uncertainty from the taste shocks however, this measure would be non-zero even if relative misperceptions had *no* effects on behavior.

As another useful summary measure, we also examine the growth in the average wage that individual i would have applied for if wage relative misperceptions were removed:

$$\frac{\sum_{j=1,2,3} \overline{\Pi}^{*}_{i,j}\left(\widetilde{\boldsymbol{W}^{1}_{i}}\right)W_{i,j} - \sum_{j=1,2,3} \overline{\Pi}^{*}_{i,j}\left(\widetilde{\boldsymbol{W}^{0}_{i}}\right)W_{i,j}}{\sum_{j=1,2,3} \overline{\Pi}^{*}_{i,j}\left(\widetilde{\boldsymbol{W}^{0}_{i}}\right)W_{i,j}}$$

B.5.2 Counterfactual employment outcomes and wages

The previous section discusses counterfactuals in application behavior. Of course, not all applications result in a hire, so another interesting counterfactual concerns how actual hiring outcomes and wages would change in the absence of wage misperceptions. To compute meaningful counterfactuals for these, we must impose some additional assumptions.

First, we need to take a stand on the likelihood that a given application results in the worker being hired. As our benchmark we simply use the worker's own reported likelihood of receiving an offer. As part of our survey, job seekers reported their own perceived likelihood that an application from them result in an offer for each of the three jobs j. We let $\widetilde{P_{i,j}}$ denote this perceived application success probability. We use this as the actual probability of being hired in our counterfactual:[41]

$$P_{i,j} = \widetilde{P_{i,j}} \tag{B.18}$$

Note here, that by imposing a fixed application success probability regardless of beliefs, our counterfactual rules out any general equilibrium effects where changes in beliefs and application behavior lead to changes in the likelihood of being hired when applying. Accounting for such equilibrium effects would require us to take a stance on how firms' vacancy postings and hiring decisions respond to changes in the number of applications. Since our data contains no direct information on this, we instead view assumption (B.18) as the most transparent and simple benchmark.

Additionally, to characterize eventual hiring outcomes, we need to take a stand on

[41] Given the additional assumptions discussed below, the results of our counterfactual calculations depend only on the *relative* likelihood of being hired in the different jobs. Our counterfactual calculation is thus unaffected by job seekers systematically overestimating or underestimating all application success probabilities.

what happens for those workers for whom the (first) application does not result in a hire. Mirroring standard job search models, we simply assume that such workers continue searching and make an application again with some fixed, individual-specific Poisson arrival rate, λ_i, while facing a stationary continuation value. The translates to the hazard rate of worker i into job j (under beliefs $\widetilde{\boldsymbol{X}}$) being

$$h_{i,j}\left(\widetilde{\boldsymbol{X}}\right) = \lambda_i P_{i,j} \overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{X}}\right) \tag{B.19}$$

Given the additional assumptions (B.18) and (B.19), we can compute the likelihood that worker i ends up in job j (conditional on being hired somewhere). Under current beliefs it is

$$\frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^0}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^0}\right)}$$

while in the counterfactual without relative wage misperceptions it is

$$\frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^1}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^1}\right)}$$

As our first summary measure comparing the two hiring likelihoods, we are interested in measuring how likely it is that worker i would shifted to a different job if relative wage misperceptions did not exist. Analogous to above, we summarize this by computing the predicted total change in the hiring shares into the different jobs:

$$\frac{1}{2}\sum_{j=1,2,3}\left|\frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^1}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^1}\right)} - \frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^0}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^0}\right)}\right|$$

Additionally, we compute the growth in the expected reemployment wage for person i

after removing relative wage misperceptions:

$$\frac{\sum\limits_{j=1,2,3} \frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^1}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^1}\right)} W_{i,j} - \sum\limits_{j=1,2,3} \frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^0}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^0}\right)} W_{i,j}}{\sum\limits_{j=1,2,3} \frac{P_{i,j}\overline{\Pi}^*_{i,j}\left(\widetilde{\boldsymbol{W}_i^0}\right)}{\sum\limits_{j'=1,2,3} P_{i,j'}\overline{\Pi}^*_{i,j'}\left(\widetilde{\boldsymbol{W}_i^0}\right)} W_{i,j}}$$

B.5.3 Utility gains

Finally, leaning further into the model structure, we can also use this framework to assess the welfare costs of relative wage misperceptions.

Specifically, we consider how the removal of relative wage misperceptions would change the total utility value of sending an application, i.e. how much does $\Omega_i(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^1})$ differ from $\Omega_i(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^0})$. To make this comparison on a meaningful scale, we ask how much all wages would have to increase in order to deliver the same welfare gain as removing relative wage misperceptions. If all wages for i are scaled up by k_i, the value of making application will be $\Omega_i(k_i\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^0})$. Accordingly, we compute the value k_i that satisfies:

$$\Omega_i(k_i\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^0}) = \Omega_i(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^1}) \tag{B.20}$$

To do this requires us to impose three additional assumptions. First, we need to take a stand on the continuation value of not getting a job, U_i. In line with assumptions in the previous section, we assume that the continuation value reflects that the job seeker will continue searching while earnings some exogenous flow utility b_i, making an application again at an exogenous Poisson rate, λ_i, and time discounting at rate ρ_i. This implies the following continuous time Bellman equation:

$$\rho_i U_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}}) = b_i + \lambda_i \Big(\Omega_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}}) - U_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}}) \Big) \tag{B.21}$$

Second, to assess actual utilities, we also need to take a stand on the relationship between perceived and actual non-wage amenities. Here we simply adopt the benchmark assumption that workers have correct beliefs about non-wage amenities in the different jobs:

$$Z_{i,j} = \widetilde{Z_{i,j}} \tag{B.22}$$

Finally, we need to impose a parameter restriction on the surplus utility from jobs. This reflects the usual challenge that choice behavior does not separately identify the scale of the taste shocks.[42] For robustness, we consider three different parameter restrictions and compute our welfare counterfactual under each of them. The first possible restriction simply imposes unit scale of the taste shocks as is sometimes done in applications:

$$\sigma = 1 \tag{B.23}$$

The second possible restriction we consider instead normalizes that the surplus value of jobs is linear in the wage:

$$\delta = 1 \tag{B.24}$$

Finally, we consider a third possible restriction that calibrates the scale of the taste shocks, σ, based on the cross-sectional relationship between application likelihoods and perceived application success probabilities. Plugging in the definition of the error term in (B.9) shows that the perceived application success probability, $\widetilde{p_{i,j}}$, enters our 2SLS second stage with a coefficient of σ^{-1}. Accordingly, in Column (2) of Table A6 we add $\widetilde{p_{i,j}}$ as a control variable to our 2SLS specification and estimate its coefficient to be $0.602 \approx 1.66^{-1}$. This suggests the following value for the scale of the taste shocks:

$$\sigma = 1.66 \tag{B.25}$$

With the additional assumptions (B.21) and (B.22) and one of the normalizations, (B.23), (B.24) or (B.25), we are able to compute the utility gain k_i from removing relative wage misperceptions for each individual in our control group (see Subsection B.5.5 further

[42]From our reduced form equations, the identified parameter governing application behavior is $\beta_1 = \frac{1}{\sigma}\delta$. Scaling up σ and δ by the same amount thus leaves application behavior unchanged.

below for more details of this derivation).

B.5.4 Implementation and inference

To generate the counterfactual results presented in the main text, we compute the different quantities of interest described above for each individual in the control group, while replacing β_1 with its estimated value $\hat{\beta}_1$ from the baseline 2SLS specification (Table 5, Panel C, Column 1). We then compute averages or percentiles over the sample as relevant. To quantify the sampling uncertainty around these estimated counterfactuals, we produce standard errors via a bootstrap procedure that resamples individuals and then recomputes both the 2SLS estimate and the counterfactual quantities of interest within each bootstrap sample. For the welfare counterfactual, we produce three sets of results corresponding to the three normalizations we consider, (B.23), (B.24) or (B.25).

B.5.5 Derivation of welfare counterfactual

To compute our measure of the counterfactual welfare gain k_i, we first reorgainze the assumed Bellman equation for the value of continued search, (B.21), to arrive at:

$$U_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}}) = \frac{b_i}{\rho_i + \lambda_i} + \frac{\lambda_i}{\rho_i + \lambda_i} \Omega_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}})$$

Plugging this into the equation for the value of an application, (B.8), and rearranging yields:

$$\Omega_i(\boldsymbol{X}, \widetilde{\boldsymbol{X}}) = \frac{b_i}{\rho_i} + \frac{\rho_i + \lambda_i}{\rho_i} \sum_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{X}}) A_{i,j}(\boldsymbol{X}, \widetilde{\boldsymbol{X}})$$

With this, the equation defining k_i, (B.20), is equivalent to:

$$\sum_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}^0_i}) A_{i,j}(k_i \boldsymbol{W}_i, \widetilde{\boldsymbol{W}^0_i}) = \sum_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}^1_i}) A_{i,j}(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}^1_i})$$

Then using the fact that $A_{i,j}(k_i \boldsymbol{X}, \widetilde{\boldsymbol{X}}) = k_i^{\delta} A_{i,j}(\boldsymbol{X}, \widetilde{\boldsymbol{X}})$, we can solve for k_i here to obtain the formula:

$$k_i = \left(\frac{\sum\limits_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^1}) A_{i,j}(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^1})}{\sum\limits_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^0}) A_{i,j}(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^0})} \right)^{1/\delta}$$

Finally, it is convenient to multiply by ι_i^δ in the fraction:

$$k_i = \left(\frac{\sum\limits_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^1}) \iota_i^\delta A_{i,j}(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^1})}{\sum\limits_{j=1,2,3} \overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^0}) \iota_i^\delta A_{i,j}(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^0})} \right)^{1/\delta}$$

With our assumptions above, this formula can be implemented given our data and reduced from results. As shown in the preceding sections, given an estimate of β_1, the actual and counterfactual application probabilities, $\overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^0})$ and $\overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^1})$ can be computed under assumption (B.17). Under any of the normalizations, (B.23), (B.24) and (B.25), the values for δ and σ can also be inferred from the reduced form estimates because $\beta_1 = \frac{\delta}{\sigma}$. Finally, the parts of the formula reflecting the relative surplus value of applying to different jobs can be inferred from the counterfactual choice probabilities under assumptions (B.18) and (B.22) because the formulas for the choice probabilities, (B.6) and (B.7), imply:

$$\left(\overline{\Pi}^*_{i,j}(\widetilde{\boldsymbol{W}_i^1})\right)^\sigma = A_{i,j}(\widetilde{\boldsymbol{W}_i^1}, \widetilde{\boldsymbol{W}_i^1}) = \iota_i^\delta A_{i,j}(\boldsymbol{W}_i, \widetilde{\boldsymbol{W}_i^1})$$

C Additional details of the survey

This section presents the full survey instructions translated into English (Section C.1), includes screenshots of the interactive figures shown in the survey (Section C.2), and adds details of the data-driven procedure to create the education-specific job types not discussed in Section 2.3 (Section C.3). The original Danish survey instructions can be found at `https://tinyurl.com/zjw6dc83`.

C.1 Full survey questionnaire (in English)

A Thank you very much for participating in this survey about job search and career choices - we greatly appreciate it!

As a token of our gratitude for your time, we are giving away a total of **20 gift cards to GoGift** among those who complete the questionnaire. These gift cards can be used for many things and are very easy to use in both stores and online. Each gift card is worth **1,000 DKK**.

Half of the gift cards are distributed at random to respondents who complete the entire questionnaire (it is expected that this will take approximately **15 minutes**).

The rest of the gift cards are given to respondents who complete the questionnaire and do especially well in the exercises that make up part of the survey (e.g., by guessing the answer to some factual questions correctly). Keep an eye out for questions marked with an emoji. At these questions, you can win additional gift cards!

[*Emoji*]

If you win one (or more) gift cards, you will receive a direct notification in your e-Boks [e-Boks is a Danish state-official email] no later than September 31st, 2023. During the survey, you are able to disconnect and return to complete the survey later if you need to. Your answers are saved and when you click the link again, you will automatically return to where you left off.

B1 We are now going to ask you some questions about your education and current employment circumstances.

B2 Are you currently undertaking education (e.g., at a university or elsewhere)?

- Yes
- No

B3 [**if B2==Yes**] Which of the following educations are you currently undertaking?
[**if B2==No**] Which of the following educations correspond to the highest level of education that you have completed?

[3-layer dropdown with first layer corresponding to the following]

- Primary or lower secondary education
- Upper secondary education (High school)
- Vocational basic course
- Vocational education
- Labor market education
- Short-cycle higher education
- Medium-cycle higher education
- Bachelor's degree
- Long-cycle higher education (e.g., Master's degree)
- PhD programme and research education

[The education categories are based on DISCED, the Danish implementation of ISCED. The second layer corresponds to the broad field, and the third layer corresponds to the narrow field. See `https://www.dst.dk/da/Statistik/dokumentation/nomenklaturer/disced15-audd`]

B4 [**if B2==No**] When did you finish studying this education?
- MM/YYYY

[**If B2==Yes**] When do you expect to finish studying this education?
- MM/YYYY

[Create custom variables that take on values depending on responses to question B3: *custom_educ*: **Detailed education (all three layers)** *jobtype*1, *jobtype*2, *jobtype*3: **three specific types of jobs that are relevant to the respondent's education. The three job types are specific to the respondent's education and based on combinations of sector, industry, occupation and firm size. See Appendix Section C.3 for details on how the job types are generated.]**

C1 Which of the following statements best describes your current labor market situation position?

- **I have a job** and my contract is **not expiring in the immediate future.**
- **I have a job**, but my contract **expires soon** and I have **not** accepted another job offer.
- **I have a job**. My contract **expires soon**, but I have **accepted** another job offer.
- **I do not have a job**, but I have **accepted** a job offer which is starting soon.
- **I do not have a job** and I have not accepted a job offer.

C2 [**if C1==5**] Are you currently searching for a job?

[**if C1==1 or 2 or 3 or 4**] Are you currently searching for another job?

Note: Searching for a job means e.g., asking employers about potential job opportunities, gathering information about vacant positions, or sending out job applications.

- Yes
- No

C3 [**if C2==Yes & C1==4 or 5**] When did you start searching for a job?

[**If C2==Yes and C1==1 or 2 or 3**] When did you start searching for another job?

[**If C2==No and C1==1**] When did you start searching for your current job?

[**If C2==No and C1==2**] When do you plan on starting to search for another job?

[**If C2==No and C1==3 or 4**] When did you start searching for the job you have accepted?

[**If C2==No and C1==5**] When do you plan on starting to search for a job?

Note: Searching for a job means e.g., asking employers about potential job opportunities, gathering information about vacant positions, or sending out job applications.

- MM/YYYY

D1 [The Industri_example* and Funktion_example* variables are job type specific variables that give examples of the industry and occupations mentioned in the job types. These variables are empty if the job types does not include industry or occupation, or examples where deemed unnecessary.]

We would now like for you to answer some questions about three categories of jobs people with your educational background ("$custom_educ$") could have:

- **$jobtype_1$**

  [If Industri_example1a !="" and Industri_example1b =="" show:]

  An example of this industry could be:

  - *Industri_example1a*

  [Else if Industri_example1a != "" and Industri_example1b !="" show:]

  Examples of this industry could be:

  - *Industri_example1a* - *Industri_example1b*

  [Do the same for Funktion instead of Industri:]

  [If Funktion_example1a !="" and Funktion_example1b =="" show:]

  An example of this type of occupation could be:

  - *Funktion_example1a*

  [Else if Funktion_example1a != "" and Funktion_example1b !="" show:]

  Examples of this type of occupation could be:

  - *Funktion_example1a* - *Funktion_example1b*

  [If none of the conditions are met, don't show anything.]

- **$jobtype_2$**

[Same as for $jobtype\_1$ *just using* $Industri\_example2a - b$ *and* $Funktion\_example2a - b$ *instead of 1]*

- **$jobtype_3$**

  [Same as for $jobtype\_1$ *just using* $Industri\_example3a - b$ *and* $Funktion\_example3a - b$ *instead of 1]*

D2 Before you continue, use a moment to <u>read and think about the three categories of jobs listed above.</u>

How well do you understand what we mean by a **$Jobtype_1$**?

• Very poorly • Poorly • Neither well nor poorly • Well • Very well

How well do you understand what we mean by a **$Jobtype_2$**?

• Very poorly • Poorly • Neither well nor poorly • Well • Very well

How well do you understand what we mean by a **$Jobtype_3$**?

• Very poorly • Poorly • Neither well nor poorly • Well • Very well

D4 Do you know someone close to you who has had a job in this job category, such as a close friend or family member? If you are close to multiple people, think of the person who is closest to you.

$Jobtype_1$:

• Very close • Close • Not so close • I am not close to anybody who has had a job in this field.

$Jobtype_2$:

- Very close
- Close
- Not so close
- I am not close to anybody who has had a job in this field.

$Jobtype_3$:

- Very close
- Close
- Not so close
- I am not close to anybody who has had a job in this field.

E1 In the following, you are asked to guess the real answers to some **factual** questions about the labor market in recent years. You will be asked a few questions about things that persons with **the same educational background as you ("$custom_educ$")** have experienced in the three job categories above.

Take your time and try to guess the correct answer!

[*Emoji*]

In the questions marked with this emoji, the 10 participants with the most correct answers will receive **a gift card of 1,000 DKK** in their e-Boks.

E2 [**Salary expectations**]

Some jobs pay better than others. This question concerns monthly **gross salary** (i.e. **before tax** and **including** contribution to pension savings) of people in full-time employment. Full-time employment refers to contracts of employment of at least 37 hours of work a week.

Consider persons with the **same** educational background as you ([**"$custom_educ$"**]), who are newly graduated and completed their education in the 2010s whereafter they began working full-time.

What do you think were the **average monthly gross earnings during the first year of work for persons with jobs** in the following fields in the 2010s?

[*Emoji*]

As a reference, the average was approximately 31,000 DKK if **ALL** newly graduated workers of the period are considered (all positions and all educational backgrounds).
- Full-time salary of **$Jobtype_1$**: _______ DKK
- Full-time salary of **$Jobtype_2$**: _______ DKK
- Full-time salary of **$Jobtype_3$**: _______ DKK

If you are one of 3 respondents with answers closest to the actual numbers calculated from data from the Danish labor market of the years 2010 to 2018, you will receive a gift card of 1,000 DKK directly to your e-Boks. Remember that your answer should be in Danish "Kroner" (DKK) per month (before tax and including pension savings).

E3 [**Expectations of job application success**]

Some jobs are harder to get than others. This question concerns how often the application of unemployed people leads to an actual hiring in the firm.

Consider a person with the **same** educational background as you ([**"$custom_educ$"**]), who is newly graduated and completed their education in the 2010s whereafter they began receiving unemployment benefits and did not get a job for at least 6 weeks.

When the average person who is newly graduated sent out a job application for a position in the following fields, **what was the applicant's chance of employment at that firm**?

[*Emoji*]

As a reference, the average success rate was approximately 27 out of 1000 if considering **<u>ALL</u>** newly graduated (all positions and all educational backgrounds).

- Rate of success when applying for **$Jobtype_1$**: _____ out of 1000

- Rate of success when applying for **$Jobtype_2$**: _____ out of 1000

- Rate of success when applying for**$Jobtype_3$**: _____ out of 1000

If you are one of 3 respondents with answers closest to the actual success rate calculated from data on application gathered from applications logged in joblog.dk and administrative data on hirings in the years 2016 to 2018, you will receive a gift card of 1,000 DKK directly to your e-Boks. Remember that your answer should be given as a share of 1000. If you answer 10, this means that you think 10 out of 1000 applications lead to employment for the applicant. If you answer 45, this means that you think 45 out of 1000 applications lead to employment instead.

E4 [Salary after 5 years]

Some jobs offer better opportunities when it comes to wage growth than others. This question concerns the monthly **gross salary** (i.e. **before tax** and **including** pension savings) **5 years after having started a job.**

Consider again a person with the **same** educational background as you ([**"$custom_educ$"**]), who is newly graduated and completed their education in the 2010s.

[*Emoji*]

What do you think was the monthly full-time gross salary after 5 years, for persons who started working in the following fields?

As a reference, you are informed that the average total salary was approximately 41,000

DKK 5 years after employment if **ALL** newly graduated workers of the period are considered (all positions and all educational backgrounds).

- Full-time salary 5 years after first employment in **$Jobtype_1$**: ___________________DKK
- Full-time salary 5 years after first employment in **$Jobtype_2$**: ___________________DKK
- Full-time salary 5 years after first employment in **$Jobtype_3$**: ___________________DKK

If you are one of 4 respondents with answers closest to the actual numbers calculated from data for the Danish labor market of the years 2010 to 2014, you will receive a gift card of 1,000 DKK directly to your e-Boks. Remember that your answer should be in Danish "Kroner" (DKK) per month (before tax and including pension savings).

F1.A [**Treatment status T1, is determined in the following way.**
It is 0 if not eligible, or eligible and randomly chosen for control.
It is 1 if eligible and chosen for treatment.
Eligibility is determined on an education basis:
Eligible educations have both a data driven specific job type classification, and the number of observed education-to-job transitions with salary information in the least common job type was at least 30.]

[**wage treatment**] [**show if T1==1**]

Earlier, we asked you about the average monthly gross salary of full-time employees in different types of jobs.

We have researched the **actual** monthly gross salary of people with the same education as you ("**$custom_educ$**") who completed their education and worked full-time (at least 37 hours) in different job categories in the 2010s. The following graph depicts the **actual** **average monthly full-time gross salary in the first year of employment after graduating (the three bars) compared to your estimates (the three dots)**.

Figure: Actual data concerning salaries vs. your estimates (see Figure C1)

Note: When proceeding you are not able to return to this graph. You can continue from this page by clicking "next" after 15 seconds.

F1.B [**Reminder wages**] [**show if T1==0**]

Previously, we asked you about the **full-time salaries of newly graduated** in different types of jobs.

One of the questions was about the monthly gross salary of people with the same education as you ([**"$custom_educ$"**]) who completed their education and worked full-time (at least 37 hours) in different fields in the 2010's. The following graph depicts **your estimates** regarding the **average monthly full-time gross salary in the first year of employment after graduation.**

Figure: Your estimates regarding salaries (see Figure C2)

Note: You can continue from this page by clicking "next" after 15 seconds.

G1 We now ask you to **consider three distinct jobs** - one from each category listed below (the same that you were asked about earlier). The jobs you consider should be relevant to a person with the same educational background as you. The positions should also be **representative** of what you think the category of jobs in question would offer **you personally** in **every aspect** (pay, probability of receiving an offer if you apply, working conditions, etc.).

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

When you have thought about each of the three jobs and what they offer, we would like for you to answer some questions about the three positions. Some of the questions can be hard to answer, but we are interested in your best guess.

Click “next” to start answering the questions.

G2 Imagine that you are looking for a job and find postings corresponding to **the three positions we asked you to think about before** (i.e. three job postings in total). You only have the time to apply to exactly one of the positions. Alternatively, you do not apply to any of the three jobs.

It is hard to know exactly how you would act in this situation, but what do you think is the probability of each of the possibilities listed below?

Please enter your answers as percentages. An answer of 10 percent means that you would apply for the position in 10 out of 100 instances while 50 percent means that you would apply for the job about half of the time. Please make sure that your answers add up to 100 percent.

[**Here the respondents answer by moving indicators on three sliders from 1 to 97%, the default value is 25%. The sliders are set up, such that they always sum to 100. In practice this is done as follows: When respondents move one of the sliders, the other sliders adjust to make them sum to a 100. The adjustment is done with priority, so, if possible, only the 4th slider adjusts, then the 3rd, then 2nd then 1st.**]

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- Probability that I will apply for **Job A**: [**slider**]
- Probability that I will apply for **Job B**: [**slider**]
- Probability that I will apply for **Job C**: [**slider**]
- Probability that I will not apply to any of the jobs: [**slider**]

[**If respondents do not move the indicator, they cannot go to the next screen; error message "You have not moved any of the sliders. You need to move at least one, to show that you have answered the question."**]

G3 Imagine that you are offered all **the positions we asked you to think about before** (i.e. three job offers in total). You can accept exactly one offer. Alternatively, you do not accept any of the three job offers.

It is hard to know exactly how you would act in this situation, but what do you think is the probability of each of the possibilities listed below?

Please enter your answers as percentages. An answer of 10 percent means that you would accept a job offer in 10 out of 100 instances while 50 percent means that you would accept a job offer about half of the time. Please make sure that your answers add up to 100 percent.

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

[**Here the respondents answer by moving indicators on three sliders from 1 to 97%, the default value is 25%. The sliders are set up, such that they always sum to 100. In practice this is done as follows: When respondents move one of the sliders, the other sliders adjust to make them sum to a 100. The adjustment is done with priority, so, if possible, only the 4th slider adjusts, then the 3rd, then 2nd then 1st.**]

- Probability that I would accept **Job A**: [**slider**]
- Probability that I would accept **Job B**: [**slider**]
- Probability that I would accept **Job C**: [**slider**]
- Probability that I would not accept any of the three jobs offered: [**slider**]

[**If respondents do not move the indicator, they cannot go to the next screen;**

error message "You have not moved any of the sliders. You need to move at least one, to show that you have answered the question."]

G4 Consider again the three jobs from earlier. How many hours a week would you end up working in each of these positions?

*Note: We are interested in **actual working hours**. A job that is stated to be 37 hours could require more or less than 37 hours a week in practice. The position does not need to be stated as full time.*

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- Weekly hours in **Job A**: _______ hours
- Weekly hours in **Job B**: _______ hours
- Weekly hours in **Job C**: _______ hours

G5 Consider once again the three jobs from earlier. What would be your starting salary in each of these positions?

*Please enter the monthly gross salary **before tax** and **including pension savings**.*

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- Starting salary in **Job A**: _______ DKK
- Starting salary in **Job B**: _______ DKK
- Starting salary in **Job C**: _______ DKK

G6 Consider again the three jobs from earlier. If you applied for these, how likely is it that you would receive a job offer?

Please enter your answers as a share of 1000. If you answer 10, this means that you think you would receive an offer 10 out of 1000 times. If you answer 45, this means that you think you would receive an offer 45 out of 1000 times.

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- The chance that I would receive an offer if I applied for **Job A**: _______ out of 1000
- The chance that I would receive an offer if I applied for **Job B**: _______ out of 1000
- The chance that I would receive an offer if I applied for **Job C**: _______ out of 1000

G7 Consider again the three positions we told you to imagine earlier. **In 5 years**, how many hours a week would you spend working if you started working in each of these positions tomorrow?

*Note: We are interested in **actual working hours**. A job that is stated to be 37 hours could require more or less than 37 hours a week in practice. The position does not need to be stated as full time.*

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- Weekly hours of work in 5 years if I started working in **Job A** tomorrow: _______ hours
- Weekly hours of work in 5 years if I started working in **Job B** tomorrow: _______ hours
- Weekly hours of work in 5 years if I started working in **Job C** tomorrow: _______ hours

G8 Consider again the three positions from earlier. What would your monthly gross salary be **in 5 years** if you started in each of these positions tomorrow?

*Please enter the monthly gross salary **before tax** and **including pension savings** after 5 years.*

Job A:	Job B:	Job C:
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- Monthly gross salary in 5 years if I started working in **Job A** tomorrow: _______ DKK

- Monthly gross salary in 5 years if I started working in **Job B** tomorrow: _______ DKK

- Monthly gross salary in 5 years if I started working in **Job C** tomorrow: _______ DKK

G9 Consider again the three positions from earlier. Imagine the experiences in the workplace that these three jobs would entail. How likely is it that you would get along really well with your coworkers in these positions?

Job A:	Job B:	Job C:
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- The probability that I would get along really well with my colleagues at **Job A**:

- **Very unlikely**
- **Unlikely**
- **A little unlikely**
- **A little likely**
- **Likely**
- **Very likely**

- The probability that I would get along really well with my colleagues at **Job B**:
[**Same dropdown**]

- The probability that I would get along really well with my colleagues at **Job C**:
[**Same dropdown**]

G10 Consider again the three jobs from earlier. Think about how challenging you would find the work in these positions. How likely is it that you would perform excellently in each of the three positions?

Job A:	Job B:	Job C:
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

- Probability that I would perform really well at **Job A**:

- **Very unlikely**
- **Unlikely**
- **A little unlikely**
- **A little likely**

- **Likely**
- **Very likely**

- Probability that I would perform really well at **Job B**:

[**Same dropdown**]

- Probability that I would perform really well at **Job C**:

[**Same dropdown**]

G11 Consider again the hypothetical situation where you receive **job offers regarding the three positions** from earlier questions. Earlier, you gave the following probabilities when answering how likely you were to accept exactly one of the offers.

- Probability that I would accept **Job A**:

[**Previous answer from G3 is inserted here**]

- Probability that I would accept **Job B**:

[**Previous answer from G3 is inserted here**]

- Probability that I would accept **Job C**:

[**Previous answer from G3 is inserted here**]

- Probability that I would not accept any of the three jobs offered:

[**Previous answer from G3 is inserted here**]

Job A:	**Job B:**	**Job C:**
$Jobtype_1$	$Jobtype_2$	$Jobtype_3$

Imagine that you again receive the same job offers, but that **the offers are atypical**, such that the starting salary is [**median answer from G5, rounded to 1000s, is inserted**] in all three of the jobs. Besides the starting salary, the jobs are exactly as before (including the weekly hours, relative wage growth and so on). How would this change your previous probabilities stated above?

Please enter your answers as percentages. An answer of 10 percent means that you would accept a job offer in 10 out of 100 instances while 50 percent means that you would accept a job offer about half of the time. Please make sure that your answers add up to 100 percent.

[**Here the respondents answer by moving indicators on three sliders from 1 to 97%, the default value is 25%. The sliders are set up, such that they always sum to 100. In practice this is done as follows: When respondents move one of the sliders, the other sliders adjust to make the sum correspond to 100. The adjustment is done with priority, so, if possible, only the 4th slider adjusts, then the 3rd, then 2nd then 1st.**]

- Probability that I would accept **Job A** when the starting salary of all three of the positions is [**median answer from G5, rounded to 1000s, is inserted**]:
[**slider**]

- Probability that I would accept **Job B** when the starting salary of all three of the positions is [**median answer from G5, rounded to 1000s, is inserted**]:
[**slider**]

- Probability that I would accept **Job C** when the starting salary of all three of the positions is [**median answer from G5, rounded to 1000s, is inserted**]:
[**slider**]

- The probability that I would turn down all three job offers is:
[**slider**]

[**If respondents do not move the indicator, they cannot go to the next screen; error message "You have not moved any of the sliders. You need to move at least one, to show that you have answered the question."**]

G12 [$Jobtype_example_chosen$ is an example of one of the three education-specific job types. Only show this question when the job types examples are available.]

Throughout this survey we have asked you about 3 different types of jobs. To better understand how you understand these, we pose the following question:

In which job category would you place the more concrete example:

$Jobtype_example_chosen$

- It is an example of **$Jobtype_1$**.
- It is an example of **$Jobtype_2$**.
- It is an example of **$jobtype_3$**.
- It does not fit into any of the three types of jobs.

H1 Finally, we would like for you to answer some questions about yourself.

How willing are you, in general to take a risk?

Please answer on a scale from 0 to 10[43] where 0 means that you are “completely unwilling to run any risks” and 10 means that you are “very willing to run a risk”. You can also use any integer between 0 and 10 to indicate where you are situated on the scale.

[The respondents can choose between the integers 1-10]

Think of all the job offers you have received in the last 2 years.

How many job offers have you received in total?

[43]Note: In the survey there was a mistake, and the text stated 0 to 10, but is was only possible to choose between 1 to 10

___

How many of these did you reject?

___

How many of these did you reject because you received an offer for another job at the same time that you liked better?

___

How many job offers did you reject although you had no other jobs available?

___

How many children would you like to have throughout your entire life? (your best estimate is fine)

______ children

What do you think is the probability that you will have a child within 4 years?

______ percent (%)

I1 [wage treatment] [T1==1]

On a previous page of this survey, we showed you some information about the **actual** monthly gross salary of people with your educational background ([**"$custom_educ$"**]) who graduated and worked full-time (at least 37 hours) in different job categories between 2010 and 2018.

The same graph that you saw earlier is depicted below. It shows the **actual average monthly full-time gross salary in the first year of employment after graduating (the three columns) compared to your estimates (the three dots)**.

Figure: Actual data concerning salaries vs. your estimates (see Figure C1)

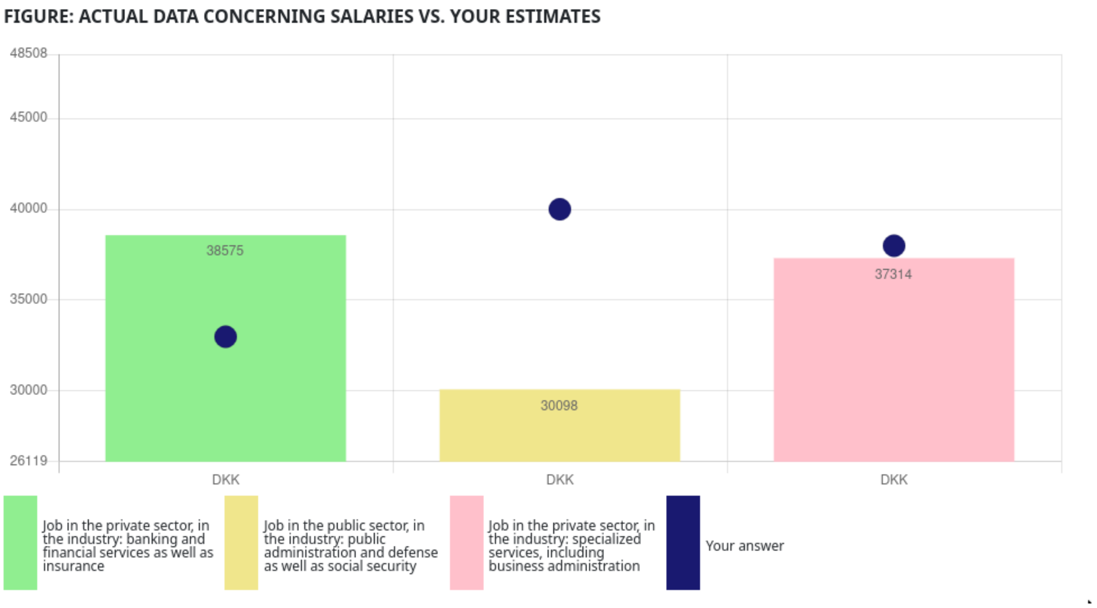
[**if T1==0 just go to next question**]

I2 If you have any thoughts or input, you would like to share with us about job search or the survey you have just completed, please write them down below:
[**Text entry box**]

Thank you for taking part in this survey!

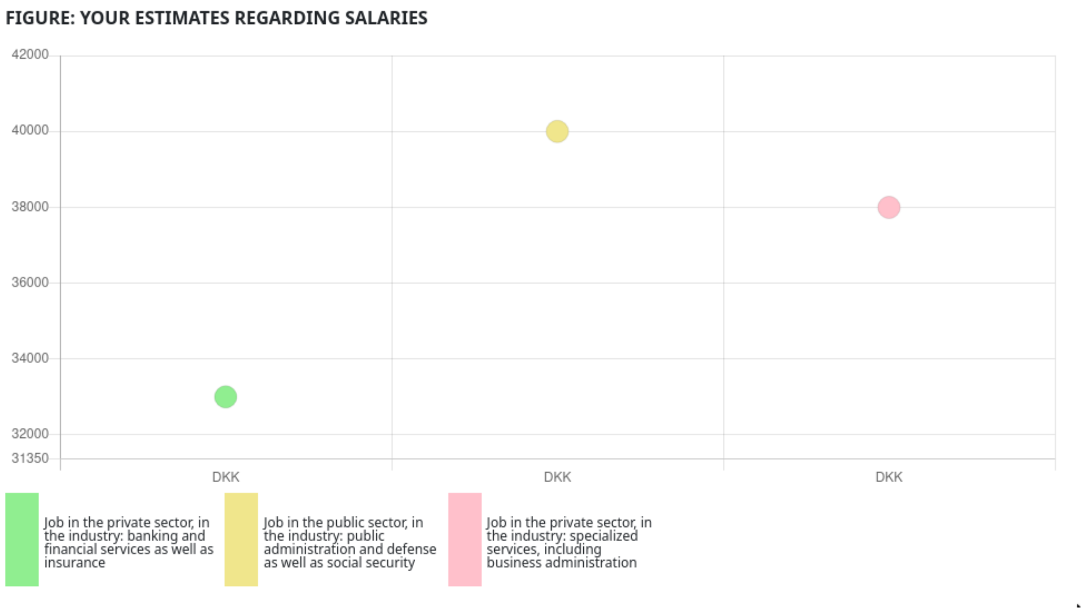
C.2 Survey screenshots and examples

Figure C1: Example of treatment slide shown to the treatment group in the survey



Note: This figure shows an example of the interactive graph (translated into English) shown to the treatment group (translated into English), reminding them of their stated answers about the average wage in each job, and informing them of the actual averages in the register data.

Figure C2: Example of treatment slide shown to the control group in the survey



Note: This figure shows an example of the interactive graph (translated into English) shown to the control group, reminding them of their stated answers about the average wage in each job.

C.3 Details on creation of job types and information treatment

C.3.1 Manual changes made to job types

After creating the job categories using the approach explained in section 2.3 all job types was read through manually to ensure consistently and interpretability. As a start we used the default labels for industry and occupation, given by Denmark Statistics. We changed 18 occupation labels and 19 industry labels. Some were smaller word changes for clarity, others where typesetting changes, and a few were a full write up of categories which default labels had been a residual.[44]

After creating our categorizations, we found a error in a particular occupation code in 2010. For Philosophy PhD (DISCED: 802520) and Information science Bachelor (DISCED: 602535) graduates, this occupation code was part of one of the chosen job types. For Information science bachelor we could just choose a different categorization since more were available[45], but for Philosophy PhD we had to rerun the procedure using only 2011-2018 cohort transitions to create an alternate categorization.

Table C1: Examples of job types asked about in the survey for different educations

Degree	**Job types**	**Categories**
Ms. Economics DISCED: 703950	- Private sector, in the industry of banking, financial, and insurance activities - Public sector, in the industry of public administration, defense, and social security - Private sector, in the industry of professional, scientific and technical Activities	- Public/private - Industry, 1-digit
Ms. Physics DISCED: 703530	- Public sector, teaching and research at a university - Public sector, teaching at the level of high school - Private sector workplace with than 50 workers, doing development and analysis of software and applications.	- Public/private/size - Occupation, 3-digit

Notes: Examples of the job types, that were showed to the respondents, for two different educations. They are translated from Danish.

C.3.2 Creation of job type examples for explanation and validation question

For some job types it was possible to create examples, by using subcategories of occupation or industry with a more granular level, 4-digit for NACE and 6-digit for DISCO. The 2

[44]For example ISCO/DISCO code 2149 is named "Other engineering work (except Electrical engineering)", referring to the other categories under the 21 code. But as our respondents would not be able to see all other categories, we created a new label, explaining what would be in the residual.

[45]So the second best out-of-sample mean squared wage predictor.

most common subcategories in the education are chosen as possible examples for industry and occupation separately. The examples were checked manually to make sure they made sense and were not the same as the upper category.[46] When available they were shown in question D1 of the survey.

At the end of the survey we had a validation question that also tested attention. We created examples of the job types by combining the subcategories of industry and occupation and asking the respondent to place a random one of them, in one of the 3 job types. The validation examples are created analogously to the examples and with a large overlap and the same manual label changes. The main difference is that when applicable, the 2 most common *combinations* of the subcategories of industry and occupation are chosen as possible examples.

[46]This involved changing 41 labels to make them interpretable and distinct from the upper category, removing 10 labels that were not distinguishable from the upper category, and adding 11 examples where none was usable in the existing labels, but examples where easy to think of.

D Post survey data handling

D.1 Procedure to correct reporting mistakes

To asses how simple reporting errors may be contributing to our measured misperceptions, Table 3 shows how measured misperceptions change if we apply a simple procedure to correct obvious reporting errors in the raw survey data.

The specific procedure we consider works as follows. All reported monthly wages outside 7,000–120,000 DKK were flagged as being implausible. For these reports, we examined the number of digits in the raw responses. If all flagged values for the same respondent had the same number of digits, we interpret it as a systematic reporting issue. In that case, the value was rescaled according to the most plausible interpretation:

- six-digit values were treated as annual earnings and divided by 12
- two-digit values were treated as monthly earnings reported in thousands and multiplied by 1,000
- three-digit values were interpreted as hourly wages and multiplied by 160.33 (approximate full-time hours per month)

If instead the number of digits varied across flagged responses, we interpreted this as an idiosyncratic reporting error and corrected each report individually as follows:

- six-digit values were assumed to contain an extra digit and divided by 10
- four-digit values were assumed to be missing a final digit and multiplied by 10
- two-digit values were again treated as thousands and multiplied by 1,000.

Finally after applying the corrections above, values within the acceptable range of 7,000–120,000 DKK were retained, while all others were coded as missing.

D.2 Lost observations due to send-out error in the second wave

Due to an error regarding the send-out of the survey in the second wave, the same link(s) were sent to multiple people, making their answers invalid and making it impossible for us to tie them to the register data. The mistake was found and the correct links were sent out within 4 hours. We are not entirely certain if any of the links sent out in the first instance were correct and unique, so we have been forced to remove all answers registered in the 4 hour period, which was 151 distinct links, answered by multiple persons.