Wage Misperceptions and Young Workers' Sorting in the Labor Market

Asker Nygaard Christensen Nikolaj Harmon Sonja Settele Daphné Skandalis

January 31, 2025

Abstract

Given the prevalence of information frictions in the labor market, young workers may misperceive their options and sort into the wrong career. To test this, we conduct a survey among young unemployed individuals and soon-to-graduate students, where each respondent receives a tailored survey version focusing on three job types that are frequent transitions for their degrees and hence represent relevant job options to them. The survey includes a randomized information treatment about the average past entry-level wage in each job type and elicits young workers' expectations about average wage and non-wage amenities in each job type as well as their planned job applications. We document that young workers largely misperceive the wage differential between job types. We show that receiving wage information makes them update their expectations about the wage they could receive in each job type, but not about non-wage amenities (such as working hours, collegiality, or job difficulty). The wage information also affects young workers' planned search behavior: the planned applications to each job type increase by 3.1-3.3% for each 1% increase in wage expectations. Our results suggest that young workers' misperception about the wage differential between their job options may adversely affect their sorting in the labor market.

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

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, 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, and seminar participants at ELSE, the University of Cornell for helpful comments and suggestions. **Funding**: Financial support through the Rockwool Foundation is gratefully acknowledged. The activities of CEBI are funded by the Danish National Research Foundation (Grant DNRF134). Settele: Funded by 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

There is growing evidence that workers are not correctly informed about the wages they can get (e.g Jäger et al. (2024)). This is likely even more true for young workers, who have little experience in the labor market. This may lead them to seek out jobs and eventually sort into jobs that offer wages below other relevant job options. Since early jobs may have a lasting effect on workers' career, this may lead to persistent earnings losses.[1] However, such misallocation is very difficult to detect in practice. Indeed, jobs differ along many dimensions besides wages. Some workers may start their career in job types that offer relatively low wages, not because they are unaware of wage differentials but because they appreciate other amenities at these jobs. Moreover, getting a job does not only depend on workers' decisions but also on employers' decisions and labor market frictions. Some workers may start their career in job types that offer relatively low wages, simply because they did not receive offers in higher paying jobs.

Do young workers hold incorrect beliefs about their wage options, and do these beliefs affect their search behavior and their sorting into careers? To answer this question, we have conducted a survey among recent graduates about their job options in Denmark. Due to the structure of the Danish unemployment insurance system and the detailed data available, we are able to target a sample of young persons, who have recently signed up for unemployment insurance, and are mostly recent graduates. New graduates are eligible for unemployment insurance in Denmark even though they have received no prior labor earnings, which means that financial constrains should have a minimal role in their job sorting decision in this context. Our aim was to send the survey during the period when many recent graduates are searching for a job, to focus on information problems which might directly affect the search and sorting process. For that reason, our survey was rolled out in the summer of 2023, i.e. around university graduation dates. We invited all expected university graduates at the largest university in the country, the University of Copenhagen, as well as all individuals below age 40 who had just registered as unemployed.

Instead of focusing on a specific degree, we aimed at covering the largest possible set of degrees. In administrative data, we identified for each type of degree, three types of jobs that can be described using intuitive categories (firm size, public vs. private sector, industry and occupation) and covered a large share of entry-level jobs of graduates in previous cohorts. We excluded some degrees in this process as they had too few students, or students who selected into too similar or too diverse jobs after graduation—such that that we could not classify students' job options into three distinct job types. Each respondent received a tailored version of the survey, where she was asked to consider three job types that are

[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), andWachter (2020)). Arellano-Bover (2024) documents long-term positive effects from getting a first job at a large firm as opposed to a small firm.

relevant to her specific university major. Our final sample consists of 1,941 respondents.

The first part of the survey elicits beliefs about the wages offered in the three relevant job types. Specifically, we ask respondents about the average wage received by previous graduates from their university major in each job type. Eliciting beliefs about population averages rather than about potential personal wages has two advantages. First, it allows us to abstract from the individuals' perceived personal characteristics (e.g., productivity or perceived fit for each job type) and to obtain inter-personally comparable beliefs. Second, it allows us to compare elicited beliefs to ground truth measures from administrative registers, and determine the extent of workers' misperceptions. Since we are interested in workers' misperceptions to the extent that they might affect their sorting in the labor market, we focus on misperceptions about *wage differentials between job types*.

In order to isolate the causal effect of job type-specific wage expectations on individuals' job search behavior, we provide a random half of our respondents with factual information about entry-level wages in the three job types that are relevant to them. After the information treatment, we elicit the respondent's personal expected starting wage if she were to start a job in each of the tree job types, as well as the likelihood of applying for a job in each of the tree job types. In order to study spillover effects of wage expectations on perceived non-wage amenities, we also ask respondents about a wider range of job amenities in different job types, such as working hours, collegiality or job difficulty. Additionally, we measure their perceived difficulty to get a job within each type by eliciting the perceived probability to receive a job offer conditional on applying.

We first document that individuals have large misperceptions about the wage differential between the job types that are relevant to them. On average, when comparing two job types, 37 percent of respondents think the factually lower-paying job pays a higher wage, 10 percent report the same wage estimate, and only slightly more than 50 percent get the wage ranking right. In general, respondents under-estimate the dispersion in wages across job types, i.e., the median perceived wage differential across job types corresponds to 25 percent of the true wage differential, despite the fact that perceived wage levels generally exceed actual wage levels. In line with a role for rational inattention, we find that perceived wage differentials are closer to the truth whenever the true wage differential is larger and incorrect perceptions are potentially more costly.

We estimate the causal effect of our information treatment on respondents' beliefs about the amenities typically offered in their job options. We show that the information we provided about population wages lead individuals to update their expectations about their own potential wage at each job type in the expected direction: after receiving the information, individuals who initially under-estimated (respectively over-estimated) the population wage increased (respectively decreased) their wage expectations. This confirms that information about prior cohorts is considered relevant to form expectations about personal wages. In terms of magnitude, our estimates suggest that people increase (decrease) their personal wage expectations in each job type by 0.31-0.38% for each 1% they under(over)-estimated the average wage in the job type before receiving the information treatment. In order to understand how wage information may affect behavior, it is important to consider how it may also affect other beliefs beyond wages. For instance, when people learn about higher wages in a given job type, they may infer that this job type would provide them a less pleasant work environment or be harder to get than they initially thought. We hence also estimate the effect of the information treatment on job seekers' expectations about various non-wage amenities and about the probability to get the job when applying. We find that in fact, the wage information does not affect any belief beyond wages. Interestingly, this suggests that workers do not implicitly believe in compensating wage differentials, since they don't think that higher wage must come at the cost of lower non-wage amenities.

We then estimate the effect of our information treatment on planned job application behavior. We show that the information we provided has large effects on planned applications in the expected direction: after receiving the information, individuals who initially under-estimated (respectively over-estimated) the wage in a job type increased (respectively decreased) their propensity to apply to that job type. In terms of magnitude, our estimates suggest that people increase their planned applications to each job type by 1.45-1.57% for each 1% they under(over)-estimated the average wage in the job type before receiving the information treatment. This confirms that information about wages of prior cohorts in different job types is considered relevant to form expectations about personal wages. Despite the misperceptions we have documented, this information is hence valuable to workers. It suggests that young workers would direct their applications to different job types and likely sort into different careers, if better informed.

Moreover, our randomized information treatment allows us to learn about the importance of wage expectations as a driver of job application decisions. Intuitively, since it generates exogenous variation in wage expectations, it provides an ideal instrumental variable to identify the causal effect of wage expectations on job applications. To guide this analysis, we base our specification on a discrete choice model of job applications. Using the information treatment as an instrument, we find that the planned applications to each job type increase by 3.1-3.3% for each 1% increase in their personal wage expectations. In other words, our results suggest an elasticity of labor supply with respect to the job type-specific wage expectation of 3.1-3.3.

Our results contribute to several strands of the literature. First, we contribute to the literature showing that workers have misperceptions about the labor market that may lead them to behave suboptimally. A strand of articles have highlighted misperceptions that lead workers to stay unemployed excessively long (Krueger and Mueller (2016), Spinnewijn (2015), Mueller et al. (2021)). Relatedly, several studies have shown that receiving search advice can shorten unemployment durations (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)).

Beyond excessive unemployment durations, misperceptions may generate other costs for workers and society that have received much less attention. Our paper highlights that misperceptions can cause young workers to start their career in the wrong job type. We hence focus on misperceptions about *wage differentials between job types* rather than the overall level of wage. Our paper is closely related to Jäger et al. (2024) who show that employed workers misperceive wages in their outside options, which leads them to stay too long at low-wage firms and negotiate too little. We confirm that such misperception about wages at relevant job options exist even among workers who are not yet employed and actively searching for jobs and show their dramatic consequences for early-career sorting decisions.

Second, we provide new estimates for the job type-specific wage elasticity of labor supply which help reconcile mixed evidence in the literature. Most estimates in prior literature are surprisingly small, i.e. between 0.1 and 1.9 (Staiger et al. (2010), Falch (2010),Marinescu and Wolthoff (2020),Belot et al. (2022b) ,Mueller et al. (2024)[2]). One exception is Bassier et al. (2023) which find estimates around 3 to 5. Manning (2011) highlights that such low firm-specific elasticities would imply an enormous amount of monopsony power and suggests various reasons why some of these estimates might be biased downward: the labor supply response is measured in the short-run though it might take longer to materialize, workers might expect the increase in wage to be temporary, firms might react to a rise in mandated wages by decreasing recruitment activities. Belot et al. (2022b) analyze job applications responses to posted wages in vacancies and find elasticities of 0.7-0.9. They argue that these elasticities could be dampened by applicants' belief that high-wage vacancies are more competitive. Our setting neutralizes many of these factors that have been hypothesized to dampen labor supply reactions in prior literature and we find higher elasticities of 3.1-3.3: we do not have to wait for hiring to materialize, our setting exploits variation in wage expectations while firms' behavior are held constant, we show that applicants in our setting do not interpret information about high wages as signals for high competition. This suggests that these factors may indeed explain the smaller elasticities found in other settings.

Third, our paper is closely connected to the literature studying the determinants of career choices (e.g., Wiswall and Zafar (2015)). The literature has mostly focused on major choices. We extend this framework by focusing on the choice of the type of the first job once education is completed. This is a very important decision as well, as earnings still vary a lot among individuals who complete the same education and early jobs have persistent consequences on the rest of the career (e.g, von Wachter and Bender (2006), Arellano-Bover (2024)). However, studying sorting decisions at the start of the career poses unique challenges. In particular, the possible options that each individual faces when entering the

[2]Mueller et al. (2024) find vacancy duration elasticities with respect to the wage of -0.07 to -0.021. The probability of filling a vacancy, θ, is the inverse of its expected duration, $d = \frac{1}{\theta}$. So in a world of constant elasticities, the elasticity of d with respect to the wage is equal to minus 1 times the elasticity of θ with respect to the wage, $\frac{\partial \ln d}{\partial w_{i,j}} = -\frac{\partial \ln \theta}{\partial w_{i,j}}$. This is e.g. explained in Bassier et al. (2023)

labor market are not pre-defined (as opposed to the set of possible majors one can choose): we hence constructed job types using past labor market transitions.

2 Setting, survey design and data

2.1 Empirical setting

Ex ante, we expect wage misperceptions to be most relevant and most costly for young workers and labor market entrants who are actively weighing different job options. Given their limited labor market experience, such workers are more likely to be poorly informed about wages. Given the persistence of early career job choices, potential distortions in search behavior may also be particularly costly for this group.

Accordingly, our analysis focuses on younger workers in Denmark who register as newly unemployed job seekers, and/or who are about to graduate from an educational degree and enter the labor market. After registering as unemployed, such workers will be assigned to receive an appropriate form of public benefits. A majority of individuals qualify to receive the relatively generous Danish Unemployment Insurance scheme. During the time period we study, newly eligible unemployment insurance (UI) recipients are eligible to receive 23.449 DKK for up to 3 months, and 19,728 DKK per month for up to 21 months thereafter[3]. The key UI eligibility criteria is membership in one of the Danish UI funds. Importantly, membership and eligibility is open to individuals enrolled in an education at no cost during their studies, so searching for job while on UI is a common way for new graduates to start their labor market career.[4] Individuals registering as unemployed who do not qualify for UI, or have exhausted their UI period, will typically receive the less generous subsistence benefit called 'Kontanthjælp'[5]. The available unemployment insurance means that financial constrains should have a minimal role in their job sorting decision.

Denmark offers an ideal setting as online vacancies typically do not include posted wages, making wage information less salient to workers deciding where to send their applications.

[3]Graduates under the age of 30, are eligible upon graduation at a rate of 14,106 DKK, which is reduced to 9.700 DKK after 3 months of unemployment, and they are covered for up to 1 year. For more information see: `https://www.retsinformation.dk/eli/retsinfo/2023/9076`

[4]A detailed study by the Ministry of Economic Affairs and the Interior found that around half of all new graduates receive unemployment benefits within 6 months of graduation, (Økonomi og indenrigsministeriet (2018))

[5]The exact amount varies depending on personal circumstances, but at the time of our study, the base amount is 7,699 DKK per month when younger than 30, and 11,944 DKK per month after reaching the age of 30. See `https://www.retsinformation.dk/eli/retsinfo/2022/10391`

2.2 Survey details

The key data source for our analysis comes from a survey we conducted in the summer of 2023[6]. Sampling into the survey was done in two distinct ways. First, in collaboration with the Danish Agency for Labor Market and Recruitment (STAR), we sampled individuals below age 40 who registered as newly unemployed in the summer of 2023.[7] We targeted this period specifically to maximize the number of new graduates sampled as most educational programs finish during the summer months. Since regular UI entry happens year round however, the sample also covers many younger workers who have had at least some prior experience in the labor market. We refer to this as the STAR sample.

Second, in collaboration with the University of Copenhagen (UCPH), we sampled all UCPH students who were close to finishing their masters degree during the spring semester of 2023.[8] This allows us to supplement the young unemployed sample with individuals who are in the process of graduating and looking for jobs but who do not necessarily enter unemployment, for example because they successfully land a job before graduation. We refer to this as the UCPH sample.

To conduct the survey, all sampled individuals were contacted via the official governmental email, Eboks. Individuals in the STAR sample were contacted at most 3 weeks after they registered as unemployed.[9] Individuals in the UCPH student sample were contacted on June 21st, just before the typical masters graduation date. This timing of the survey aimed to primarily reach actively job searching individuals who had not yet started a new job. Indeed, 68 percent of the final sample report actively searching for a job at the time of answering the survey and only 20 percent report currently being employed.

A total of 43,622 individuals were contacted as part of the STAR sample and an additional 2,609 individuals were contacted as part of the UCPH sample. Around 15% of the contacted population responded to the survey and 9% fully completed it. This is comparable to the typical response rates for such surveys in the Danish context, although on the lower end, likely reflecting a general lower participation rate among our population of young, mainly unemployed individuals. More information on the population and the difference between the waves is provided in the Appendix, Table A1.

For the purpose of understanding wage misperceptions and their potential distortions, the survey has two key aims: First, the survey aims to provide objectives measures of whether and how much job seekers misperceive the difference in wages typically offered in

[6] The survey experiment was pre-registered in the AEA RCT Registry under AEARCTR-0011592

[7] We sampled all individuals who signed up in 3 waves. The first two waves covered individuals who signed up in May 29th-July 10th. The third wave covered individuals who signed up in July 15th-August 15th.

[8] Formally, the UCPH sample includes all registered UCPH master's students who were less than 40 ECTS from finishing their studies in March 2023.

[9] STAR allowed us to sample people in 3 waves over the summer of 2023, at each time sampling all persons who had signed up for unemployment benefits in the previous 3 weeks.

different jobs. Second, the survey aims to measure the extent to which such misperceptions distort job seekers search behavior. The next sections describe how the survey design aims to achieve this goal.

2.2.1 Defining job types

Our survey deals with misperceptions about the wages offered in different types jobs and related job search decisions. Naturally, this requires us to operationalize the concept of job *types*. In doing so we face three key constraints: First, the number of job options we ask workers about must be parsimonious enough to fit in a survey. Second, for results to be meaningful in practice, each worker needs to be asked about a a set of jobs that reasonably matches their actual set of considered jobs; while it is very likely that workers are poorly informed about wages in occupations that they could never work in, such misperceptions will not matter for individual decisions and outcomes. Third, since a key aim of the survey is to provide objectives measures of misperceptions, we must define our job types in a way that allows us to reliably construct ground truth measures of typical wages based on available data. To achieve this, the survey was designed to ask each worker about three individualized job types defined in terms of standard administrative definitions of industry, occupation, sector and/or firm size.

To determine which jobs are relevant we rely on the educational background of the respondent. Most educational degrees in Denmark are specialized, so education strongly shapes possible jobs in our sample of graduates and early career workers. To group individual jobs into a meaningful set of *types* for each educational background, we rely on a data-driven approach leveraging administrative data on job transitions of past graduates within each education background from Statistics Denmark. We use the BFL register for information on monthly labor market outcomes at the person level. And the UDDA register for information on educations at the person-level. We look at educations at the 6-digit DISCED level of level and field.[10]

For each education, we attempt to create a set of three specific, relevant job types, that are identifiable in the administrative data and cover a large share of the job types that graduates go into. The job types are created in a data-driven way, using data on the first job transition of Danish graduates in 2010-2018. They can be based on any combination of sector and firm size (jobs in public firms, private firms with 50 or fewer employees, or private firms with more than 50 employees), occupation (using either 1, 2, or 3 digits of the International Standard Classification of Occupations (ISCO-08) codes[11]), and industry

[10] This is the danish version of the International Standard Classification of Education (ISCED).

[11] For more information see: `https://ilostat.ilo.org/resources/concepts-and-definitions/classification-occupation/`

(using either 1. letter or 2 digits of the NACE classification[12]). For each education, we construct job types of all possible combinations of these categories, i.e. we combine all the different levels of granularity across each topic, but not within.[13] From all these possible combinations, we then select those that satisfy the following criteria: (1) The 3 most common job types cover more than 40% of transitions. (2) The 3rd most common job type covers more than 5% of transitions.

For some educations, no combination satisfies these criteria, these are given three exhaustive, but non-specific job types. For the rest, if there is only one combination, the three most common transitions in this combination become the job types for the education. If there are multiple combinations we choose the combination that is the best at predicting logged wages.[14]

The resulting job types were hand-checked for clarity. Some wording was modified, and minor adjustments were made to the actual job types.

For some educations it was not possible to create a specific job type classification,[15] or it was possible but the number of observed transitions with wage information in the least common job type was below 30, meaning the average wage would be a noisy, less informative measure. Persons from these educations were classified as not eligible for treatment with the information experiment but were still surveyed as they were still able to answer relevant questions.

At end of the process above, for each possible educational background of our respondents, we have a set of three main relevant *job types*, each of which are defined through some combination of industry, occupation, sector and/or firm size. Throughout the survey, respondents then receive questions specifically about the three job types corresponding to their education. Table A6 shows an example of the three job types for respondents with a masters in Economics and a masters in Physics.

A crucial question is whether the process above succeeded in creating job types that are also well understood by respondents in the survey. To validate this, after describing the job types to the respondents, we asked the respondents how well they felt they understood the job types. Figure 1 shows the answers to this question, showing that the respondents mostly stated that their understanding was good or very good. The jobs are sorted by the number of transitions for each education, meaning that the respondents tended to understand the

[12]For more information see: `https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF`

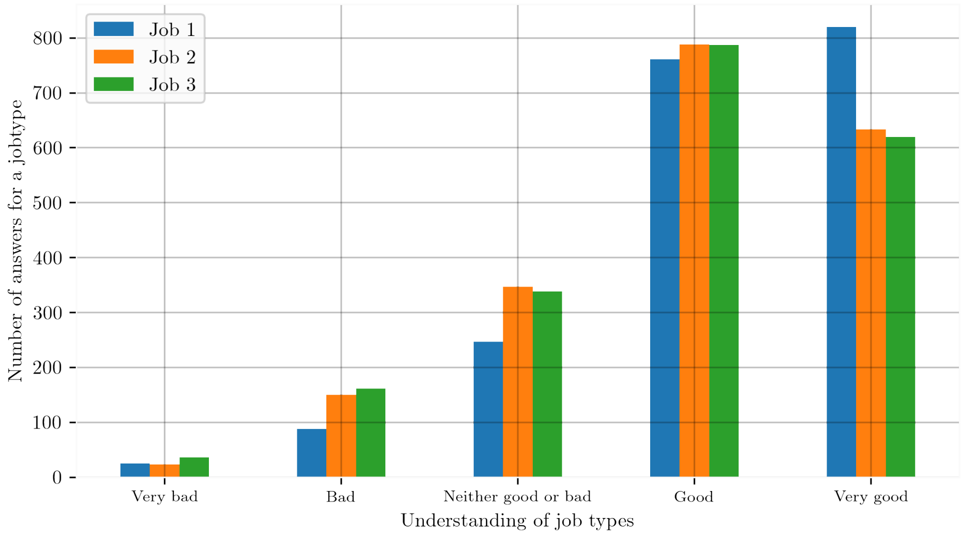
[13]So 1-digit ISCO-08 is not combined with 2-digit ISCO-08, but they are both combined with all levels of NACE and sector and firm size, along with combinations with only one of those, or not combined at all

[14]For each combination, we do 5-fold cross-validation and calculate the mean squared error of a regression of log wages on dummies from the job types resulting from the combination, and choose the combination with the lowest error to create the job types for the education.

[15]For these educations we created three exhaustive, but non-specific job types: jobs in public firms, private firms with 50 or fewer employees, or private firms with more than 50 employees.

most common job of their education slightly better.

Figure 1: '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?'. The jobs are ordered by how common they were to transition to in the past, such that job 1 is the most common job type to go into for persons with the respondents' education. This figure shows the part of the sample that was eligible for treatment.

Job type examples For some job types, it was possible to create examples, by creating job type examples with a subcategory of occupation or industry with a more granular level, 4-digit for NACE and 6-digit for ISCO-08. The 2 most common subcategories in the education are chosen as possible examples. The examples were checked manually to make sure they made sense and were not the same as the upper category. They were then given to the survey respondents when presenting the job types.

At the end of the survey, we had a validation question that tested attention and understanding of respondents. 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 three job types. The validation examples are created analogously to the examples and with a large overlap. The main difference is that, when applicable, the two most common combinations of the subcategories of industry and occupation are chosen as possible validation questions. While the examples are created at for each category. So the most common subcategories of occupation, and the most common subcategories of industry, separately.

77% of the respondents who were given a validation question, answered it correctly, further suggesting that the job types were understandable. A further breakdown of the answers across educations is shown in the Appendix, Table A4.

2.2.2 Measuring misperceptions about typical wages

The first part of the survey is aimed at measuring wage misperceptions. For the three job types relevant to their education, $j \in \{1, 2, 3\}$, respondent i is asked what they thought the past average earnings were for people in the first year after graduating with the same education as them.[16] To minimize idiosyncratic errors from people misunderstanding units etc., respondents were informed of the overall average wage paid to new hires over the period (irrespective of educational background and job type) in conjunction with the question. We refer to respondent's answer to this question as their perceived average wage, $\widetilde{W}^a_{i,j}$. Respondents are incentivized to exceed effort and answer correctly through monetary incentives.[17] Importantly, using administrative data, we can in fact compute the ground truth actual average wage, $W^a_{i,j}$ for each of the job types and respondents. This allows us to document misperceptions about what different jobs typically offer.

In addition to asking about typical (starting) wages, the survey also included additional questions about typical later wage growth and about the likelihood of getting hired into the different jobs if applying. We return to misperceptions in these dimensions when discussing additional results in Section 5.3.3.

2.2.3 Information treatment and misperceptions about offered wages

After eliciting workers beliefs about the typical wages offered in the different treatment. The survey included a randomly information treatment, aimed at examining whether wage misperceptions causally affect job search decisions. Half of the respondents were randomly assigned to treatment, with the rest acting as controls. Respondents in the treatment group were shown the true actual average wage, for the three job types they had been asked about $W^a_{i,j}\ \forall j \in \{1, 2, 3\}$. These actual average wages were shown visually next to the respondents own guess. To keep the survey experiment comparable across treatment and control, the control group were shown an analogous figure only containing their own guesses (thus

[16] The wording of the question was (translated from Danish):

"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 is newly graduated and completed their education in the 2010s whereafter they began working full-time.

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

Where $custom_educ$ is the name of their education.

[17] 10 gift cards worth 1,000 DKK would be distributed randomly among respondents who finished the survey. 3 further gift-cards were given to the 3 respondents who answers were the closest to the the average calculated from register data. 4 and 3 gift cards were distributed to respondents who answered closest to correctly for similar questions regarding, respectively, the average likelihood of getting a job given an application, and the average earnings 5 years after graduation, given initially working in the job type.

receiving no additional information about wages). Examples of the treatment is shown in Figures A2 and A3.

After the information treatment step, all respondents were then asked questions about their planned job search and expected wage offers. Specifically, participants are asked to imagine three different jobs, one from each job type, that are representative of what they think a job of that job type would offer them. Participants are then asked what they expect the offered starting wage to be. We refer to their answers here as $\widetilde{W}_{i,j}$. Note that these expected wage offers may differ from the previously elicited beliefs about average typical wages for at least two reasons: either because the individual expects themself to differ from the average worker, or because they expect wage offers in the future to differ from what was paid in the past. For the purpose of making future job search decisions, the relevant variable is of course wage offer the respondent expects to receive. As will become clear however, workers beliefs about wages offered to them is strongly correlated with beliefs about typical wages in the past. Accordingly, our information treatment about typical wages in the past generates exogenous variations in beliefs about offered wages, which we leverage to estimate the causal effects of beliefs on search.

To measure job search intentions, respondents are asked to imagine a hypothetical scenario, where they the following day will find 3 vacancy postings each corresponding to one of the job types. The are asked to imagine that they only have time to apply for one of the jobs and then answer the likelihood they would apply to each job, including the likelihood that they would not apply to any of the jobs. Answers to this question will be referred to as $\kappa_{i,j}$. The answers were recorded using three sliders that were adjusted to make sure the probabilities always summed to 100% by automatically adjusting the other sliders. They were all bounded between 1 and 97%. This likelihood will be referred to as $\kappa_{i,j}$ for person i in job j, with $j = 0$ being the likelihood of not applying for any of the three jobs.

2.2.4 Winsorizing

All numeric variables that were entered in free fields, like perceived earnings and worked hours were winsorized at the 2.5th and 97.5th percentile on the eligible for treatment sample who finished the survey. Winsorizing is quite common in the literature (see e.g. Epper et al. (2020); Hvidberg et al. (2023); Roth et al. (2022)). In a similar environment Wiswall and Zafar (2018) does not winsorize outliers, but instead uses a Least absolute deviations (LAD) estimator, which essentially estimates the median effect instead of the average, which means it is not sensitive to outliers. The LAD-estimator is not easily combined with the IV methods used in the analysis and we therefore opted for winsorizing. In the survey, the elicited choice probabilities were also bounded at minimum to be one, in order to avoid issues arising from rounding when the true probability is close to zero, and the log is taken. This issue is mentioned in Blass et al. (2010).

2.3 Descriptive statistics

Table 1 presents descriptive statistics for the survey respondents collectively and split on the different population sources.

The university graduate population from UCPH all registered as master graduates which fits with the sampling condition. They are subsequently better educated than the STAR sample, they are also slightly younger and more likely to be employed. They are all are either still studying or recently graduated, which again fits their sampling condition. In the STAR sample on the other hand, 60% recently graduated, and a much larger proportion are actively searching for a job.

One of the ways in which this paper adds to the literature on misperceptions about the labor market is through the large share of the total sample, 65%, that are actively searching for a job. Most earlier papers that have documented misperceptions about earnings like Jäger et al. (2024), use a sample of employed workers. Active job seekers have a higher incentive to be better informed about the labor market, and should therefore be less likely to be misinformed. Mueller et al. (2021) use a sample of job seekers, but document misperceptions on job finding rate as opposed to earnings. Other papers elicit beliefs from students before they start searching for full-time jobs, like Conlon and Patel (2023) and Cortés et al. (2024).

Table A2 in the appendix, shows that the randomization for treatment worked as expected, showing no systematic difference between the eligible control and treatment samples.

3 The extent of wage misperceptions

Figure 2 illustrates the misperceptions regarding mean earnings for each job type among survey respondents, relative to the administrative data. The figure plots a histogram of the log ratio of the perceived average monthly earnings in each job type, $\widetilde{W}^a_{i,j}$, over the actual average calculated in the data, $W^a_{i,j}$, along with its empirical cumulative distribution. The figure shows a large discrepancy between respondents' beliefs about average earnings and the actual average earnings based on the register data. While the median respondent slightly overestimates the earnings, the figure shows a lot of dispersion in both sides.

The survey was not set up to discover absolute misperceptions, (recall that respondents were shown an anchor of an overall average wage). So this figure is not sufficient to say something about the absolute level of misperceptions, whether job seekers on average over or underestimate the average earnings, as their answers were anchored to the given information about the overall average. Misunderstandings in the inclusion of taxes and benefits, extra hours, etc., could also play a role. The main point of figure 2 is to show the dispersion in beliefs relative the actual averages.

Table 1: Descriptive statistics

	UCPH only	UCPH and STAR	STAR only	All	Eligible
Total invited	2,609	366	43,256	46,231	
Total answers	482	72	6,609	7,163	2,928
Completed answers	296	45	4,019	4,360	1,941
Has custom jobtypes	243	33	2,485	2,761	1,941
Eligible for treatment	173	30	1,738	1,941	1,941
Age	27.89	27.73	28.65	28.59	28.10
Female	0.65	0.71	0.63	0.63	0.63
Higher education	1.00	1.00	0.81	0.83	1.00
Masters	1.00	1.00	0.43	0.48	0.64
Graduated at most 2 years ago	0.61	0.87	0.60	0.60	0.70
Currently studying	0.39	0.11	0.05	0.07	0.07
Expect to graduate in 2 months	0.24	0.11	0.01	0.03	0.04
Expect to graduate in 1 year	0.38	0.11	0.02	0.05	0.05
Employed	0.53	0.09	0.18	0.20	0.19
Active job searcher	0.36	0.62	0.70	0.68	0.65
Active job searcher and neither employed nor studying	0.15	0.58	0.60	0.57	0.56
Studying and employed	0.25	0.07	0.01	0.03	0.03
Not studying and not employed	0.33	0.87	0.79	0.76	0.77
Studying and active job searcher	0.13	0.04	0.02	0.03	0.03

Notes: UCPH is the population made available by the University of Copenhagen of person close to finishing their masters. STAR is the population of persons that have signed up for unemployment benefits in the summer of 2023. The eligible column only includes respondents who had an education that was eligible for treatment, meaning we had enough historical observations from the education to provide an informative treatment about the transition, as described in Section 2.2.1.

The focus of this section is to document misperceptions regarding *relative* wage differences between jobs. Looking at relative wages between two jobs takes out misunderstandings in how the level is supposed to be interpreted. For example, if a respondent thinks job j pays more than job j', they will state this, whether they think taxes or pensions are included or not. One issue might arise if respondents believe benefit levels differ across jobs and misunderstand which to include, but this should be a minor concern since the survey explicitly clarified this aspect.

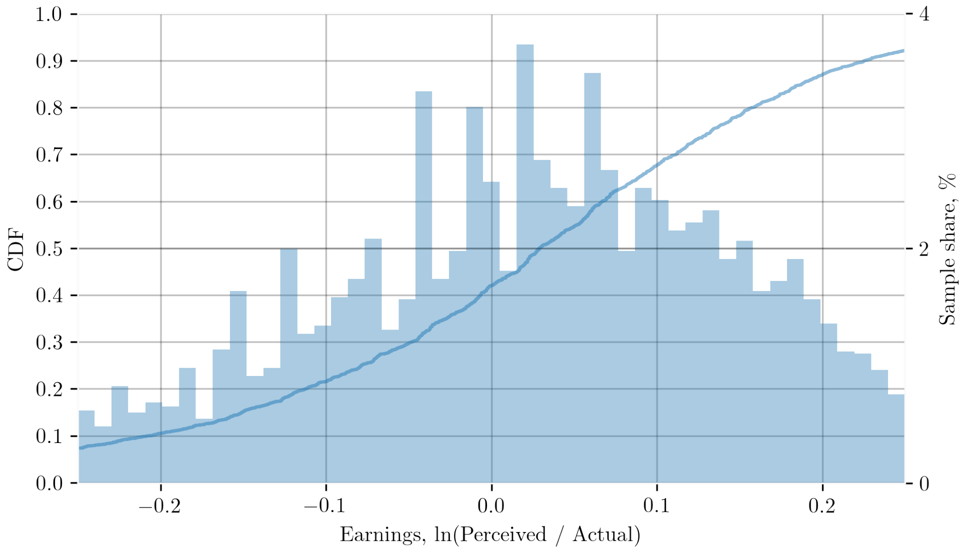
In order to document misperceptions in the earnings difference between jobs, we calculate a measure that, for all respondents, i, and all pairs of jobs j, j', computes the perceived gap in average earnings between jobs, relative to the actual gap in the register data. We call this the Perceived Difference in Percent of the Actual:

$$\text{PDPA}_{i,j} = \frac{\widetilde{W}^a_{i,j} - \widetilde{W}^a_{i,j'}}{W^a_{i,j} - W^a_{i,j'}} \tag{1}$$

The jobs are ordered such that the denominator is always positive.[18] The measure gives a sense of how the respondent perceives the earnings comparison between the two jobs,

[18]So for each respondent there are three jobs and we calculate three gaps: Highest paying vs. medium paying, highest paying vs. lowest paying, and medium paying vs. lowest paying.

Figure 2: Misperceptions about earnings for each job



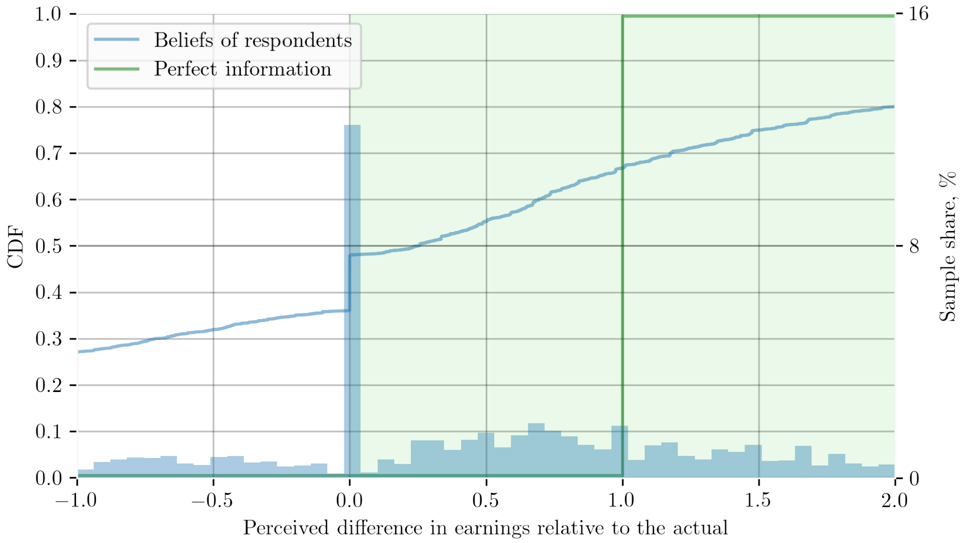
Note: This figure plots empirical cumulative distribution function (left axis) and the histogram (right axis) of the log ratio of perceived to actual initial earnings in each job, $\ln \widetilde{W}^a_{i,j} - \ln \widetilde{W}^a_{i,j}$. The sample is the survey population from educations that were eligible for treatment.

relative to the comparison in the register data. If the measure is positive the respondent is correct about which of the two jobs pays the most. If $\text{PDPA}_{i,j} \in (0,1)$, respondent i knows which job pays the most, but thinks that the difference is smaller than what it actually is. $\text{PDPA}_{i,j} = 1$ means the respondent knows the exact gap between jobs. $\text{PDPA}_{i,j} > 1$ means the respondent overestimates the gap between the jobs.

Figure 3 plots the distribution of the Perceived Difference in Percent of the Actual, $\text{PDPA}_{i,j}$. The shaded green areas highlights the respondents who get the ranking right, and the green line shows a hypothetical distribution if all respondents knew the actual relative earnings between jobs. The plots show that there are substantial misperceptions, 37%, think that the lower paying job, pays more. The median belief is 25% of the true gap, understating the dispersion in wages. There is also not a substantial mass around the truth at 1, but rather the mass is around $0.6-0.7$, i.e. answers that underestimate the gap by around $30-40\%$.

Around 11% of the sample answer that they think the jobs pay the exact same, this seems to be a rounding down problem, such that jobs the respondents thinks are roughly similar, are entered as paying the exact same. The $\text{PDPA}_{i,j}$ measure does not say much for these respondents, as the actual difference between the jobs might be very small, which means that guessing a difference of zero might not be a bad guess. In general small actual gaps, might overstate the level of misperceptions, as it would cause the denominator in equation (1) to be very small. Therefore, Figure 4 plots $PDPA_{i,j}$, but conditions on the size of the actual gap in the registers. Conditioning on the difference between the jobs

Figure 3: Perceived difference in percent of the actual



Note: This figure plots empirical cumulative distribution function (left axis) and the histogram (right axis) of the perceived Difference in Percent of the Actual for the initial earnings: $\text{PDPA}_{i,j} = \frac{\widetilde{W}^a_{i,j} - \widetilde{W}^a_{i,j'}}{W^a_{i,j} - W^a_{i,j'}}$. The green shaded area highlights respondents who correctly ranked the earnings differences between the two jobs. Perfect Information plots the CDF in the hypothetical scenario, where all respondents had correct beliefs. The sample is the survey population from educations that were eligible for treatment.

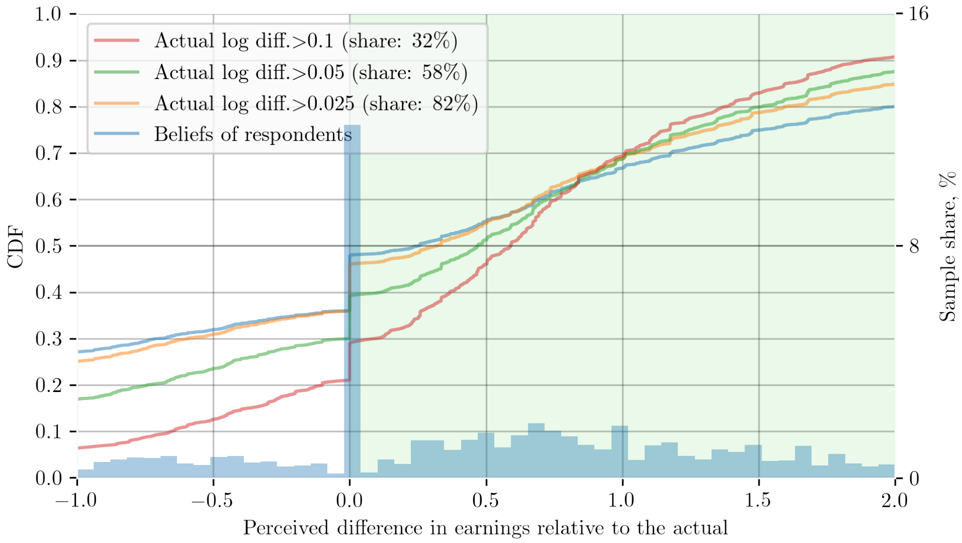
being at least 2.5 log points, removes 18% of the observations. The remaining observations show slightly less misperception, 37% still get the ranking wrong. But the ones that know the right ranking, guess closer to the true gap in the data. Conditioning on a very large actual gap of 10 log points, removes 68% of the observations. Still despite the gap being very large, 20% think the difference between the jobs is the opposite, and 10% think there is no difference, in spite of the substantial actual difference.

The PDPA measure is less interpretable when the actual gap is very small, therefore Figure B6 plots the perceived log difference between jobs, conditional on the size of the actual gap. When the actual difference between two jobs is smaller than 2.5 log points, 20 % answer that the difference is 0, while the remaining answers are widely distributed between -20 and 20 log points.

Despite the significant misperceptions, there is a correlation between beliefs and data. Figure B1, a binned scatter plot of $\ln \widetilde{W}^a_{i,j}$ and $\ln W^a_{i,j}$ shows an overall robust correlation, while Figure B2 plots a binned scatter plot of the perceived differences, $\left(\ln \widetilde{W}^a_{i,j} - \ln \widetilde{W}^a_{i,j'}\right)$ against the actual differences $\left(\ln W^a_{i,j} - \ln W^a_{i,j}\right)$. The correlation is small when the data gap is smaller than around 6%, but there is a clear correlation for higher gaps.

The appendix includes further robustness checks: Figure B5 shows that answers are not tied to the anchor we gave the respondents. Figure B3 shows a version of Figure 4 where the difference that is being conditioned on, is in absolute terms. Figure B4 shows the perceived log difference in percent of the actual difference showing a similar picture.

Figure 4: Relative misperceptions gap, conditioning on the actual gap



Note: This figure plots empirical cumulative distribution function (left axis) and the histogram (right axis) of the perceived Difference in Percent of the Actual for the initial earnings: $\text{PDPA}_{i,j} = \frac{\widetilde{W}^a_{i,j} - \widetilde{W}^a_{i,j'}}{W^a_{i,j} - W^a_{i,j'}}$. The green shaded area highlights respondents who correctly ranked the earnings differences between the two jobs. The actual log diff. is the log difference between the average earnings in the jobs in the register data, $\left(\ln W^a_{i,j} - \ln W^a_{i,j}\right)$. How big a share of the initial sample that is still included after the condition is stated in the label. The sample is the survey population from educations that were eligible for treatment.

Table B2 and B2 lists the share of the sample that answers the correct ranking of which jobs that pays the most.

4 Wage misperceptions and search behavior: Theoretical framework and identification

Having found evidence of large wage misperceptions in the previous section, the rest of the paper examines whether these misperceptions have causal effects on actual job search behavior. In this section, we start by discussing a general theoretical framework and use it to derive relevant empirical specifications. Additionally we discuss potential concerns with the identification of causal effects, and how we use our randomized treatment to overcome them.

4.1 Belief updating

To quantify the exogenous variation in beliefs induced by the survey treatment, we set up a model of belief formation. It is common in the literature of information experiments to base the specification on a model of Bayesian learning and then add elements that are more based on econometric necessity than on the model but can be theoretically motivated (see e.g. Haaland et al. (2023), Roth and Wohlfart (2020), Jäger et al. (2024) Cullen and Perez-Truglia (2022), and Fuster et al. (2022). We will do the same in this section, but include the added elements directly in the modeling.

There are two notable deviations from standard practice in our approach. The first is not uncommon, namely that the elicited prior and information signal is not the exact same as what is asked about in the posterior. Instead, the elicited prior and information treatment pertains to the past average wage in each job type, while the 'posterior' is changes in beliefs about what the respondents themselves would make. This makes the prior/posterior language somewhat imprecise, and for this section the posterior will refer to the unobserved posterior belief about the average wage in the job. This divergence could mitigate 'Experimenter demand' effects as noted in Haaland et al. (2023),[19] and is for example also done in Haaland and Roth (2023). Second, rather than predicting the degree of updating, i.e. the difference between the posterior and some benchmark, we aim to predict the changes in the log level of the posterior, as that is what we wish to instrument. This is rarer, but is for example done in Cullen and Perez-Truglia (2022). It is also done in Haaland and Roth (2023) but their setting is quite different. We will show that predicting the level is assumes less about the updating rule than predicting the gap. First we'll show the simple updating model, and then an extended model that is more practical for estimation.

We assume that the wage a person, i would make in a given job type j is the mean wage in that job type, $\bar{W}_{i,j} = \exp(\bar{w}_{i,j})^{20}$, times some factors, a_i and $a_{i,j}$, that depends on how productive they are relative to the average and how well they match with the specific job type. $\bar{W}_{i,j}$ is the true mean, while $W^a_{i,j} = \exp(w^a_{i,,j})$ is the calculated average, corresponding to the signal we will give to the treated.

We assume wages for a given person in a given job type, will be the true mean times a person-specific ability adjustment and a person-job-specific adjustment:

$$W_{i,j} = \exp(a_i)\exp(a_{i,j})\cdot\bar{W}_{i,j} \tag{2}$$

$$\Leftrightarrow w_{i,j} = a_i + \bar{w}_{i,j} + a_{i,j} \tag{3}$$

We assume that a_i and $a_{i,j}$, which represent the individual fixed effect and the job-specific match effect, are both known to the respondent,[21] and that the respondent's prior is that $\bar{w}_{i,j}$ is normally distributed: $\bar{w}_{i,j}|\tilde{w}^a_{i,j} \sim \mathcal{N}(\tilde{w}^a_{i,j}.\frac{1}{\rho})$, $\tilde{w}^a_{i,j}$ is the log of their stated prior about

[19] The issue can be summarized as the following: if a survey asks the respondents about what they think some value, A, is, then tell the treated the true value of A, and then again ask everybody what they think the value of A is. The treated will feel expected to answer their treatment value, and the non-treated will be confused that they are asked about the same thing twice.

[20] All variables that pertain to a job type, retains the i subscript, as the job types are education specific. For two different people $j = 1$ thus only refers to the same job type if they have the same education.

[21] One could add uncertainty about a_i and $a_{i,j}$, but, as long as one assumes that the signal does not change beliefs about a_i and $a_{i,j}$ the interpretation is the same.

what the past average wage is, which is assumed to also be their belief about the true mean. ρ is the precision of the distribution, indicating their degree of certainty about the true mean. When treated with information about the past average, this is a noisy signal $w^a_{i,j} = \bar{w}_{i,j} + n_{i,j}$, with the noise being $n_{i,j} \sim \mathcal{N}(0, \frac{1}{\rho_n})$. Both ρ and ρ_n are assumed to be commonly known. The posterior after a signal is then normally distributed as:

$$\bar{w}_{i,j} | \{\tilde{w}^a_{i,j}, w^a_{i,j}\} \sim \mathcal{N}\big(\rho', \tilde{w}^{a,p}_{i,j}\big) = \mathcal{N}\left(\rho + \rho_n, \frac{\rho_n}{\rho+\rho_n} w^a_{i,j} + \frac{\rho}{\rho+\rho_n} \tilde{w}^a_{i,j}\right) \tag{4}$$

This is a known result, with a proof provided in appendix D.1. Which is to say the posterior about the mean, is a weighted average of the prior and the signal, were the weights depend on the relative certainty about the precision of them. It is often written as:

$$\underbrace{\tilde{w}^{a,p}_{i,j} - \tilde{w}^a_{i,j}}_{\text{Updating}} = -\frac{\rho_n}{\rho+\rho_n} (\underbrace{\tilde{w}^a_{i,j} - w^a_{i,j}}_{\text{Perception gap}}) \tag{5}$$

The left-hand side, is then how much the posterior is updated, and the right hand side is the difference between the signal and the prior. It is common to use this formula as the basis of the first stage, and plot $(\tilde{w}^{a,p}_{i,j} - \tilde{w}^a_{i,j})$ against $(w^a_{i,j} - \tilde{w}^a_{i,j})$ for the treated and the non-treated to visually show the degree of updating.

There can be reversion or spurious learning towards the truth for the non-treated, so the proceeding regression includes both perception gap interacted and not interacted with a treatment dummy, this also ensures identification, as the research design typically ensures that the treatment dummy is randomized, while the prior perception gap is not (Haaland et al. (2023)). If ϱ is the degree of spurious learning or reversion towards the signal without treatment, and we assume that spurious learning happens for both the treated and non-treated, we can then write the posterior about the average wage as:

$$\tilde{w}^{a,p}_{i,j} - \tilde{w}^a_{i,j} = -\varrho(\tilde{w}^a_{i,j} - w^a_{i,j}) - \frac{\rho_n}{\rho+\rho_n} T_i (\tilde{w}^a_{i,j} - w^a_{i,j}) \tag{6}$$

Where, equation (6) suggests a regression of predicting the degree of updating relative to the prior information gap, in the mean wage. To create a formula for expectations about the *level* of the wage the respondent would earn *themselves* in the job, we start from (3) and insert the posterior belief about $\bar{w}_{i,j}$, i.e. $\tilde{w}^{a,p}_{i,j}$:

$$\tilde{w}_{ij} = a_i + \tilde{w}^a_{i,j} - \varrho(w^a_{i,j} - \tilde{w}^a_{i,j}) - \left(\frac{\rho_n}{\rho+\rho_n}\right) T_i (w^a_{i,j} - \tilde{w}^a_{i,j}) + a_{i,j} \tag{7}$$

This restricts the coefficient on $\tilde{w}^a_{i,j}$ to unity, which either suggests dropping it, or allowing it to differ from 1, but including it in the regression might prompt the question of having it interacted with the treatment dummy.

Another approach to include more flexibility, is to follow Cavallo et al. (2017), and use a flexible belief formation instead of (3). This is essentially a behavioral assumption. To assume that that while $w_{i,j}$ is determined by (3), beliefs about $w_{i,j}$ is some flexible beliefs function, f^b:

$$\tilde{w}_{i,j} = f^b(\tilde{w}^{a,p}_{i,j}, T_i, a_i, a_{i,j}) \equiv a_i + \psi_1 \tilde{w}^{a,p}_{i,j} + \psi_2 T_i \tilde{w}^{a,p}_{i,j} + a_{i,j} \quad (8)$$

Treatment is included here because it might change certainty about $\bar{w}_{i,j}$, and this is allowed to influence how much their belief about the mean, affects their belief about $w_{i,j}$. This is the same as (3) if $\psi_1 = 1, \psi_2 = 0$. We then insert the posterior from (6) to get the general formula:

$$\tilde{w}_{i,j} = a_i + \psi_1 \tilde{w}^a_{i,j} - \psi_1 \varrho(w^a_{i,j} - \tilde{w}^a_{i,j}) - \left(\frac{\rho_n(\psi_1 + \psi_2)}{\rho + \rho_n} + \psi_2 \varrho\right) T_i(w^a_{i,j} - \tilde{w}^a_{i,j}) + \psi_2 T_i \tilde{w}^a_{i,j} + a_{i,j} \quad (9)$$

This equation outlines different ways of specifying the first stage, depending on how one assumes beliefs are formed. The most flexible approach would be to use (9) as the first stage without any changes. Starting from the above and setting $\psi_2 = 0$, would be taking the approach of Cullen and Perez-Truglia (2022). Other papers, like Jäger et al. (2024) and Roth and Wohlfart (2020), predict the gap on the left-hand side, which essentially forces ψ_1 to be equal to 1. They also set $\psi_2 = 0$. When doing level-prediction, ψ_1 could be allowed to vary, like Cullen and Perez-Truglia (2022), or one could simply exclude $\tilde{w}^a_{i,j}$ from the regression. Haaland and Roth (2023) predict in levels and essentially takes the exclusion approach, however, their prior and posterior variables are so different in definition, that there is no reason to think that the level should be predictive, which is not the case in our setting. As noted in section 4.2 all these approaches satisfy the independence assumption required for a valid IV instrument, however they should not be expected to be equally efficient.

Throughout, we assume a linear, homogeneous updating rate. Using the methodology of Cattaneo et al. (2024), we formally test this in Appendix F, and find no evidence to the contrary.

4.2 Empirical approach

The empirical approach is motivated by a discrete choice discrete choice logit model, following McFadden and Train (2000) and Train (2009), where individuals select which job to apply for based on their perceived utility. They way we relate choice probabilities to the survey questions, follow the approach of Blass et al. (2010) and Wiswall and Zafar (2018). The model is formulated in Appendix E.

The functional form derived from the model specifies that the log of the likelihood of applying to a job relative to not applying to any job, is linear in the log of the perceived wage and the the log of the perceived likelihood of being offered the job. It also specifies that a person fixed effect should be included in the regression.

We denote this relative log likelihood as $k_{i,j} = \ln \kappa_{i,j} - \ln \kappa_{i,0}$, where $\kappa_{i,j}$ is the likelihood that i chooses to apply for j, and $\kappa_{i,0}$ is the likelihood that i chooses not to apply for any job. These variables correspond to the elicited stated likelihood explained in section 2.2.3.

Our specification will use the derived equation from the model as the second stage equation:

$$k_{i,j} = \mu_i + \beta_1 \hat{\tilde{w}}_{i,j} + \beta_2 \ln \tilde{p}_{i,j} + \beta_3 \tilde{w}^a_{i,j} + \beta_4 (\tilde{w}^a_{i,j} - w^a_{i,j}) + u^{\text{2stage}}_{i,j} \tag{10}$$

$\tilde{p}_{i,j}$ is the perceived likelihood of being offered the job if the respondent had applied to the job, which we also elicit in the survey. μ_i is the person fixed effect. Controlling for the log of the perceived average wage, $\tilde{w}^a_{i,j}$, and the log gap between the perceived and actual average, $(\tilde{w}^a_{i,j} - w^a_{i,j})$, has no structural interpretation but are necessary to include if they are present in the first stage.

β_1 is our main parameter of interest as it explains the relationship between the the perceived wage with the likelihood of applying for the job. Appendix E.1 shows how it β_1 can be adjusted to estimate the elasticity between the two.

$\hat{\tilde{w}}_{i,j}$ is predicted from the first stage from (9), adding $\ln \tilde{p}_{i,j}$ as a control from the first:

$$\tilde{w}_{i,j} = a_i + \tau_1 \ln \tilde{p}_{i,j} + \tau_2 (w^a_{i,j} - \tilde{w}^a_{i,j}) + \tau_3 T_i (w^a_{i,j} - \tilde{w}^a_{i,j}) + \tau_4 \tilde{w}^a_{i,j} + \tau_5 T_i \tilde{w}^a_{i,j} + u^{\text{1stage}}_{i,j} \tag{11}$$

The independence of both instruments is ensured by the exogeneity of the treatment dummy as it was randomly assigned and by controlling for the interacted variables. However, the exclusion restriction, i.e. the assumption that treatment does not affect $k_{i,j}$ through other factors than wage is not. This does not hold if treatment is correlated with beliefs about other characteristics of the job type, if this assumption does not hold, the interpretation of the estimate of β_1 changes slightly as will be discussed in section 4.2.1.

The independence still holds for $T_i(w^a_{i,j} - \tilde{w}^a_{i,j})$, if $T_i \tilde{w}^a_{i,j}$ is not included as an instrument, this is relevant because $T_i \tilde{w}^a_{i,j}$ is possibly a weak instrument if ψ_2 in equation (10) is close to zero, i.e. treatment does not affect the degree to which beliefs about the mean wage of the job affects beliefs about the wage the respondent think they would make in the job. Because the types of jobs presented are different across educations the standard errors are clustered at the education level.

The inclusion of $\tilde{w}^a_{i,j}$ as a control does not affect the inclusion restriction, but it can increase the efficiency of the estimate by decreasing the amount of noise in the first stage, as it's inclusion is guided by updating rule discussed in the previous section. Allowing its associated coefficient to differ from unity essentially adds flexibility to reduce noise. It is therefore our preferred specification to include this, while excluding . In the empirical section, the different approaches will be compared.

4.2.1 Threats to exclusion restriction

Given the randomly controlled trial (RCT) design in the survey, the treatment effect on beliefs is cleanly causally identified, relying only on the random assignment of treatment.

However, interpreting and estimating β_1 requires further assumptions, which, if violated breaks the exclusion restriction of the IV-approach, which compromises the second stage's interpretation as solely reflecting the effect of perceived wage on the log relative likelihood of applying for a job, and extend the interpretation to include other changes in beliefs associated with the information treatment and their effect on the log relative likelihood.

Spillovers to other amenities If the information about past average wages affects beliefs about other amenities about the job, this would contaminate the estimate of β_1. If there are spillovers, the estimated elasticity of perceived wage on the relative application likelihood would include changes induced by changes in beliefs about other amenities. However, these changes in beliefs would likely also occur in the real world, e.g. if a firm posts a vacancy with a higher wage and this causes beliefs about other amenities to change as well.

The survey asks about beliefs about some amenities after the treatment, which allows us to test the spillover assumption for these amenities, by running the first stage regression with the amenity as the outcome variable. This is described in section 5.3.2, where it is shown that there does not appear to be spillovers on hours worked, likelihood of getting a job offer conditional on application, how well they will get along with their colleagues, nor their own performance in the job.

Something that is not tested for explicitly, is spillovers to beliefs about the value of being unemployed. It is likely that information about expected wage should change the continuation value of being unemployed, however this is dealt with using the fixed effects approach as it takes out the average change in application likelihood relative to not applying for any of the jobs.

Spillover to other jobs One could also worry that the treatments are correlated, so treatment about job j' also affects beliefs about j. One way to check this is to include the misperception gap in the other 2 jobs, interacted with treatment, as controls in the first stage regression, to test if they have a causal effect on belief about the wage after controlling for the information treatment of the job. This is discussed further in Appendix C.2, and is empirically shown to not be an issue. As the treated are given information about all three job types, any spillover to the other two jobs is also likely attenuated by the direct information given about the job. We cannot test whether beliefs about all other jobs types are changed, but any behavioral effect in the second stage that influences all jobs, like changes in the outside option, is captured by the person fixed effect.

5 Wage misperceptions and search behavior: Results

Here we present our results regarding the causal effects of wage misperceptions and information on job search, building on the theory and empirical specifications from the preceeding section.

5.1 First stage: Belief updating

We first examine whether and how the information provision in our treatment affected beliefs about the wages individuals would be offered in the different job types. In addition to being of independent interest, this relationship forms the first stage of our IV approach to estimating the causal effect of wage beliefs.

Starting from the general equation (11) and imposing our preferred restrictions, our preferred (first stage) specification can be written as follows:

$$\tilde{w}_{i,j} = a_i + \tau_1 \ln \tilde{p}_{i,j} + \tau_2 (w^a_{i,j} - \tilde{w}^a_{i,j}) + \tau_3 T_i (w^a_{i,j} - \tilde{w}^a_{i,j}) + \tau_4 \tilde{w}^a_{i,j} + u^{1\text{stage}}_{i,j} \tag{12}$$

The key coefficient of interest is the learning rate, τ_3. It measures how respondents belief about wages offered in job j changes if our information treatment informed them that their intial belief overestimated the average typical wages by 1 percent (1 log point).

Column (4) of Table 2 presents the regression results from our preferred specification. The coefficient of interest is estimated to be 0.329 and significantly different from zero. Our information treatment thus lead respondents to update their belief in the expected direction. In terms of magnitude, the estimated learning rate is smaller than what has been found in for example Jäger et al. (2024) and Cullen and Perez-Truglia (2022) who finds rates about $50-60\%$. This is likely due to the signal variable not being the same as the posterior variable. Haaland and Roth (2023), in a different topic, find large variability in learning rates when there is a discrepancy between the signal and the posterior variable.

The other columns of Table 2 present alternative specifications, all of which show estimates in the range of 0.33-0.38 for the parameter of interest. Columns (1)-(2) are included mostly for completeness and are atheoretical, in the sense that they exclude the person-fixed effects suggested by our learning framework and the model shown in Appendix E. Column (3) is the Jäger et al. (2024) and Haaland and Roth (2023)-like approach, with $\psi_2 = 0$ and excluding $\tilde{w}^a_{i,j}$ from the regression in equation 9. Despite leaving out the highly predictive $\tilde{w}^a_{i,j}$, the precision in the estimate of the learning rate in this columns is only marginally lower than the preferred estimates. Column (5) includes the extra treatment effect from the level of the perceived average wage, which is not significant, suggesting it should not be included as an instrument.

As noted in Section 4.2.1, a potential concern with this first stage is learning spillovers to both other amenities and to other jobs. As also discussed however, we have evidence that this is not the case, at least for the subset of amenities where we have data (see Section

Table 2: The effect of log perceived average wage on log perceived wage

	ln perceived wage, $\tilde{w}_{i,j}$				
	(1)	(2)	(3)	(4)	(5)
Treatment, T_i	-0.009	-0.016**			
	(0.006)	(0.007)			
$\ln \tilde{p}_{i,j}$	0.009*	0.007***	0.002	0.004**	0.004**
	(0.004)	(0.002)	(0.004)	(0.002)	(0.002)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	0.627***	0.560***	0.499***	-0.129**	-0.134**
	(0.047)	(0.031)	(0.046)	(0.052)	(0.054)
$T_i \times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	-0.381***	-0.353***	-0.310***	-0.329***	-0.317***
	(0.048)	(0.038)	(0.045)	(0.043)	(0.043)
ln perceived avg. wage, $\tilde{w}^a_{i,j}$				0.844***	0.851***
				(0.045)	(0.051)
$T_i \times$ ln perceived avg. wage, $T_i \tilde{w}^a_{i,j}$					-0.015
					(0.042)
Education FE		✓			
Person FE			✓	✓	✓
Constant	✓				
Educations	83	82	82	82	82
Persons	1895	1894	1859	1859	1859
N	5515	5514	5479	5479	5479

This table shows the first stages regressions, estimating the degree of updating in the log perceived wage of the job, given a signal about the average wage. $\tilde{p}_{i,j}$ is the perceived probability of receiving an offer if the respondent applies to the job type. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

5.3.2). We also find no evidence that there is learning about about the wages in other jobs in section C.2.

Finally, the coefficient on $\ln \tilde{p}_{i,j}$ interestingly shows a slightly positive relationship between perceived own wages and perceived likelihood of a job offer given an application. However Table C3 shows that there is no significant correlation when only controlling for fixed effects, consistent with column (3).

Table C4 shows that inclusion of $\ln \tilde{p}_{i,j}$ is not important for the results.

5.2 The effect of perceived wage on job search

Next, we turn to the key question of how wage (mis)perceptions causally affect job search outcomes. We estimate this using an IV approach. Restating it from Section 4.2, the preferred specification (10) is:

$$k_{i,j} = \mu_i + \beta_1 \hat{\tilde{w}}_{i,j} + \beta_2 \ln \tilde{p}_{i,j} + \beta_3 \tilde{w}^a_{i,j} + \beta_4 (\tilde{w}^a_{i,j} - w^a_{i,j}) + u^{\text{2stage}}_{i,j} \tag{13}$$

The coefficient of interest is β_1 which measures the effect of (mis)perceiving that job j will offer a 1% higher starting wage, on the log relative likelihood of applying to job j, k_{ij}.

In estimation, we instrument the perceived wage $\tilde{w}^a_{i,j}$ leveraging the information treatment, with (12) serving as the first stage.

Table 3: The effect of log perceived wage on log relative application ratio, IV

	Log likelihood of applying relative to not applying, $k_{i,j} \equiv \ln \frac{\kappa_{i,j}}{\kappa_{i,0}}$			
	(1)	(2)	(3)	(4)
ln perceived wage, $\tilde{w}_{i,j}$	1.846**	4.684***	4.757***	4.671***
	(0.940)	(1.314)	(1.252)	(1.236)
$\ln \tilde{p}_{i,j}$	0.090**	0.534***	0.535***	0.536***
	(0.040)	(0.042)	(0.042)	(0.042)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	0.998*	2.188***	1.406	1.382
	(0.592)	(0.667)	(0.977)	(0.978)
ln perceived avg. wage, $\tilde{w}^a_{i,j}$			1.002	1.074
			(1.329)	(1.318)
Person FE		✓	✓	✓
Constant	✓			
Instruments:				
$T_i(\tilde{w}^a_{i,j} - w^a_j)$	✓	✓	✓	✓
$T_i \tilde{w}^a_{i,j}$				✓
T_i	✓			
First stage F(K-P)	34.15	46.92	58.28	33.03
Average elasticity $\widehat{\mathcal{E}}^a$	1.274	3.245	3.296	3.236
Educations	83	82	82	82
Persons	1895	1859	1859	1859
N	5515	5479	5479	5479

This table shows second stage estimates instrumenting log of the perceived wage, $\ln \tilde{w}_{i,j}$, with the information treatment. The outcome variable is the log likelihood of applying for a job, relative to none. $\tilde{p}_{i,j}$ is the perceived probability of receiving an offer if the respondent applies to the job type. $\widehat{\mathcal{E}}^a$ is the sample-average elasticity of the likelihood of applying for a job with respect to the perceived wage suggested by the estimate. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Column (3) of Table 3 shows estimates from the preferred specification. The coefficient of interest is 4.757 and highly significant; (mis)perceiving that a given job pays more thus indeed increases the likelihood of applying. To most easily gauge the magnitude, the fourth last row of Table 3 converts the estimate to an average elasticity (accounting for the fact that the outcome variables stems from a discrete choice problem). Perceiving a 1% higher starting wage in a given job, increases the likelihood of applying by 3.296%. This suggests that beliefs starting wages have a large influence on stated job application behavior. Accordingly, wage misperceptions can significantly distort search behavior.

The estimated elasticity is substantially larger than what would be expected when compared to papers like Mueller et al. (2024), Belot et al. (2022b), and Marinescu and Wolthoff (2020), but are close to what is found in Bassier et al. (2023).

For completeness and robustness, the other columns of the Table 3 show results from alternative specifications. Column (1) shows the importance of the inclusion of the person fixed effect, as the estimate is substantially smaller. This is likely because the exclusion restriction fails to hold without the person fixed effect, as treatment can chain application intentions through changes in the perceived outside option.

The estimate of β_2 in the second row shows a positive relationship between the likelihood of receiving an offer and applying for a job. This is purely correlational, as $\ln \tilde{p}_{i,j}$ is possibly endogenous in this regression, but the correlation is intuitive. Table C5 shows that excluding $\ln \tilde{p}_{i,j}$ does not change the magnitude of the wage elasticity.

5.3 Additional results

5.3.1 OLS and reduced form results

Figure 4 shows naive OLS estimates β_1 for comparison with the estimates in Table 3. Columns (1)-(4) include the same controls as in the second stage, while columns (5)-(6) exclude those controls, more directly resembling a natural OLS approach to estimating β_1, as they were motivated by the updating rule of the first-stage regression. Without the IV approach controlling for these variables changes the interpretation of β_1, as it uses variation in what the respondent think they would earn, while keeping information about the average wage fixed. The remaining variation could be interpreted as a form of variation in the perceived wage premium relative to other workers in the job type. This makes the coefficient in columns (5)-(6) the ones to focus on.

The estimate in column (5) is slightly larger than the IV, suggesting an upwards bias in the naive estimates.

Columns (3) and (6) introduces controls for a set of elicited amenities asked about in the survey, that the respondent expect if they were to get a job in the different job types. These amenities are what the respondent think will be the yearly wage growth, the hours worked in the job, how well they will get on with their colleagues, and how well they will perform in the job.

Column (6) shows that controlling for these amenities incidentally pushes the estimate of β_1 into the same ballpark of the IV estimate, however this result is sensitive to not controlling for all of the amenities, and might change if one was able to add more amenities. Interestingly, the estimate related to the perceived likelihood of getting an offer, is also slightly reduced when controlling for amenities, suggesting that $\tilde{p}_{i,j}$ is positively related to a positive valuation of the jobs based on these amenities.

Table 5 shows the reduced form version of Table 3. The estimate on the treatment interacted with the perception gap in column (4) suggests that telling the respondent that

Table 4: The association between log perceived wage and the log relative application ratio

	Log likelihood of applying relative to not applying, $k_{i,j} \equiv \ln \frac{\kappa_{i,j}}{\kappa_{i,0}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
ln perceived wage, $\tilde{w}_{i,j}$	0.970***	4.285***	3.497***	3.601***	5.149***	4.486***
	(0.209)	(0.317)	(0.363)	(0.375)	(0.298)	(0.323)
$\ln \tilde{p}_{i,j}$	0.098**	0.535***	0.541***	0.374***	0.526***	0.362***
	(0.041)	(0.043)	(0.043)	(0.039)	(0.043)	(0.040)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	1.393***	2.330***	1.058	0.970		
	(0.484)	(0.530)	(0.973)	(0.775)		
ln perceived avg. wage, $\tilde{w}^a_{i,j}$			2.059**	0.849		
			(0.843)	(0.750)		
Yearly wage growth				12.514***		14.880***
				(1.490)		(1.371)
ln Hours				1.924***		2.003***
				(0.352)		(0.357)
Colleagues				0.271***		0.285***
				(0.032)		(0.034)
Performance				0.368***		0.367***
				(0.030)		(0.030)
Persons FE		✓	✓	✓	✓	✓
Constant	✓					
N	5515	5479	5479	5346	5479	5346
Average elasticity $\widehat{\mathcal{E}}^a$	0.669	2.968	2.422	2.488	3.567	3.100

This table shows OLS estimates. The outcome variable is the log likelihood of applying for a job, relative to none. $\tilde{p}_{i,j}$ is the perceived probability of receiving an offer if the respondent applies to the job type. $\widehat{\mathcal{E}}^a$ is the sample-average elasticity of the likelihood of applying for a job with respect to the perceived wage suggested by the estimate. *Yearly wage growth*, uses perceived wages in 5 years and perceived initial wages to calculate an equivalent yearly wage growth. ln Hours is the log of the stated expected monthly hours worked, observations are excluded if wages in 5 years or hours were stated to be zero. *Colleagues* is how they think they would get on with their colleagues in the job type, scalled from 1-6. *Performance* is how well they think they would perform in the job type, scalled from 1-6. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

the average wage in a job type is 1% than they thought it was, increases the likelihood that they will choose to apply for that job relative to not applying for any of the three jobs, at a given point in time, by 1.6%. Given the large dispersion in misperception, this suggests substantial space for influencing application behavior.

5.3.2 Spillovers to other amenities

Table 6, directly tests whether the wage information treatment affects beliefs about other amenities of the job asked about in the survey, after the treatment. This is done by running the first stage regression, but with the other perceived amenities as the outcome variable,

Table 5: Reduced form regression of experiment

	Log likelihood of applying relative to not applying, $k_{i,j} \equiv \ln \frac{\kappa_{i,j}}{\kappa_{i,0}}$				
	(1)	(2)	(3)	(4)	(5)
Treatment, T_i	0.024	0.005			
	(0.089)	(0.086)			
$\ln \tilde{p}_{i,j}$	0.106**	0.141***	0.544***	0.554***	0.554***
	(0.042)	(0.031)	(0.048)	(0.044)	(0.043)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	2.181***	1.459***	4.527***	0.794	1.337
	(0.555)	(0.336)	(0.523)	(1.112)	(1.200)
$T_i\times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	-0.760**	-0.573	-1.453***	-1.566***	-2.699***
	(0.368)	(0.348)	(0.402)	(0.422)	(0.799)
ln perceived avg. wage, $\tilde{w}^a_{i,j}$				5.015***	4.286***
				(0.788)	(0.964)
$T_i\times$ ln perceived avg. wage, $T_i\tilde{w}^a_{i,j}$					1.499*
					(0.874)
Education FE		✓			
Person FE			✓	✓	✓
Constant	✓				
Educations	83	82	82	82	82
Persons	1895	1894	1859	1859	1859
N	5515	5514	5479	5479	5479

This table shows the reduced form regresions, estimating the effect of given a signal about the average wage, on the log likelihood of applying to a job, relative to none. $\tilde{p}_{i,j}$ is the perceived probability of receiving an offer if the respondent applies to the job type. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

instead of the perceived wage. In the appendix, Table C3 shows the base correlations between perceived earnings and these amenities, as a motivation for why there could be spillover effects. If the respondent knows these amenities are correlated with $\tilde{w}_{i,j}$, then when receiving a signal that makes them update beliefs about $\tilde{w}_{i,j}$, they might also update their beliefs about the other amenities of the job.

The first column shows a clear treatment effect on perceived wages five years after graduation. This follows from the argument that shifts in beliefs will also shift beliefs about future earnings through shifting the base from which the earnings will grow. The estimate is however smaller than the learning rate estimated in Table 2, suggesting that a 1% increase in beliefs about starting wages, is associated with less than a 1% increase in perceived wages 5 years after. The coefficient from column (2) supports this, as it suggests that information about the average wage being lower than the respondent had expected, causes the treated respondent to revise their estimate of the earnings growth upwards, undoing some of the initial wage drop, in the future. We do not consider this spillover a violation of the exclusion restriction, as it operates through the pecuniary incentive.

Columns (3)-(6) show that there seems to be no significant spillover effects on the other recorded amenities. *Appl likelihood* is the stated likelihood between 0 and 100%, that given

Table 6: Effect of log perceived wage on other job amenities

	(1) $\tilde{w}_{i,j,t+5}$	(2) Wage growth	(3) ln Hours	(4) Appl likelihood	(5) Colleagues	(6) Performance
ln perceived avg. wage, $\tilde{w}^a_{i,j}$	1.046***	0.047***	0.323***	-8.675	-0.089	-0.199
	(0.069)	(0.011)	(0.077)	(8.812)	(0.274)	(0.420)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	-0.250***	-0.024**	-0.109	7.136	0.785***	0.743*
	(0.078)	(0.010)	(0.067)	(8.855)	(0.284)	(0.429)
$T_i\times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	-0.243***	0.013**	-0.027	-1.025	-0.045	-0.223
	(0.046)	(0.006)	(0.029)	(4.358)	(0.250)	(0.369)
Person FE	✓	✓	✓	✓	✓	✓
N	5659	5659	5532	5706	5706	5706

This table shows first stage regressions, replacing the outcome variable with perceived job amentites. $\tilde{w}_{i,j,t+5}$ is the log of the perceived earnings in 5 years after starting to work in the job type. *Yearly wage growth*, uses perceived wages in 5 years and perceived initial wages to calculate an equivalent yearly wage growth. ln Hours is the log of the stated expected monthly hours worked, observations are excluded if wages in 5 years or hours were stated to be zero. *Appl likelihood* is the stated likelihood that, given that they had applied to the job, how likely it would be that they would receive a job offer, graded from 0-100. *Colleagues* is how they think they would get on with their colleagues in the job type, scalled from 1-6. *Performance* is how well they think they would perform in the job type, scalled from 1-6. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

that had they applied to the job, how likely it would be that they would have received a job offer.[22] *Colleagues* is how well they think they would get on with their colleagues in the job type, scaled from 1 to 6. And *Performance* is how well they think they would perform in the job type, scaled from 1-6. This test is suggestive evidence that our information experiment did not cause spill overs to beliefs about non-pecuniary amenities. However, the test does entirely exclude the possibility that there are information spillovers onto other non-recorded amenities. Table C1 shows that results are the same when not controlling for $\tilde{w}^a_{i,j}$.

Table C7 recreates the first stage regression of column (4) of Table 2, but includes controls for treatment interacted with the misperception gap in the two other jobs, to check whether there is a spillover of beliefs about wages between jobs. There does not seem to be spillovers across jobs, as the estimate associated with $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$ is not much different from the first stage without the controls, nor is there an effect from treatment interacted with the misperception in other jobs.

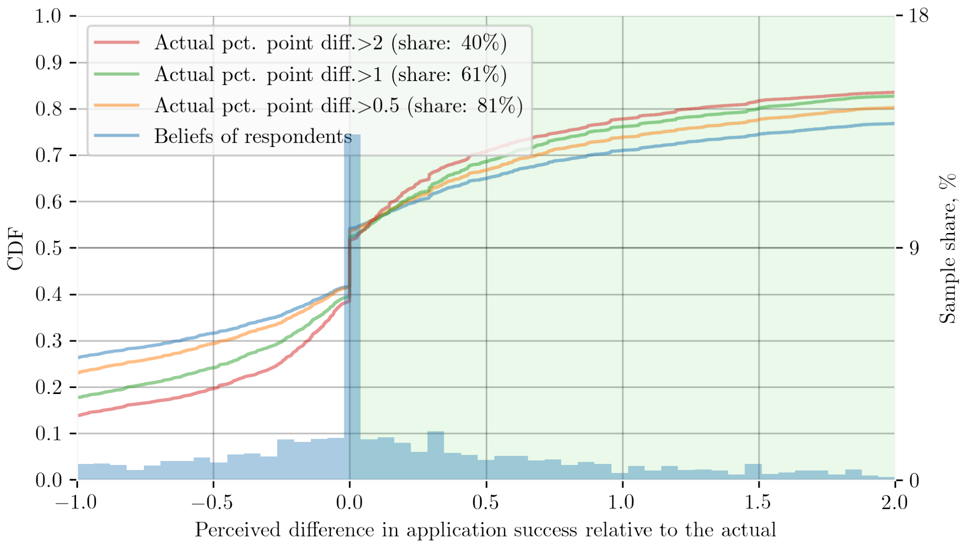
[22]Table C2 shows that replacing the log of perceived hours worked, with just the perceived hours worked causes the coefficient to be significant at the 10%-level, suggesting a slight increase in expected hours worked when the perceived wage is increased, this is however not robust to excluding answers that thought they would work 0 hours if they worked in the job job type. (These observations were not winsorized because just above 2.5% of the sample answers 0)

5.3.3 Misperceptions in application success

Figure 5 shows the PDPA measure, but for likelihood of a successful application instead of earnings.

Through data on logged application behavior of Danish unemployment insurance recipients and linked data to their later employment outcomes, we can calculate the rate of applications to firms that are successful, in the sense of leading to the person being hired in the firm, and average these for the different job types.

Figure 5: Relative misperceptions gap for likelihood of successful application



Note: This figure plots the histogram and empirical cumulative distribution function of the perceived difference relative to the actual likelihood of a successful application for different jobs. The green shaded area highlights respondents who correctly ranked the earnings differences between the two jobs. The sample is the survey population from educations that were eligible for treatment.

In the survey, prior to information treatment, we ask respondents to guess what they think this statistics will be, for each job type, to gauge how well they understand the likelihood of getting a job in each job type for unemployed persons.[23] Similarly to the question about average wages, we anchor the respondents by telling them the average likelihood of a successful application across educations and job types, this anchor was 2.7%.

If $\tilde{p}^a_{i,j}$ is the perceived average likelihood of successful application for person i in job type j, and $p^a_{i,j}$ is the actual past average for job seekers with person i's education that search for a job in job type j, we calculate the PDPA measure as:

$$\text{PDPA}_{i,j} = \frac{\tilde{p}^a_{i,j} - \tilde{p}^a_{i,j'}}{p^a_{i,j} - p^a_{i,j'}} \tag{14}$$

[23]The question was asked to be answered as the number of successful applications out of a 1000

This is the measure plotted in Figure 5. The likelihood of successful application is arguably a less prominent and more diffuse statistic, than initial earnings, and this suggestion is born out in the fact that the relative misperceptions are even more dispersed than earnings. A smaller share are able to point out the job where a successful application is more likely, and there is no noticeable mass around 1, where respondents correctly perceive the difference between jobs. Removing job comparisons where the actual difference is small, also only marginally alleviates the misperceptions.

6 Conclusion

Our findings show that young job seekers have substantial misperceptions about the relative starting wages in different jobs. In our survey, nearly 37% of respondents misrank the relative wages of job types and the median person strongly underestimates the true dispersion in earnings across job pairs. By providing tailored information on actual wages, we demonstrate that wage expectations are elastic to new information, and that individuals adjust their planned job search behavior accordingly, targeting jobs they perceive as offering higher pay. The elasticity of job application behavior with respect to wage expectations suggests that information frictions cause significant distortions early-career labor market sorting, as misperceptions cause workers to target suboptimal job options in their search.

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.

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.

Bassier, Ihsaan, Alan Manning, and Barbara Petrongolo, "Vacancy Duration and Wages," IZA Discussion Papers 16371, Institute of Labor Economics (IZA) August 2023.

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.

Blass, Asher A., Saul Lach, and Charles F. Manski, "Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability*," *International Economic Review*, 2010, *51* (2), 421–440. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2010.00586.x.

Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng, "On Binscatter," *American Economic Review*, May 2024, *114* (5), 1488–1514.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments," *American Economic Journal: Macroeconomics*, July 2017, *9* (3), 1–35.

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," Working Paper 2023. https://scholar.harvard.edu/devpatel/publications/what-jobs-come-mind-stereotypes-about-fields-study.

Cortés, Patricia, Jacob French, Jessica Pan, and Basit Zafar, "Gender Differences in Negotiations and Labor Market Outcomes: Evidence from an Information Intervention with College Students," Working Paper 32154, National Bureau of Economic Research February 2024.

Cullen, Zoë and Ricardo Perez-Truglia, "How Much Does Your Boss Make? The Effects of Salary Comparisons," *Journal of Political Economy*, March 2022, *130* (3), 766–822. Publisher: The 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.

Epper, Thomas, Ernst Fehr, Helga Fehr-Duda, Claus Thustrup Kreiner, David Dreyer Lassen, Søren Leth-Petersen, and Gregers Nytoft Rasmussen, "Time Discounting and Wealth Inequality," *American Economic Review*, April 2020, *110* (4), 1177–1205.

Falch, Torberg, "The Elasticity of Labor Supply at the Establishment Level," *Journal of Labor Economics*, 2010, *28* (2), 237–266.

Fuster, Andreas, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar, "Expectations with Endogenous Information Acquisition: An Experimental Investigation," *The Review of Economics and Statistics*, September 2022, *104* (5), 1059–1078.

Haaland, Ingar and Christopher Roth, "Beliefs about Racial Discrimination and Support for Pro-Black Policies," *The Review of Economics and Statistics*, January 2023, *105* (1), 40–53.

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

Hvidberg, Kristoffer B, Claus T Kreiner, and Stefanie Stantcheva, "Social Positions and Fairness Views on Inequality," *The Review of Economic Studies*, November 2023, *90* (6), 3083–3118.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer, "Worker Beliefs About Outside Options*," *The Quarterly Journal of Economics*, January 2024, p. qjae001.

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.

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, "Imperfect Competition in the Labor Market," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 991–1041.

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. Publisher: The University of Chicago Press.

McFadden, Daniel and Kenneth Train, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 2000, *15* (5), 447–470. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1099-1255%28200009/10%2915%3A5%3C447%3A%3AAID-JAE570%3E3.0.CO%3B2-1.

Mueller, Andreas I, Damian Osterwalder, Josef Zweimüller, and Andreas Kettemann, "Vacancy Durations and Entry Wages: Evidence from Linked Vacancy–Employer–Employee Data," *The Review of Economic Studies*, May 2024, *91* (3), 1807–1841.

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.

Roth, Christopher and Johannes Wohlfart, "How Do Expectations about the Macroeconomy Affect Personal Expectations and Behavior?," *The Review of Economics and Statistics*, October 2020, *102* (4), 731–748.

_ , Sonja Settele, and Johannes Wohlfart, "Beliefs about public debt and the demand for government spending," *Journal of Econometrics*, November 2022, *231* (1), 165–187.

Spinnewijn, Johannes, "UNEMPLOYED BUT OPTIMISTIC: OPTIMAL INSURANCE DESIGN WITH BIASED BELIEFS: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, February 2015, *13* (1), 130–167.

Staiger, Douglas O., Joanne Spetz, and Ciaran S. Phibbs, "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment," *Journal of Labor Economics*, 2010, *28* (2), 211–236.

Train, Kenneth, *Discrete choice methods with simulation*, 2nd ed ed., Cambridge ; New York: Cambridge University Press, 2009. OCLC: ocn349248337.

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

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

Appendices

A Additional information about the survey

A.1 Survey structure

This section describes the different sections of the survey. Background information and beliefs about general averages are asked before F, the treatment section. After the treatment section, the respondents are asked to imagine representative jobs from each job type and asked what they believe they themselves would experience in each job.

A. **Welcome:**

The survey begins with an introduction thanking participants for their involvement in the study on job search and career choices. As a token of appreciation, participants are informed about the chance to win one of twenty gift cards, each worth 1,000 DKK. Half of these gift cards will be distributed randomly among those who complete the entire questionnaire. The other half will be awarded to participants who perform particularly well in specific tasks throughout the survey, to incentives focus on these questions.

B-C. **Demography and Background and Labor Information:**

In this section, participants are asked about their educational labor market status. They are asked about their highest completed education, which determines their educational group and the job type examples they are shown. If they are still studying their educational group is their current study. They choose their education through a detailed drop-down, first choosing a level, then a broad topic, and then a specific education. The full list of educations is shown in Appendix G. They are also asked whether they are currently employed and whether they are searching for a job.

D. **Perception of Job Types:**

Given their stated education, we can now provide the three relevant job types described in Section 2.2.1. They are introduced in this section, and the respondents are asked if they feel they understand the job types, and whether they know persons who work in them.

E. **Priors:**

These are the incentivized questions in which the respondents answer what they think the average starting wage, wage growth, and likelihood of getting a job in each job type is.

The question ask specifically about the monthly gross salary gross salary (i.e. before tax and including contribution to pension savings) of people who graduated in

2010-2018 and began working full-time in each job type after. Their answer will be referenced as respondent i's perceived average wage for job j, $\widetilde{W}^a_{i,j}$.

Because we are primarily interested in the relative beliefs, the respondents are given the anchor that the overall average in this period was 31,000 DKK. It is also stated that full-time employment refers to contracts of employment of at least 37 hours of work a week.

F. **Treatment:**

Half of the respondents with eligible educations are randomly chosen for treatment. These respondents are shown the actual average monthly wage for the three jobs in comparison with their stated answers. The control group is reminded of their previous answers in a scatter plot with three jobs on the x-axis and the stated average wage on the y-axis. The treatment group are shown the same plot but with bars indicating the average calculated from the register data. Examples of these are shown in the appendix, Figures A2 and A3.

The actual average calculated with the register data will be denoted $W^a_{i,j}$. It was calculated as the average monthly wage of persons who graduated in 2010-2018, and worked full-time in the job type within one year of graduation.

G. **Behavior and Posteriors:**

In this section, participants are asked to imagine three different jobs, one from each job type, that are representative of what they think a job of that job type would offer them. They are then asked specific questions on what they think about the job and what they would expect to experience in the job if they were hired.

They are asked to imagine a hypothetical scenario, where they the following day will find vacancy postings for all three jobs. The are asked to imagine that they only have time to apply for one of the jobs and then answer the likelihood they would apply to each job, including the likelihood that they would not apply to any of the jobs. This variable will be referred to as $\kappa_{i,j}$. The answers were recorded using three sliders that were adjusted to make sure the probabilities always summed to 100% by automatically adjusting the other sliders. They are all bounded between 1 and 97%. This likelihood will be referred to as $\kappa_{i,j}$ for person i in job j, with $j = 0$ being the likelihood of not applying for any of the three jobs.

They are also asked to imagine the monthly gross wage they think they would make themselves if they worked at the job. This variable will be referred to as $\widetilde{W}_{i,j}$

H. **Additional Background Information:**

This section gathers more personal information about the participants.

I. **Reminder:**

In the concluding section, participants in the treatment group are reminded of the actual wage data they were shown earlier, with the same graph being presented again. They are also given the opportunity to provide any additional thoughts or feedback about the survey or job search in general. The survey concludes with a thank you message for their participation.

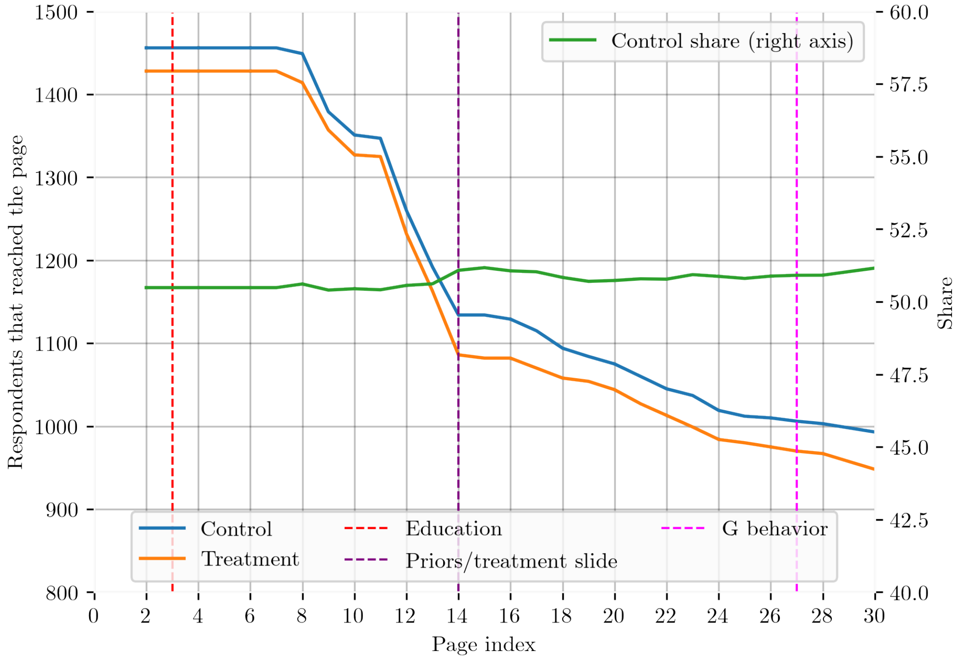
A.2 Survey descriptive statistics

Table A1: Descriptive statistics on survey waves

	UCPH only	UCPH and STAR	STAR only wave 1	STAR only wave 2	STAR only wave 3	All	Eligible
Total invited	2,609	366	16,471	21,773	5,012	46,231	
Total answers	482	72	2,193	3,598	818	7,163	2,928
Completed answers	296	45	1,308	2,226	485	4,360	1,941
Has custom jobtypes	243	33	773	1,459	253	2,761	1,941
Eligible for treatment	173	30	549	1,008	181	1,941	1,941
Age	27.89	27.73	29.26	28.56	27.44	28.59	28.10
Female	0.65	0.71	0.65	0.63	0.58	0.63	0.63
Higher education	1.00	1.00	0.79	0.84	0.75	0.83	1.00
Masters	1.00	1.00	0.41	0.46	0.39	0.48	0.64
Graduated at most 2 years ago	0.61	0.87	0.52	0.67	0.46	0.60	0.70
Currently studying	0.39	0.11	0.05	0.04	0.07	0.07	0.07
Expect to graduate in 2 months	0.24	0.11	0.02	0.01	0.00	0.03	0.04
Expect to graduate in 1 year	0.38	0.11	0.03	0.02	0.02	0.05	0.05
Employed	0.53	0.09	0.16	0.18	0.18	0.20	0.19
Active job searcher	0.36	0.62	0.71	0.68	0.75	0.68	0.65
Active job searcher and neither employed or studying	0.15	0.58	0.61	0.59	0.63	0.57	0.56
Studying and employed	0.25	0.07	0.01	0.01	0.02	0.03	0.03
Not studying and not employed	0.33	0.87	0.81	0.79	0.76	0.76	0.77
Studying and active job searcher	0.13	0.04	0.03	0.02	0.05	0.03	0.03

Notes: UCPH is the population made available by the University of Copenhagen of persons close to finishing their masters. STAR is the population of persons that have signed up for unemployment benefits in the summer of 2023. The eligible column only includes respondents who had an education that was eligible for treatment, meaning we had enough historical observations from the education to provide an informative treatment about the transition, as described in Section 2.2.1.

Figure A1: Survey attrition among eligible for treatment



Note: The figure shows the share of respondents that stay in the survey for each page they answer, for the control and treatment. The green line shows the share the control group makes of the sample. The sample is the survey population from educations that were eligible for treatment.

Table A2: Balance table

	(1) Fully completed answers	(2) Not-completed answers	(3) Eligible for Treatment	(4) Not eligible	(5) Control group	(6) Treatment group	(7) p-value (1) = (2)	(8) p-value (3) = (4)	(9) p-value (5) = (6)
Age	28.59	28.44	28.10	28.99	28.17	28.02	0.156	0.000	0.385
Female	0.63	0.70	0.63	0.63	0.61	0.64	0.000	0.741	0.321
Higher education	0.83	0.72	1.00	0.69	1.00	1.00	0.000	0.000	
Masters	0.48	0.37	0.64	0.35	0.63	0.65	0.000	0.000	0.403
Graduated at most 1 year ago	0.54	0.52	0.64	0.45	0.64	0.65	0.188	0.000	0.528
Graduated at most 2 years ago	0.60	0.59	0.70	0.52	0.68	0.72	0.631	0.000	0.067
Currently studying	0.07	0.09	0.07	0.07	0.07	0.07	0.007	0.874	0.916
Expect to graduate in 2 months	0.03	0.03	0.04	0.02	0.04	0.04	0.537	0.001	0.857
Expect to graduate in 1 year	0.05	0.04	0.05	0.04	0.05	0.06	0.370	0.016	0.746
Employed	0.20	0.23	0.19	0.20	0.21	0.17	0.000	0.379	0.017
Studying and is employed	0.03	0.03	0.04	0.03	0.04	0.03	0.402	0.121	0.857
Neither studying or employed	0.76	0.72	0.77	0.75	0.75	0.79	0.000	0.177	0.037
Active jobbsearcher	0.68	0.65	0.65	0.70	0.63	0.67	0.041	0.002	0.113
Willingness to take a risk (1-10)	6.34	.	6.20	6.44	6.27	6.14		0.000	0.162
Job offers last 2 years	3.16	.	2.27	3.87	2.41	2.13		0.000	0.151
Wanted children	2.02	.	2.02	2.03	2.04	2.00		0.755	0.333
Likelihood of child next 4 years	45.50	.	46.56	44.66	47.48	45.58		0.108	0.273
Sum of abs. relative diff. btw. earnings belief and actual	0.59	0.74	0.58	0.63	0.55	0.60	0.067	0.448	0.598
Sum of abs. relative diff. btw. earnings growth belief and actual	0.61	0.54	0.58	0.68	0.55	0.61	0.777	0.605	0.428
Sum of abs. diff. btw. application success belief and actual	382.27	461.92	365.85	416.73	370.61	360.87	0.002	0.025	0.697
Observations	4360	2803	1941	2419	993	948			

Notes: Balance table between different subgroups of the survey population. Along with statistics. Column (9) shows the p-values of a t-test, testing the difference between the control and treatment groups. This table only includes respondents who finished the survey.

Table A3: Balance table - including non-finished observations

	(1) Fully completed answers	(2) Not-completed answers	(3) Eligible for Treatment	(4) Not eligible	(5) Control group	(6) Treatment group	(7) p-value (1) = (2)	(8) p-value (3) = (4)	(9) p-value (5) = (6)
Age	28.59	28.44	28.12	28.82	28.08	28.12	0.156	0.000	0.803
Female	0.63	0.70	0.66	0.65	0.66	0.67	0.000	0.201	0.537
Higher education	0.83	0.72	1.00	0.64	1.00	1.00	0.000	0.000	
Masters	0.48	0.37	0.62	0.31	0.61	0.63	0.000	0.000	0.458
Graduated at most 1 year ago	0.54	0.52	0.64	0.45	0.64	0.64	0.188	0.000	0.825
Graduated at most 2 years ago	0.60	0.59	0.70	0.52	0.69	0.71	0.631	0.000	0.397
Currently studying	0.07	0.09	0.07	0.08	0.07	0.07	0.007	0.015	0.889
Expect to graduate in 2 months	0.03	0.03	0.03	0.02	0.04	0.03	0.537	0.001	0.523
Expect to graduate in 1 year	0.05	0.04	0.05	0.04	0.05	0.05	0.370	0.026	0.784
Employed	0.20	0.23	0.20	0.22	0.22	0.19	0.000	0.163	0.032
Studying and is employed	0.03	0.03	0.03	0.03	0.03	0.03	0.402	0.311	0.841
Neither studying or employed	0.76	0.72	0.76	0.74	0.74	0.78	0.000	0.019	0.043
Active jobbsearcher	0.68	0.65	0.65	0.68	0.64	0.66	0.041	0.005	0.210
Willingness to take a risk (1-10)	6.34	.	6.20	6.44	6.27	6.14		0.000	0.162
Job offers last 2 years	3.16	.	2.27	3.87	2.41	2.13		0.000	0.151
Wanted children	2.02	.	2.02	2.03	2.04	2.00		0.755	0.333
Likelihood of child next 4 years	45.50	.	46.56	44.66	47.48	45.58		0.108	0.273
Sum of abs. relative diff. btw. earnings belief and actual	0.59	0.74	0.59	0.70	0.57	0.61	0.067	0.140	0.561
Sum of abs. relative diff. btw. earnings growth belief and actual	0.61	0.54	0.58	0.66	0.55	0.60	0.777	0.631	0.504
Sum of abs. diff. btw. application success belief and actual	382.27	461.92	384.75	421.24	384.83	384.68	0.002	0.082	0.995
Observations	1.00	0.00	0.66	0.57	0.68	0.66		0.000	0.299
thisstat23	4360	2803	2928	4235	1456	1428			

Notes: Balance table between different subgroups of the survey population. Along with statistics. Column (9) shows the p-values of a t-test, testing the difference between the control and treatment groups. This table includes recorded respondents' answers, also for respondents who did not finish the survey.

Table A4: Validation question

	Has custom jobtypes			Eligible for treatment		
	Job 1	Job 2	Job 3	Job 1	Job 2	Job 3
Eligible answers	755	817	680	494	558	442
Share of correct answers	0.81	0.75	0.80	0.80	0.72	0.79

Notes: Shows the results of the validation question. Some respondents were asked to place an example of one of the 3 job types in the correct corresponding job type. This table shows the share that got it right. The jobs are ordered by how common they have been for education in the past, such that job 1 is the most common job type to go into for persons with the respondents' education.

Table A5: Validation question, by education

	Has custom jobtypes			Eligible for treatment		
	Job 1	Job 2	Job 3	Job 1	Job 2	Job 3
Total answers	632	688	553	451	510	395
Unique educations	49	48	47	34	33	33
Average correct answer weighted by educations	0.81	0.76	0.78	0.79	0.73	0.72
std	0.23	0.23	0.25	0.25	0.25	0.27
min	0.00	0.00	0.00	0.00	0.00	0.00
10%	0.50	0.50	0.42	0.51	0.44	0.40
25%	0.71	0.59	0.68	0.69	0.55	0.60
50%	0.87	0.83	0.83	0.85	0.81	0.78
75%	1.00	0.93	1.00	1.00	0.92	0.93
max	1.00	1.00	1.00	1.00	1.00	1.00

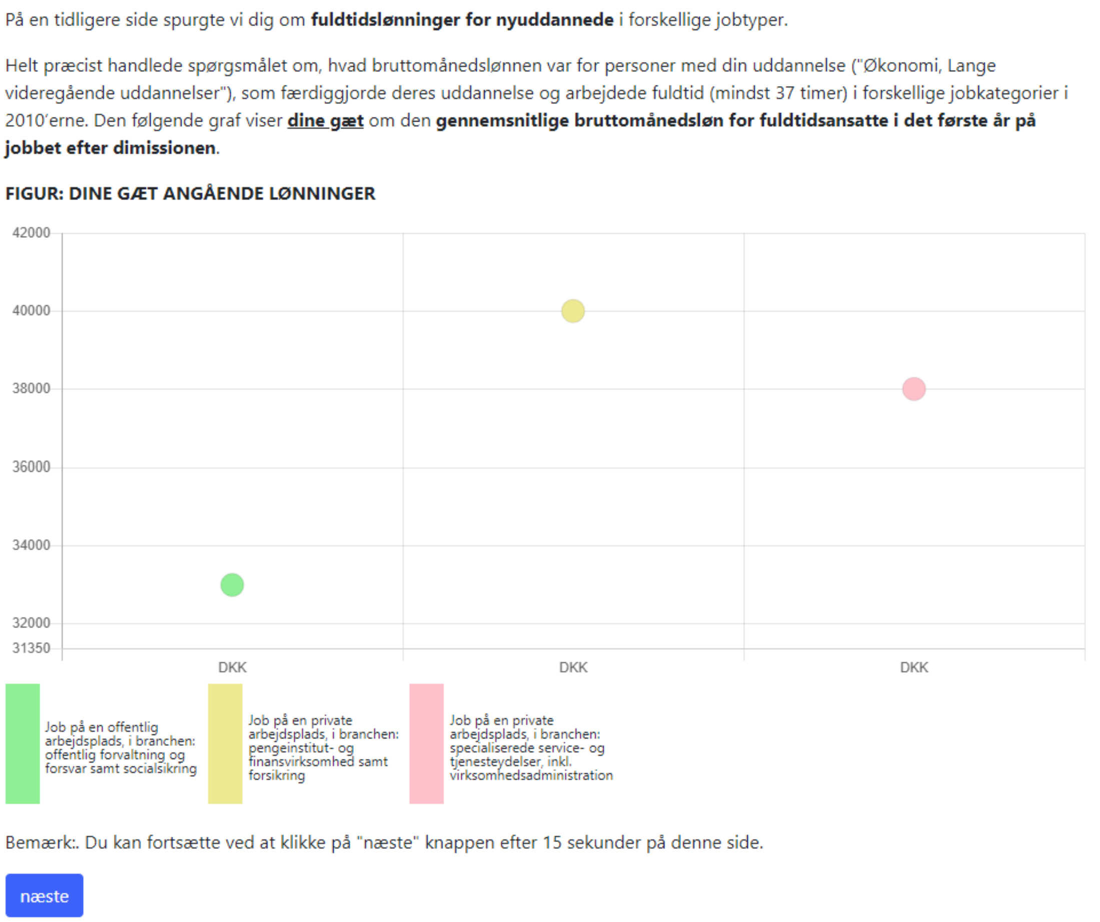
Notes: Shows the results of the validation question. Some respondents were asked to place an example of one of the 3 job types in the correct corresponding job type. This table shows the share that got it right, weighted by educations. The jobs are ordered by how common they have been for education in the past, such that job 1 is the most common job type to go into for persons with the respondents' education.

Table A6: Examples of some job types associated with two different educations.

Degree	**Job types**	**Categories**
Ms. Economics	- 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
Ms. Physics	- 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

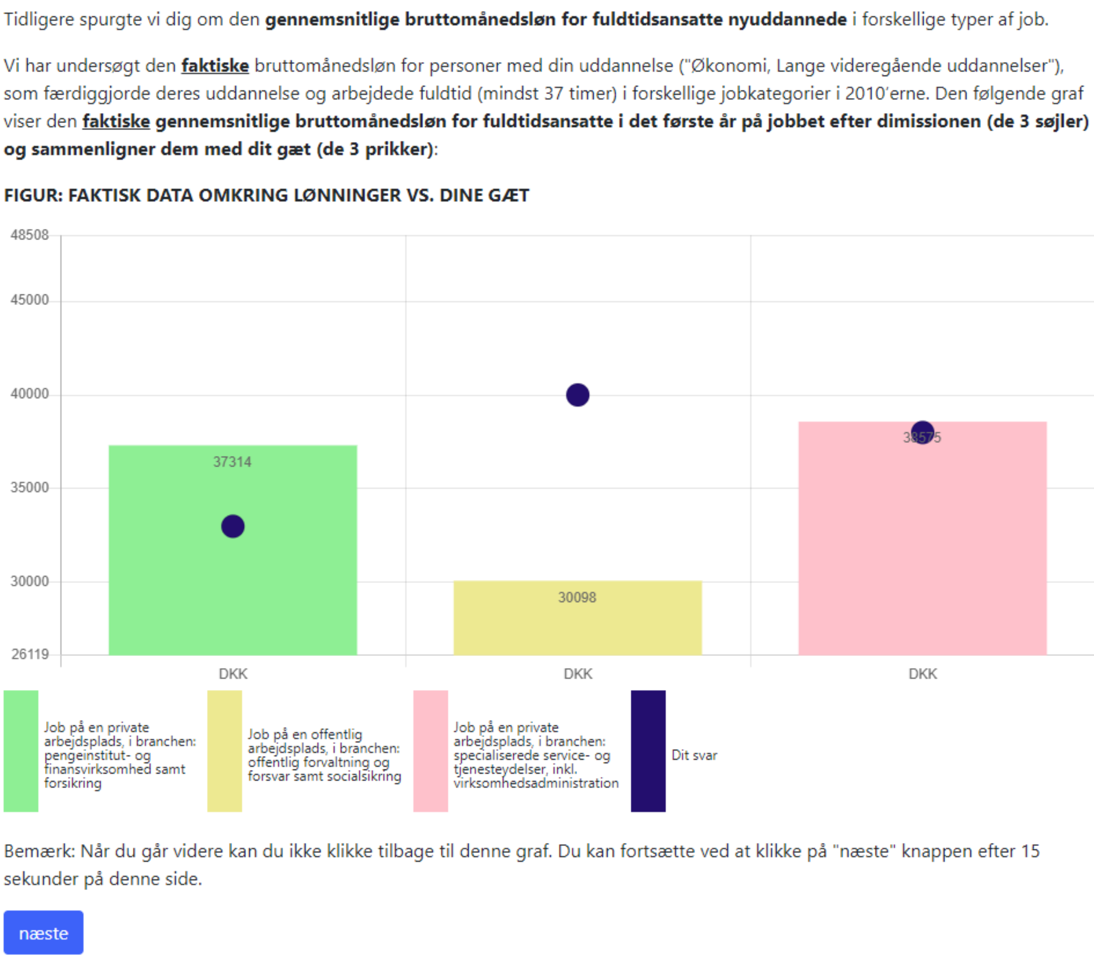
Notes: Examples of the job types, that were showed to the respondents, for two different educations. They are translated from Danish.

Figure A2: Example of treatment slide shown to the control group in the survey



Note: This figure shows an example of the treatment slide shown to the control group, reminding them of their stated answers about the average wage in each job.

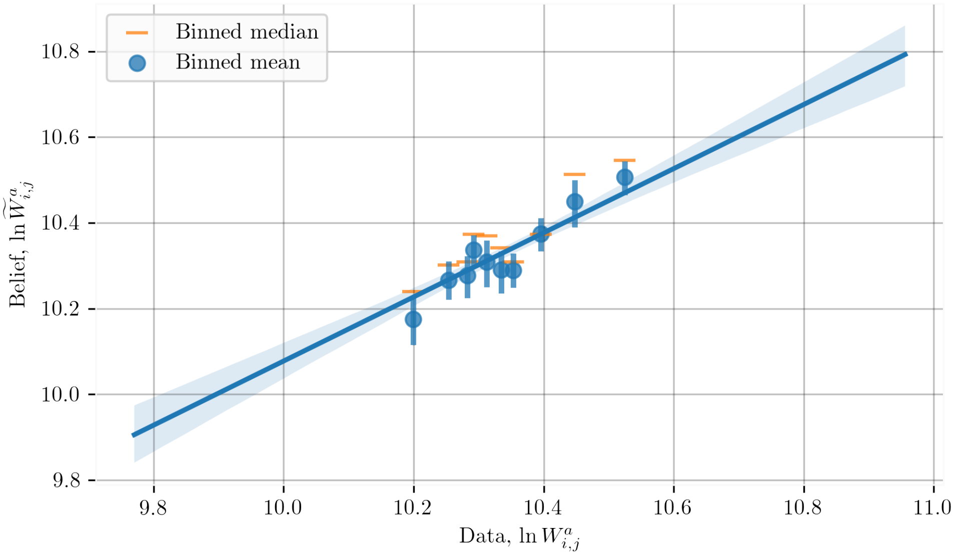
Figure A3: Example of treatment slide shown to the treatment group in the survey



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

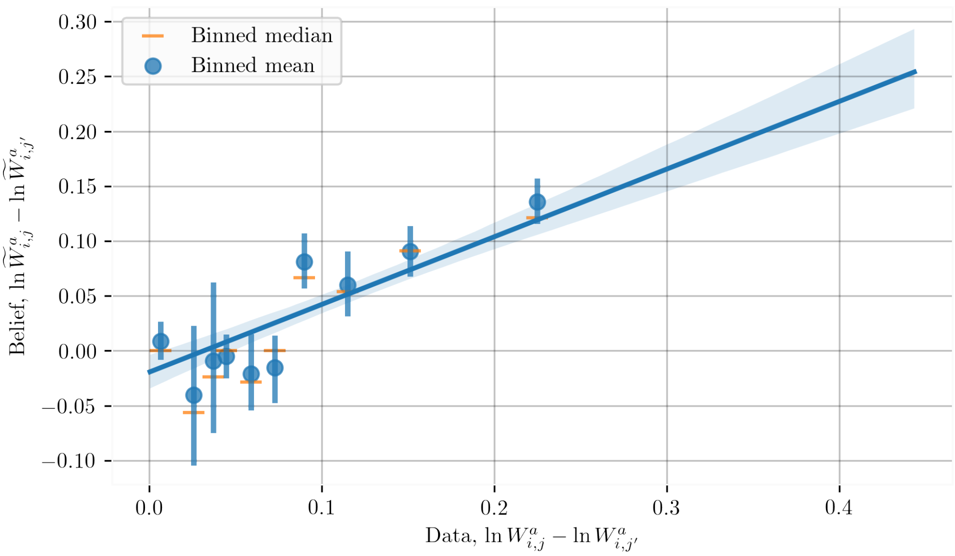
B Additional descriptives on misperceptions

Figure B1: (Log) Starting average wages, respondent beliefs vs. ground truth, binscatter



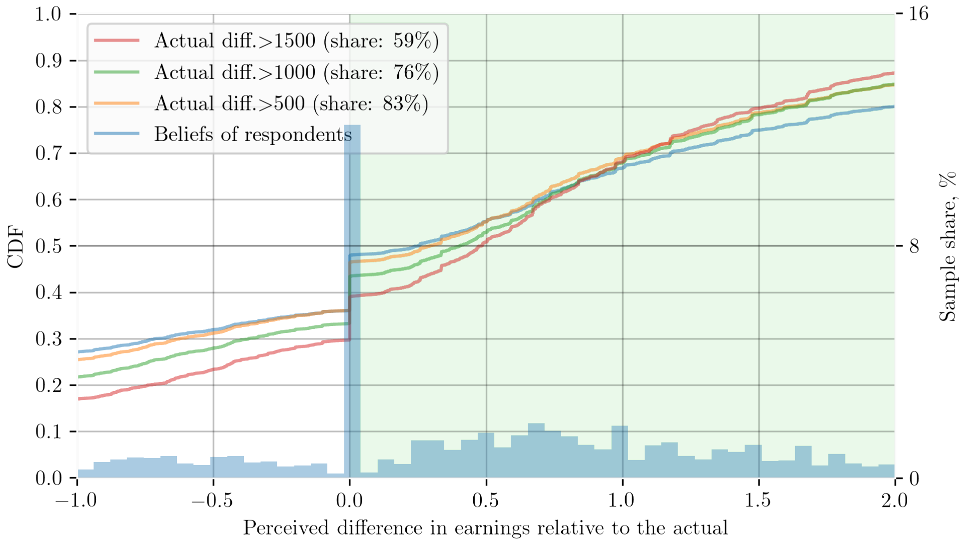
Note: This figures plots the log of the perceived average starting wage against the log of the average starting wage in the register data. The sample is the survey population from educations that were eligible for treatment. folder

Figure B2: Between jobs average earnings log ratio, respondent beliefs vs. ground truth, binscatter



Note: This figures plots the log gap between the perceived average starting wages of job types, against the log gap between the average starting wage in the register data of job types. The sample is the survey population from educations that were eligible for treatment. folder

Figure B3: Perceived difference in percent of the actual, conditioning on absolute actual difference



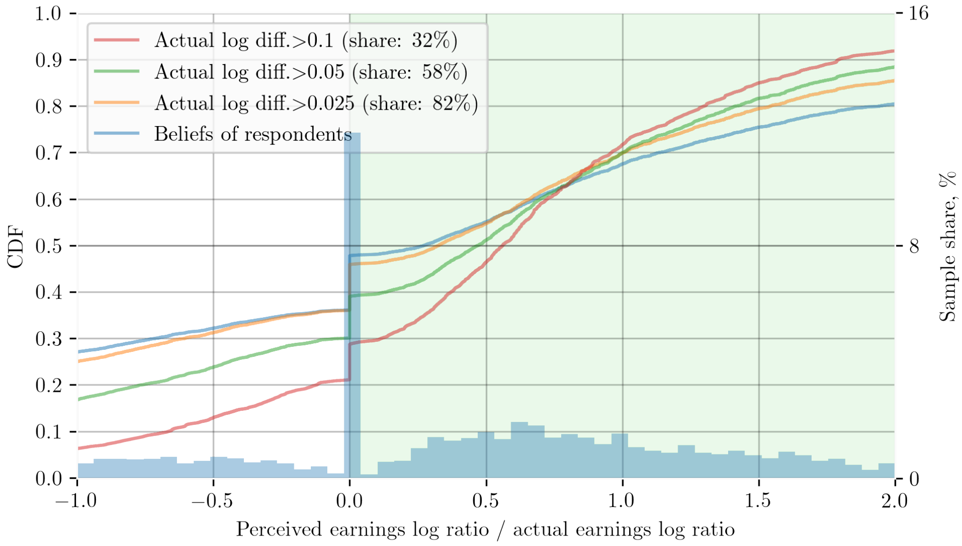
Note: This figure plots empirical cumulative distribution function (left axis) and the histogram (right axis) of the Perceived Difference in Percent of the Actual: $\text{PDPA}_{i,j} = \frac{\widetilde{W}^a_{i,j} - \widetilde{W}^a_{i,j'}}{W^a_{i,j} - W^a_{i,j'}}$. The green shaded area highlights respondents who correctly ranked the earnings differences between the two jobs. The actual diff. is the difference between the average earnings in the jobs in the register data, $\left(W^a_{i,j} - W^a_{i,j}\right)$, measured in Danish Kroner (DKK). How big a share of the initial sample that is still included after the condition is stated in the label. The sample is the survey population from educations that were eligible for treatment.

Table B1: Share getting the earnings ranking of the job type right

		All	Custom jobtypes	Eligible	Example Q Incorrect	 Correct	Min understanding ≥ 4	 ≥ 5
Earnings	Entire ranking	17.95	18.63	20.51	15.69	21.17	19.61	21.09
	Correct top job type	35.91	37.33	39.40	27.45	38.75	37.86	40.53
	Correct bottom job type	36.56	37.15	37.83	33.55	42.40	38.03	38.06
	N	3,498	3,311	2,482	459	1,644	1,775	607
Application success	Entire ranking	15.40	14.38	15.53	13.92	14.91	17.53	20.71
	Correct top job type	33.02	29.98	32.48	29.93	30.17	35.05	39.29
	Correct bottom job type	29.06	28.72	29.31	29.00	29.27	30.45	32.94
	N	2,526	2,378	1,829	431	1,435	1,261	425
5 year earnings	Entire ranking	23.60	22.45	23.54	19.11	25.45	24.26	24.79
	Correct top job type	41.96	40.91	41.29	36.13	42.72	42.22	42.09
	Correct bottom job type	48.93	47.60	49.03	41.62	51.65	50.54	51.71
	N	2,581	2,437	2,209	382	1,332	1,381	468

Notes: Share answering the relative ranking between jobs correctly compared to actual register data. Respondents are given the benefit of the doubt and seems as having answered correctly, if they answered a difference of 0.

Figure B4: Perceived log difference in percent of the actual log difference



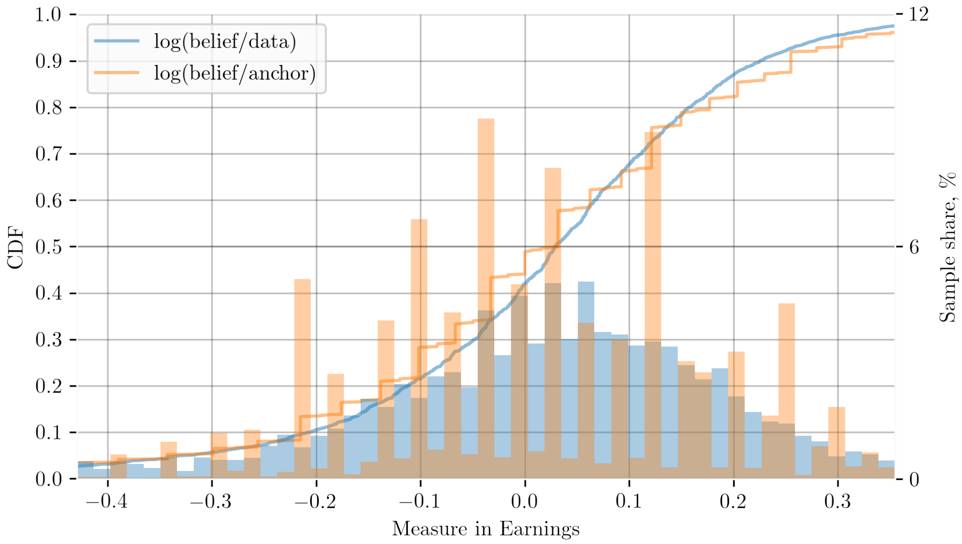
Note: The figure plots $\frac{\ln \widetilde{W}^a_{i,j} - \ln \widetilde{W}^a_{i,j'}}{\ln W^a_{i,j} - \ln W^a_{i,j'}}$ as opposed to $\frac{\widetilde{W}^a_{i,j} - \widetilde{W}^a_{i,j'}}{W^a_{i,j} - W^a_{i,j'}}$. The green shaded area highlights respondents who correctly ranked the earnings differences between the two jobs. The actual log diff. is the log difference between the average earnings in the jobs in the register data, $\left(\ln W^a_{i,j} - \ln W^a_{i,j}\right)$. How big a share of the initial sample that is still included after the condition is stated in the label. The sample is the survey population from educations that were eligible for treatment.

Table B2: Share getting the earnings ranking of the job type right, lenient

		All	Custom jobtypes	Eligible	Example Q Incorrect	Example Q Correct	Min understanding ≥ 4	Min understanding ≥ 5
Earnings	Entire ranking	33.36	34.07	36.78	30.28	37.04	35.49	37.40
	Correct top job type	48.91	50.32	52.74	41.39	51.40	50.70	53.21
	Correct bottom job type	49.43	49.74	50.48	47.28	53.35	51.15	51.57
	N	3,498	3,311	2,482	459	1,644	1,775	607
Application success	Entire ranking	30.29	28.09	29.91	29.70	26.27	32.99	39.53
	Correct top job type	45.88	42.94	45.82	45.48	40.28	48.37	55.06
	Correct bottom job type	43.19	41.76	43.03	45.01	40.00	45.44	50.82
	N	2,526	2,378	1,829	431	1,435	1,261	425
5 year earnings	Entire ranking	40.49	38.57	39.66	40.58	41.59	42.00	44.44
	Correct top job type	56.22	54.78	55.36	54.19	55.71	57.86	58.97
	Correct bottom job type	62.57	61.10	62.70	58.12	63.14	64.59	66.88
	N	2,581	2,437	2,209	382	1,332	1,381	468

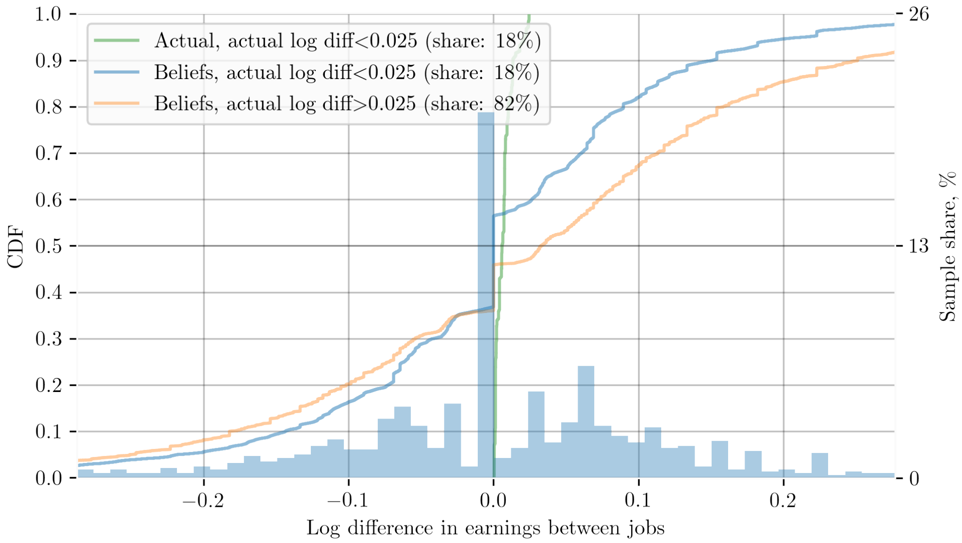
Notes: Share answering the relative ranking between jobs correctly compared to actual register data. If a respondent has stated that two job types pay the exact same, this is counted as getting the ranking right.

Figure B5: Earnings belief and data relative to anchor



Note: This figure plots the log difference between earnings belief and the common anchor shown to all respondents. And the log difference between earnings beliefs and the actual earnings in the register data. The sample is the survey population from educations that were eligible for treatment.

Figure B6: Log difference between jobs



Note: This figure plots the believed and actual log difference between jobs, i.e. $\left(\ln \widetilde{W}^a_{i,j} - \ln \widetilde{W}^a_{i,j}\right)$ and $\left(\ln W^a_{i,j} - \ln W^a_{i,j}\right)$, split on whether the actual log difference is below or above 0.025. The sample is the survey population from educations that were eligible for treatment.

C Auxiliary regressions

Table C1: Effect of log perceived wage on other job amenities, not controlling for $\tilde{w}^a_{i,j}$

	(1) $\tilde{w}_{i,j,t+5}$	(2) Wage growth	(3) ln Hours	(4) Appl likelihood	(5) Colleagues	(6) Performance
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	0.528***	0.011**	0.134***	0.653	0.718***	0.594
	(0.056)	(0.005)	(0.028)	(4.594)	(0.217)	(0.393)
$T_i\times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	-0.219***	0.014***	-0.020	-1.186	-0.047	-0.227
	(0.046)	(0.005)	(0.025)	(4.345)	(0.249)	(0.368)
Person FE	✓	✓	✓	✓	✓	✓
N	5659	5659	5532	5706	5706	5706

This table shows first stage regressions, replacing the outcome variable with perceived job amentites. $\tilde{w}_{i,j,t+5}$ is the log of the perceived earnings in 5 years after starting to work in the job type. *Yearly wage growth*, uses perceived wages in 5 years and perceived initial wages to calculate an equivalent yearly wage growth. ln Hours is the log of the stated expected monthly hours worked, observations are excluded if wages in 5 years or hours were stated to be zero. *Appl likelihood* is the stated likelihood that, given that they had applied to the job, how likely it would be that they would receive a job offer, graded from 0-100. *Colleagues* is how they think they would get on with their colleagues in the job type, scalled from 1-6. *Performance* is how well they think they would perform in the job type, scalled from 1-6. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table C2: Effect of log perceived wage on other job amenities, alternative amenities

	(1) 5-year earnings change	(2) 5-year earnings growth	(3) ln Hours	(4) Hours	(5) Hours$\neq 0$
ln perceived avg. wage, $\tilde{w}^a_{i,j}$	17969.737***	0.295***	0.323***	13.631***	14.183***
	(2946.243)	(0.071)	(0.077)	(3.698)	(3.190)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	-8153.007***	-0.155**	-0.109	-0.805	-5.015*
	(2592.670)	(0.063)	(0.067)	(4.072)	(2.788)
$T_i\times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	440.123	0.086**	-0.027	-2.474*	-1.296
	(1058.151)	(0.036)	(0.029)	(1.342)	(1.264)
Person FE	✓	✓	✓	✓	✓
N	5659	5659	5532	5706	5532

This table shows first stage regressions, replacing the outcome variable with perceived job amentites. *5-year earnings change* is the difference between perceived wages in 5 years and perceived initial wages. *5-year earnings growth* is the percentage growth. *Hours*$\neq 0$ is the perceived monthly hours the respondent would work in the firm, excluding the respondents that answered 0. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table C3: Correlation between log perceived starting wage, and other perceived amenities

	(1) $\tilde{w}_{i,j,t+5}$	(2) Wage growth	(3) ln Hours	(4) Appl likelihood	(5) Colleagues	(6) Performance
Panel A: Full eligible sample						
ln perceived wage, $\tilde{w}_{i,j}$	0.818***	-0.021**	0.227***	-0.857	0.437**	0.737**
	(0.033)	(0.008)	(0.037)	(4.371)	(0.173)	(0.340)
N	5659	5659	5532	5706	5706	5706
Panel B: Control sample						
ln perceived wage, $\tilde{w}_{i,j}$	0.837***	-0.013*	0.250***	-2.586	0.558**	0.761*
	(0.037)	(0.008)	(0.035)	(4.513)	(0.249)	(0.405)
N	2897	2897	2823	2928	2928	2928
Panel C: Treatment sample						
ln perceived wage, $\tilde{w}_{i,j}$	0.791***	-0.032**	0.194***	1.556	0.269	0.703**
	(0.058)	(0.013)	(0.048)	(5.451)	(0.195)	(0.353)
N	2762	2762	2709	2778	2778	2778
Person FE	✓	✓	✓	✓	✓	✓

Notes: This table shows the person-adjusted correlation between amenities and log of perceived wages in the job type. Panel A uses the full eligible sample. Panel B uses the control group. Panel C uses the treatment group. *Yearly wage growth*, uses perceived wages in 5 years and perceived initial wages to calculate an equivalent yearly wage growth. ln Hours is the log of the stated expected monthly hours worked, observations are excluded if wages in 5 years or hours were stated to be zero. *Appl likelihood* is the stated likelihood that, given that they had applied to the job, how likely it would be that they would receive a job offer, graded from 0-100. *Colleagues* is how they think they would get on with their colleagues in the job type, scalled from 1-6. *Performance* is how well they think they would perform in the job type, scalled from 1-6. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

C.1 Not controlling for $\ln \tilde{p}_{i,j}$

Table C4: The effect of log perceived average wage on log perceived wage

	ln perceived wage, $\tilde{w}_{i,j}$				
	(1)	(2)	(3)	(4)	(5)
Treatment, T_i	-0.008	-0.015**			
	(0.006)	(0.007)			
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	0.615***	0.553***	0.499***	-0.131**	-0.140**
	(0.050)	(0.031)	(0.042)	(0.056)	(0.059)
$T_i\times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	-0.360***	-0.337***	-0.320***	-0.336***	-0.316***
	(0.046)	(0.036)	(0.044)	(0.042)	(0.043)
ln perceived avg. wage, $\tilde{w}^a_{i,j}$				0.843***	0.855***
				(0.045)	(0.052)
$T_i\times$ ln perceived avg. wage, $T_i\tilde{w}^a_{i,j}$					-0.026
					(0.042)
Education FE		✓			
Person FE			✓	✓	✓
Constant	✓				
Educations	83	83	83	83	83
Persons	1902	1902	1902	1902	1902
N	5706	5706	5706	5706	5706

This table shows the first stages regressions, estimating the degree of updating in the log perceived wage of the job, given a signal about the average wage. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table C5: The effect of log perceived wage on log relative application ratio, IV

	Log likelihood of applying relative to not applying, $k_{i,j} \equiv \ln \frac{\kappa_{i,j}}{\kappa_{i,0}}$			
	(1)	(2)	(3)	(4)
ln perceived wage, $\tilde{w}_{i,j}$	1.799*	4.708***	4.741***	4.638***
	(0.988)	(1.353)	(1.298)	(1.272)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	1.058*	2.206***	1.758	1.729
	(0.595)	(0.757)	(1.091)	(1.089)
ln perceived avg. wage, $\tilde{w}^a_{i,j}$			0.577	0.664
			(1.435)	(1.414)
Person FE		✓	✓	✓
Constant	✓			
Instruments:				
$T_i(\tilde{w}^a_{i,j} - w^a_j)$	✓	✓	✓	✓
$T_i \tilde{w}^a_{i,j}$				✓
T_i	✓			
First stage F(K-P)	33.18	52.80	64.61	34.91
Average elasticity $\widehat{\mathcal{E}}^a$	1.256	3.285	3.308	3.237
Educations	83	83	83	83
Persons	1902	1902	1902	1902
N	5706	5706	5706	5706

This table shows second stage estimates instrumenting log of the perceived wage, $\ln \tilde{w}_{i,j}$, with the information treatment. The outcome variable is the log likelihood of applying for a job, relative to none. $\widehat{\mathcal{E}}^a$ is the sample-average elasticity of the likelihood of applying for a job with respect to the perceived wage suggested by the estimate. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table C6: OLS regression

	Log likelihood of applying relative to not applying, $k_{i,j} \equiv \ln \frac{\kappa_{i,j}}{\kappa_{i,0}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
ln perceived wage, $\tilde{w}_{i,j}$	1.135***	4.281***	3.774***	3.774***	5.132***	4.507***
	(0.195)	(0.361)	(0.344)	(0.386)	(0.376)	(0.390)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	1.355***	2.356***	1.485	1.093		
	(0.471)	(0.596)	(1.089)	(0.868)		
ln perceived avg. wage, $\tilde{w}^a_{i,j}$			1.388	0.543		
			(0.944)	(0.833)		
Yearly wage growth				11.890***		13.910***
				(1.407)		(1.559)
ln Hours				1.571***		1.641***
				(0.413)		(0.414)
Colleagues				0.308***		0.320***
				(0.033)		(0.034)
Performance				0.476***		0.472***
				(0.033)		(0.032)
Persons FE		✓	✓	✓	✓	✓
Constant	✓					
N	5706	5706	5706	5514	5706	5514
Average elasticity $\widehat{\mathcal{E}}^a$	0.792	2.987	2.634	2.623	3.581	3.133

This table shows OLS estimates. The outcome variable is the log likelihood of applying for a job, relative to none. $\widehat{\mathcal{E}}^a$ is the sample-average elasticity of the likelihood of applying for a job with respect to the perceived wage suggested by the estimate. *Yearly wage growth*, uses perceived wages in 5 years and perceived initial wages to calculate an equivalent yearly wage growth. ln Hours is the log of the stated expected monthly hours worked, observations are excluded if wages in 5 years or hours were stated to be zero. *Colleagues* is how they think they would get on with their colleagues in the job type, scalled from 1-6. *Performance* is how well they think they would perform in the job type, scalled from 1-6. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

C.2 Spillovers to other jobs

Table C7: 1 stage controlling for wage treatment in other jobs

	ln perceived wage, $\tilde{w}_{i,j}$			
	(1)	(2)	(3)	(4)
Treatment, T_i	-0.015**	-0.015**		
	(0.007)	(0.007)		
ln perceived avg. wage, $\tilde{w}^a_{i,j}$	0.843***	0.843***	0.843***	0.843***
	(0.045)	(0.045)	(0.045)	(0.045)
Perception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$	-0.173***	-0.173***	-0.131**	-0.130**
	(0.055)	(0.055)	(0.054)	(0.060)
$T_i\times$ perception gap, $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$	-0.339***	-0.339***	-0.339***	-0.333***
	(0.033)	(0.033)	(0.044)	(0.044)
$\tilde{w}^a_{i,j'} - w^a_{j'}$	-0.042*		-0.001	
	(0.022)		(0.022)	
$T_i(\tilde{w}^a_{i,j'} - w^a_{i,j'})$	-0.007		-0.006	
	(0.026)		(0.028)	
$\tilde{w}^a_{i,j''} - w^a_{j''}$	-0.041**			0.001
	(0.018)			(0.022)
$T_i(\tilde{w}^a_{i,j''} - w^a_{i,j''})$	-0.000			0.006
	(0.023)			(0.028)
$\frac{1}{2}\big((\tilde{w}^a_{i,j'} + w^a_{i,j'}) + (\tilde{w}^a_{i,j'} - w^a_{i,j'})\big)$		-0.083**		
		(0.033)		
$T_i\frac{1}{2}\big((\tilde{w}^a_{i,j'} - w^a_{i,j'}) + (\tilde{w}^a_{i,j'} - w^a_{i,j'})\big)$		-0.007		
		(0.040)		
Education FE	✓	✓		
Person FE			✓	✓
N	5706	5706	5706	5706

This table shows the effect of the information treatment on log of perceived wages in the job, and includes controls for earnings perceptions gap in other jobs types that where also treated for the persons, interacted with treatment. The sample is the survey population from educations that were eligible for treatment. Std. errors are clustered at education. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

Table C7 recreates the first stage regression of column (4) of Table 2, but includes controls for treatment interacted with the misperception gap in the two other jobs, to check whether there is a spillover of beliefs about wages between jobs. Columns (1)-(2) cannot include person-fixed effects, because of perfect multicollinearity. There does not seem to be spillovers across jobs, as the estimate associated with $T_i(\tilde{w}^a_{i,j} - w^a_{i,j})$ is not much different from the first stage without the controls, nor is there an effect from treatment interacted with the misperception on the other jobs.

D Beliefs formation

D.1 Bayesian updating with Gaussian prior and signal

Starting from the wage for person i in job type j given the mean wage in that job type $\bar{w}_{i,j}$:

$$w_{i,j} = a_i + \bar{w}_{i,j} + a_{i,j} \tag{D.1}$$

Let's first assume that a_i and $a_{i,j}$ are known and independent of $\bar{w}_{i,j}$. And then also that the prior is that $\bar{w}_{i,j}|\tilde{w}^a_{i,j} \sim \mathcal{N}(\tilde{w}^a_{i,j}, \frac{1}{\rho})$. When treated with information about the average they assume this is a noisy signal $w^a_{i,j} = \bar{w}_{i,j} + n_{i,j}$, with the noise being $n_{i,j} \sim \mathcal{N}(0, \frac{1}{\rho_n})$. Prior probability densities are then:

$$p_{\bar{w}_{i,j}}(x|\tilde{w}^a_{i,j}) = \frac{1}{\sqrt{2\pi\rho^{-1}}} \exp\left(-\frac{1}{2}\left(x - \tilde{w}^a_{i,j}\right)^2 \rho\right) \tag{D.2}$$

$$p_{w^a_{i,j}}(y|\bar{w}_{i,j}) = \frac{1}{\sqrt{2\pi\rho_n^{-1}}} \exp\left(-\frac{1}{2}(y - \bar{w}_{i,j})^2 \rho_n\right) \tag{D.3}$$

Bayes theorem allows us to write:

$$\begin{aligned}
p_{\bar{w}_{i,j}}(x|\tilde{w}^a_{i,j}, w^a_{i,j}) &= \frac{p_{w^a_{i,j}}(w^a_{i,j}|\tilde{w}^a_{i,j}, \bar{w}_{i,j} = x) p_{\bar{w}_{i,j}}(x|\tilde{w}^a_{i,j})}{p_{w^a_{i,j}}(w^a_{i,j}|\tilde{w}^a_{i,j})} && \text{(D.4)}\\
&\propto p_{w^a_{i,j}}(w^a_{i,j}|\tilde{w}^a_{i,j}, \bar{w}_{i,j} = x) p_{\bar{w}_{i,j}}(x|\tilde{w}^a_{i,j}) && \text{(D.5)}\\
&\propto \exp\left(-\frac{1}{2}\left(w^a_{i,j} - x\right)^2 \rho_n\right) \exp\left(-\frac{1}{2}\left(x - \tilde{w}^a_{i,j}\right)^2 \rho\right) && \text{(D.6)}\\
&= \exp\left(-\frac{1}{2}\left(\left(w^a_{i,j} - x\right)^2 \rho_n + \left(x - \tilde{w}^a_{i,j}\right)^2 \rho\right)\right) && \text{(D.7)}\\
&= \exp\left(-\frac{1}{2}\left(((w^a_{i,j})^2 + x^2 - 2xw^a_{i,j})\rho_n + \left(x^2 + (\tilde{w}^a_{i,j})^2 - 2x\tilde{w}^a_{i,j}\right)\rho\right)\right) && \text{(D.8)}\\
&= \exp\left(-\frac{1}{2}\left(x^2(\rho_n + \rho) - 2x(w^a_{i,j}\rho_n + \tilde{w}^a_{i,j}\rho) + (w^a_{i,j})^2\rho_n + (\tilde{w}^a_{i,j})^2\rho\right)\right) && \text{(D.9)}\\
&= \exp\left(-\frac{1}{2}\left(x^2(\rho_n + \rho) - 2x(\rho_n + \rho)(w^a_{i,j}\frac{\rho_n}{\rho_n + \rho} + \frac{\rho}{\rho_n + \rho}\tilde{w}^a_{i,j}) + (w^a_{i,j})^2\rho_n + (\tilde{w}^a_{i,j})^2\rho\right)\right) && \text{(D.10)}
\end{aligned}$$

One can always simplify by removing the parts not related to x and say:

$$p_{\bar{w}_{i,j}}(x|\tilde{w}^a_{i,j}, w^a_{i,j}) \propto \exp\left(-\frac{1}{2}\left((\rho_n + \rho)\left(x - \left(w^a_{i,j}\frac{\rho_n}{\rho_n + \rho} + \frac{\rho}{\rho_n + \rho}\tilde{w}^a_{i,j}\right)\right)^2\right)\right) \tag{D.11}$$

Which is a normal distribution, such that

$$\bar{w}_{i,j}|\tilde{w}^a_{i,j}, w^a_{i,j} \sim \mathcal{N}(w^a_{i,j}\frac{\rho_n}{\rho_n + \rho} + \frac{\rho}{\rho_n + \rho}\tilde{w}^a_{i,j}, \frac{1}{\rho_n + \rho}) \tag{D.12}$$

E Application model

The model considers a job seeker i, choosing which job to apply for, and whether to apply at all, given a set of jobs J.

Assuming all jobs are accepted, the perceived value of being offered a job is $v\tilde{v}_{i,j}$. We assume the value of being unemployed for another period is $\tilde{v}_{i,0}$. People perceive the probability of getting a job offer for this type to be $\tilde{p}_{i,j}$. We allow a behavioral component, α to determine how much the job seeker considers $\tilde{p}_{i,j}$, when deciding which job to apply for (with $\alpha = 1$, being the rational case). The expected value of applying for job is then:

$$A_{i,j>0} = (\tilde{p}_{i,j})^{\alpha}\tilde{v}_{i,j} + (1-(\tilde{p}_{i,j})^{\alpha})\tilde{v}_{i,0} = (\tilde{p}_{i,j})^{\alpha}(\tilde{v}_{i,j} - \tilde{v}_{i,0}) + \tilde{v}_{i,0} \tag{E.1}$$

$(\tilde{v}_{i,j} - \tilde{v}_{i,0})$ is the surplus value of getting the job, which we models as $s_{i,j}(\epsilon_{i,j})^{\sigma}$. Which include an expected part and the taste shocks, which are learned upon seeing the specific job openings. This gives:

$$A_{i,j>0} = (\tilde{p}_{i,j})^{\alpha} s_{i,j}(\epsilon_{i,j})^{\sigma} + \tilde{v}_{i,0} \tag{E.2}$$

The agent can also choose not to apply for any job $j = 0$m and receive perceived utility $\tilde{v}_{i,0}$ plus a positive utility component, b_i (intuitively this can be the value of free time from not having to write an application), we then have:

$$A_{i,0} = \tilde{v}_{i,0} + b_i \tag{E.3}$$

After looking through job postings (discovering $\epsilon_{i,j}$) the agent applies to the job with the highest expected utility:

$$j_i^* = \arg\max_{j \in J} A_{i,j} \tag{E.4}$$

We can subtract $\tilde{v}_{i,0}$ and take logs for the same j_i^*:

$$j_i^* = \arg\max_{j} O_i(j) \tag{E.5}$$

$$O_i(j) = \ln(A_{i,j} - \tilde{v}_{i,0}) = \begin{cases} \alpha \ln \tilde{p}_{i,j} + \ln s_{i,j} + \sigma \ln \epsilon_{i,j} & j > 0 \\ \ln b_i & j = 0 \end{cases} \tag{E.6}$$

For the usual logit solution, we assume that $\ln \epsilon_{i,j}$ is logistically distributed.

Then we have that before seeing the actual vacancies and learning $\epsilon_{i,t}$, the likelihood of applying to each job is:

$$\kappa_{i,j>0} = \frac{\exp\left(\frac{\alpha}{\sigma}\ln \tilde{p}_{i,j} + \frac{1}{\sigma}\ln s_{i,j}\right)}{b_i^{\frac{1}{\sigma}} + \sum_{j' \neq 0} \exp\left(\frac{\alpha}{\sigma}\ln \tilde{p}_{i,j'} + \frac{1}{\sigma}\ln s_{i,j'}\right)} \tag{E.7}$$

And the likelihood of not applying to any job is:

$$\kappa_{i,0} = \frac{b_i^{\frac{1}{\sigma}}}{b_i^{\frac{1}{\sigma}} + \sum_{j' \neq 0} \exp\left(\frac{\alpha}{\sigma}\ln \tilde{p}_{i,j'} + \frac{1}{\sigma}\ln s_{i,j'}\right)} \tag{E.8}$$

We can write up the log ratio of $\kappa_{i,j>0}$ to $\kappa_{i,0}$:

$$k_{i,j} = \ln \kappa_{i,j} - \ln \kappa_{i,0} = \frac{\alpha}{\sigma} \ln \tilde{p}_{i,j} + \frac{1}{\sigma} \ln s_{i,j} - \frac{1}{\sigma} \ln b_i \tag{E.9}$$

We assume the functional form of the expected part of surplus value to be linear in the log of perceived wage, $\tilde{w}_{i,j} \equiv \ln \widetilde{W}_{i,j}$ plus the vector product $\boldsymbol{\delta}_{2,i}\tilde{\boldsymbol{z}}_{i,j}$, which returns a scalar specifying the perceived utility to be gained from the job not related to the starting wage, which is not modeled explicitly and includes any amenity that is important to the person except for the wage:

$$\ln s_{i,j} = \sigma(\delta_1 \tilde{w}_{i,j} + \boldsymbol{\delta}_{2,i}\tilde{\boldsymbol{z}}_{i,j}) \tag{E.10}$$

Giving the equation:

$$k_{i,j} = -\frac{1}{\sigma} \ln b_i + \frac{\alpha}{\sigma} \ln \tilde{p}_{i,j} + \delta_1 \tilde{w}_{i,j} + \boldsymbol{\delta}_{2,i}\tilde{\boldsymbol{z}}_{i,j} \tag{E.11}$$

Which includes a person fixed component, $\mu_i = -\frac{1}{\sigma} \ln b_i$, the wage term, and something analogous to an error term $\varepsilon_{i,j} = \boldsymbol{\delta}_{2,i}\tilde{\boldsymbol{z}}_{i,j}$.

An similar equation can also be obtained by rewriting the model in way that is analogue to rewriting a fixed effects model to a first difference model. Because there are only 3 jobs the estimates from the two methods are indistinguishable. This allows one to view the regression as comparisons between two jobs for each person, and would be the following:

$$\ln \kappa_{i,j} - \ln \kappa_{i,j'} = \frac{\alpha}{\sigma}(\ln \tilde{p}_{i,j} - \tilde{p}_{i,j'}) + \delta_1(\tilde{w}_{i,j} - \tilde{w}_{i,j'}) + \boldsymbol{\delta}_{2,i}\tilde{\boldsymbol{z}}_{i,j} - \boldsymbol{\delta}_{2,i}\tilde{\boldsymbol{z}}_{i,j'} \tag{E.12}$$

E.1 Elasticity

Elasticity To transform an estimate of δ_1 into something that is comparable with the literature, one can derive the elasticity of perceived wage on the likelihood of applying for a job j.

Starting from equation (E.7) and using the functional form from (E.10), the elasticity can be derived as:

$$\mathcal{E}_{i,j} = \frac{\partial \ln \kappa_{i,j}}{\partial \tilde{w}_{i,j}} = \frac{\partial\Big(\big(\frac{\alpha}{\sigma} \ln \tilde{p}_{i,j} + \frac{1}{\sigma} \ln s_{i,j}\big) - \ln\Big(b_i^{\frac{1}{\sigma}} + \sum_{k=1}^{J} \exp\big(\frac{\alpha}{\sigma} \ln \tilde{p}_{i,k} + \frac{1}{\sigma} \ln s_{i,k}\big)\Big)\Big)}{\partial \tilde{w}_{i,j}} \tag{E.13}$$

$$= \frac{\partial \frac{1}{\sigma} \ln s_{i,j}}{\partial \tilde{w}_{i,j}} - \frac{1}{\Big(b_i^{\frac{1}{\sigma}} + \sum_{k=1}^{J} \exp\big(\frac{\alpha}{\sigma} \ln \tilde{p}_{i,k} + \frac{1}{\sigma} \ln s_{i,k}\big)\Big)} \frac{\partial\Big(b_i^{\frac{1}{\sigma}} + \sum_{k=1}^{J} \exp\big(\frac{\alpha}{\sigma} \ln \tilde{p}_{i,k} + \frac{1}{\sigma} \ln s_{i,k}\big)\Big)}{\partial \tilde{w}_{i,j}} \tag{E.14}$$

$$= \delta_1 - \frac{\partial \frac{1}{\sigma} \ln s_{i,j}}{\partial \tilde{w}_{i,j}} \cdot \frac{\exp\big(\frac{\alpha}{\sigma} \ln \tilde{p}_{i,j} + \frac{1}{\sigma} \ln s_{i,j}\big)}{\Big(b_i^{\frac{1}{\sigma}} + \sum_{k=1}^{J} \exp\big(\frac{\alpha}{\sigma} \ln \tilde{p}_{i,k} + \frac{1}{\sigma} \ln s_{i,k}\big)\Big)} \tag{E.15}$$

$$= \delta_1(1 - \kappa_{i,j}) \tag{E.16}$$

Following Train (2009) the aggregate elasticity can be found by sample enumeration, which amounts to using an estimate of δ_1, $\hat{\delta}_1$, and calculating the average elasticity in the sample:

$$\widehat{\mathcal{E}}^a(\hat{\delta}) = \frac{1}{N \cdot J} \sum_{i=1,j=1}^{N,J} \hat{\delta}_1 (1 - \kappa_{i,j}) = \hat{\delta}_1 \frac{1}{N \cdot J} \sum_{i=1,j=1}^{N,J} (1 - \kappa_{i,j}) \tag{E.17}$$

It should be noted that this elasticity is the elasticity with respect to the hypothetical choice probabilities and therefore not directly comparable to the elasticity of likelihood of receiving an application to a vacancy with respect to the posted wage. Interpreting this as an elasticity requires trusting the logit structure, which includes the independence of irrelevant alternatives-assumption. This requires assuming that the relative change induced in $\kappa_{i,0}$ by a change in $\tilde{w}_{i,j}$ is the same as the relative change induced in $\kappa_{i,j'}$ for $j' \neq j$.

F Testing for non-linearity

In this section, we briefly outline the enhanced binscatter methodology introduced by Cattaneo et al. (2024), which we will use to test the assumption about a homogeneous updating rule. Cattaneo et al. (2024) provide a comprehensive framework for binscatter plots that includes estimating conditional means with optimal binning and quantifying uncertainty through confidence bands. This approach is out-of-the-box ready and fixes some issues in traditional binscatter plots when controlling for covariates which comes from the breakdown of the Frisch–Waugh–Lovell Theorem under non-linearity, meaning partialling out is not valid.

The method involves partitioning the support of the covariate x into a fixed number of bins and then plotting the average outcome for observations within each bin. We use the misperception gap, $\tilde{w}^a_{i,j} - w^a_{i,j}$ as the covariate, and log of the perceived wage $\tilde{w}_{i,j}$ as the outcome for both the treated and the control group, if the linearity assumption holds, both these plots should be linear. Following equation (9) we control for $\tilde{w}^a_{i,j}$. Concretely the estimand is:

$$\begin{bmatrix} \boldsymbol{\beta}^C_B \\ \rho^C_B \end{bmatrix} = \underset{\boldsymbol{\beta} \in R^B, \rho^C \in R}{arg\,min}\, \mathbb{E}\Big[\big(\tilde{w}_{i,j} - \boldsymbol{b}(\tilde{w}^a_{i,j} - w^a_{i,j})'\boldsymbol{\beta} - \rho^C \tilde{w}^a_{i,j}\big)^2\Big] \tag{F.1}$$

Where B is the number of chosen bins and $\boldsymbol{b}(\tilde{w}^a_{i,j} - w^a_{i,j})$ is a function that returns a $B \times 1$ vector of dummies for whether $\tilde{w}^a_{i,j} - w^a_{i,j}$ is inside the brackets of each bin.

One can also include person-fixed effects by estimating the estimand:

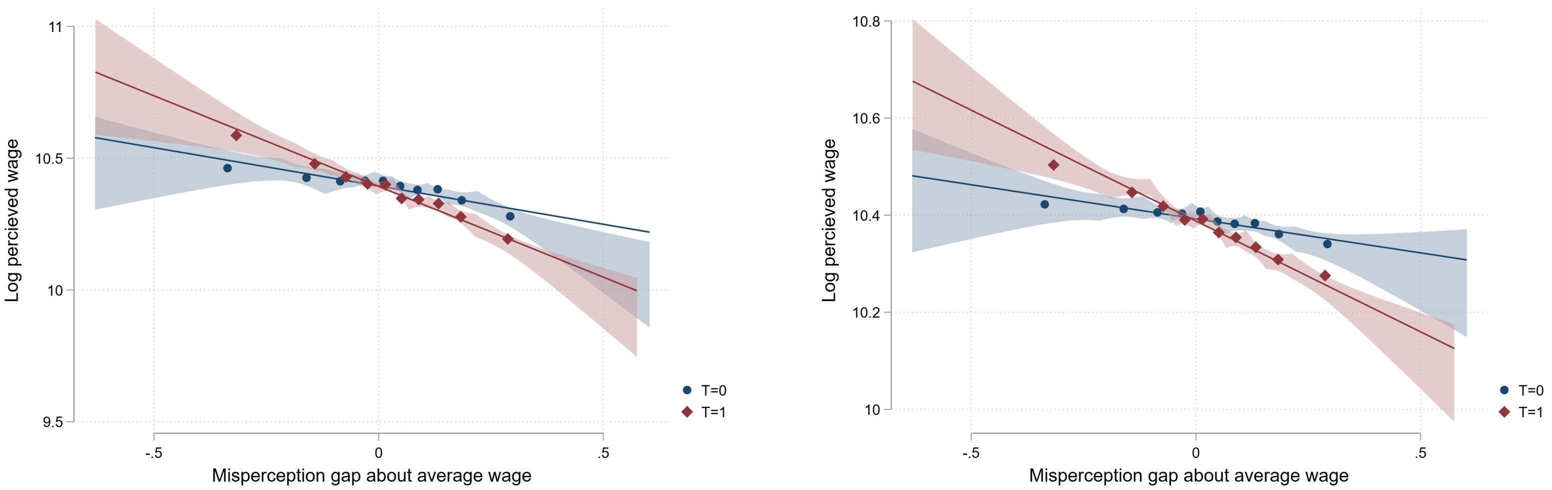
$$\begin{bmatrix} \boldsymbol{\beta}^{FE}_B \\ \rho^{FE}_B \\ \boldsymbol{\mu}_B \end{bmatrix} = \underset{\boldsymbol{\beta} \in R^B, \rho^{FE} \in R, \boldsymbol{\mu} \in R^N}{arg\,min}\, \mathbb{E}\Big[\big(\tilde{w}_{i,j} - \boldsymbol{b}(\tilde{w}^a_{i,j} - w^a_{i,j})'\boldsymbol{\beta} - \mu_i - \rho^{FE} \tilde{w}^a_{i,j}\big)^2\Big] \tag{F.2}$$

Where $\boldsymbol{\mu}_B$ is a vector of the fixed effects.

These are estimated for both the treated sample and the control sample. The predicted values of $\tilde{w}_{i,j}$ for each bin are plotted in Figure F1. The software of Cattaneo et al. (2024) then allows for making bootstrapped confidence bands, clustered at the education level.

F.1 Results of non-linearity test

Figure F1: Correlation between misperception gap and perceived wages for treated and non-treated



(a) Controlling for $\tilde{w}^a_{i,j}$ (b) Also including respondent fixed effects

Note: This figure plots predicted values of $\tilde{w}_{i,j}$ across $(\tilde{w}^a_{i,j} - w^a_{i,j})$, following Cattaneo et al. (2024) and as explained in Section F. Both include control for the log of the perceived average wage, $\tilde{w}^a_{i,j}$. Confidence bands are bootstrapped following Cattaneo et al. (2024) with 10,000 draws across 50 bins, clustered at the education level.

Figure F1a shows the approach described in Section F, plotting the predicted $\tilde{w}^a_{i,j}$ across the misperception gap, $(\tilde{w}^a_{i,j} - w^a_{i,j})$ for both the treatment and the control group. These predicted values essentially come from equation (9) from Section 4.1, forcing $\psi_2 = 0$,[24] and allowing $\psi_1 \varrho$ and $\frac{\rho_n \psi_1}{\rho + \rho_n}$ to be heterogeneous across the pre-treatment perception gap in average wages. The slope of the line for the control group shows the degree of spurious learning under linearity, $\psi_1 \varrho$. The slope of the treatment group is the degree of learning about the average wage and its effect on beliefs about own wage, which also includes the base degree of spurious learning following. The difference in slopes is the actual treatment effect related to the initial misperception gap. The inclusion of fixed effects in figure F1b, increases precision but does not change the overall picture. The plots add more suggestive evidence that the linearity assumption for the first stage is plausible.

[24] No treatment effect from the level of log perceived average wage $\tilde{w}^a_{i,j}$.

G Names of degrees, 6-digit DISCED codes

Table G1: Translated names of presented degrees

2 digit name	4 digit name	6 digit name
Primary or lower secondary education	Danish elementary school, less than 9th grade	
	Elementary school, 9th or 10th grade	
Upper secondary education	General upper secondary education	Stx
	General upper secondary education	Hf
	General upper secondary education	Student course
	Vocational upper secondary education	Hhx 1 year
	Vocational upper secondary education	Hhx 3 year
	Vocational upper secondary education	Htx
	International upper secondary education	Pre International Baccalaureate
	International upper secondary education	International Baccalaureate
	International upper secondary education	Danish-French Bacalaurétte
	International upper secondary education	Other foreign upper secondary education
	International upper secondary education	European Baccalaureate
Vocational basic courses	Care, health and pedagogy, basic course	Health, care and pedagogy, basic course
	Clerical, trade and business services, basic course	The mercantile field, basic course
	Clerical, trade and business services, basic course	Building and user service, basic course
	Food, agriculture and experiences, basic course	Food, agriculture and experiences, basic course
	Technology, construction and transportation, basic course	Technology, construction and transportation, basic course
Vocational education	Care, health and pedagogy	Health, care and pedagogy without further specification
	Care, health and pedagogy	Pedagogical courses
	Care, health and pedagogy	Social and healthcare education
	Care, health and pedagogy	Hospital technical assistants
	Care, health and pedagogy	Dental clinic assistants
	Care, health and pedagogy	Health, care and pedagogy, other
	Clerical, trade and business services	Mercantile education without further specification
	Clerical, trade and business services	Clerical courses
	Clerical, trade and business services	Finance courses
	Clerical, trade and business services	Retail courses
	Clerical, trade and business services	Trade courses
	Clerical, trade and business services	Customer contact training
	Clerical, trade and business services	Health care aervice secretary
	Clerical, trade and business services	Mercantile education, other
	Clerical, trade and business services	Property and other services, without further specification
	Clerical, trade and business services	Property service technician
	Clerical, trade and business services	Service assistant courses
	Clerical, trade and business services	Security guard
	Food processing	Baker and confectioner
	Food processing	Butcher etc.
	Food processing	Intestinal cleanser
	Food processing	Dairy education
	Food processing	Nutritional assistants
	Food processing	Gastronomy courses
	Food processing	Waiter educations
	Food processing	Receptionists
	Food processing	Food processing, other
	Agriculture and nature	Agricultural courses
	Agriculture and nature	Gardening courses
	Agriculture and nature	Greenkeeper, groundsman etc.
	Agriculture and nature	Forestry and nature engineering
	Agriculture and nature	Pet sitter
	Agriculture and nature	Veterinary nursing
	Agriculture and nature	Agriculture and nature, other
	Personal services	Event coordinator
	Personal services	Fitness instructor
	Personal services	Hairdresser
	Personal services	Cosmetology courses
	The construction field	Construction, without further specification
	The construction field	Paver and skilled construction worker etc.
	The construction field	Building painter
	The construction field	Glazier
	The construction field	Bricklayer
	The construction field	Woodworker and machine carpenter etc.
	The construction field	Carpenter etc.
	The construction field	Plumbing technology
	The construction field	Roofer
	The construction field	Other building and construction educations
	The technology field, power and electronics, etc.	Power and electronics, without further specification
	The technology field, power and electronics, etc.	Automation and automation process
	The technology field, power and electronics, etc.	Computer and communication engineering

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	The technology field, power and electronics, etc.	Electrician
	The technology field, power and electronics, etc.	Electronics and process operators etc.
	The technology field, power and electronics, etc.	Electronics and low voltage training
	The technology field, power and electronics, etc.	Other electricity and electronics courses
	The technology field, Graphical engineering and media production	Graphical technology and media production without further specification
	The technology field, Graphical engineering and media production	Film and television production
	The technology field, Graphical engineering and media production	Photographer
	The technology field, Graphical engineering and media production	Digital media and web integrator
	The technology field, Graphical engineering and media production	Graphical designer
	The technology field, Graphical engineering and media production	Sign technician
	The technology field, bicycle, car and ship mechanics, etc.	Bicycle, car and ship mechanics without further specification
	The technology field, bicycle, car and ship mechanics, etc.	Bicycle, car and ship mechanics etc.
	The technology field, bicycle, car and ship mechanics, etc.	Bodywork courses
	The technology field, bicycle, car and ship mechanics, etc.	Carriage painter
	The technology field, mechanical engineering and production	Mechanical engineering and production without further specification
	The technology field, mechanical engineering and production	Industrial technician and CNC technician
	The technology field, mechanical engineering and production	Industrial operators and producers
	The technology field, mechanical engineering and production	Wind turbine operator
	The technology field, mechanical engineering and production	Refrigeration technician and oil furnace technician
	The technology field, mechanical engineering and production	Blacksmith training
	The technology field, mechanical engineering and production	Foundry technician
	The technology field, mechanical engineering and production	Tool training
	The technology field, mechanical engineering and production	Plastic maker
	The technology field, mechanical engineering and production	Maritime trades
	The technology field, mechanical engineering and production	Ship engineering and ship assembly
	The technology field, mechanical engineering and production	Precision mechanic and watchmaker
	The technology field, mechanical engineering and production	Textile and clothing crafts
	The technology field, mechanical engineering and production	Dental laboratory technician
	The technology field, mechanical engineering and production	Orthopedist
	The technology field, mechanical engineering and production	Shoemaker and orthopedic shoemaker
	The technology field, mechanical engineering and production	Mechanical engineering and production, other educations
	Technical and industrial education, other	Technical and industrial education, other
	Transportation and the logistics fields	Transportation and logistics without further specification
	Transportation and the logistics fields	Warehouse, port and terminal training
	Transportation and the logistics fields	Airport training
	Transportation and the logistics fields	Train preparation courses
	Transportation and the logistics fields	Rescue training
	Transportation and the logistics fields	Driver
	Transportation and the logistics fields	Transportation and logistics, other
	Other vocational training	Commercial fishing etc.
	Other vocational training	Maritime education
	Other vocational training	Defense
	Other vocational training	Other vocational training
	Basic course	Basic course
	Vocational education without further specification	Vocational education without further specification
Labor market education	Care, health and pedagogy	Health and care
	Care, health and pedagogy	Pedagogical
	Clerical, trade and business services	Management
	Clerical, trade and business services	The clerical field
	Clerical, trade and business services	Wholesale and retail trade
	Clerical, trade and business services	Personal services
	Clerical, trade and business services	Property and other services
	Clerical, trade and business services	Prison and police
	Food processing	Baker and confectioner
	Food processing	Processing and production of food
	Food processing	Kitchen, hotel and restaurant
	Agriculture and nature	Agriculture
	Agriculture and nature	Gardening
	Agriculture and nature	Forest and nature
	Agriculture and nature	Agriculture and nature, other
	The construction field	Construction, without further specification
	The construction field	Building painter
	The construction field	Bricklayer
	The construction field	Carpenter etc.
	The construction field	Plumbing technology
	The construction field	Roofer

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	The construction field	The construction field, other
	The technology field, power and electronics, etc.	Power and electronics, without further specification
	The technology field, power and electronics, etc.	Automation and automation process
	The technology field, power and electronics, etc.	Computer and communication engineering
	The technology field, power and electronics, etc.	Power and electronics, other
	The technology field, Graphical engineering and media production	Film and television production
	The technology field, Graphical engineering and media production	Graphical technique
	The technology field, bicycle, car and ship mechanics, etc.	Bicycle, car and ship mechanics, etc.
	The technology field, mechanical engineering and production	Mechanical engineering and production without further specification
	The technology field, mechanical engineering and production	Industrial operators and producers
	The technology field, mechanical engineering and production	Refrigeration and oil furnace technology
	The technology field, mechanical engineering and production	Blacksmith
	The technology field, mechanical engineering and production	Tool training
	The technology field, mechanical engineering and production	The plastic field
	The technology field, mechanical engineering and production	Mechanical engineering and production, other
	Transportation and the logistics fields	Transportation and logistics without further specification
	Transportation and the logistics fields	Warehouse, port and terminal
	Transportation and the logistics fields	Driver
	Transportation and the logistics fields	Transportation and logistics, other
	Other labor market education	Commercial fishing etc.
	Other labor market education	Competency assessment
	Other labor market education	Without further specification
	Further training, special work, skilled, and other groups	Care, health and pedagogy
	Further training, special work, skilled, and other groups	Clerical, trade and business services
	Further training, special work, skilled, and other groups	Food processing
	Further training, special work, skilled, and other groups	Agriculture and nature
	Further training, special work, skilled, and other groups	The construction field
	Further training, special work, skilled, and other groups	The technology field, power and electronics, etc.
	Further training, special work, skilled, and other groups	The technology field, Graphical engineering and media production
	Further training, special work, skilled, and other groups	The technology field, bicycle, car and ship mechanics, etc.
	Further training, special work, skilled, and other groups	The technology field, mechanical engineering and production
	Further training, special work, skilled, and other groups	Technical and industrial education, other
	Further training, special work, skilled, and other groups	Transportation and the logistics fields
	Further training, special work, skilled, and other groups	Other labor market education
Short-cycle higher education	Pedagogical	Other teacher education
	Media and communication	Communication and dissemination without further specification
	Media and communication	Multimedia design
	Media and communication	Graphical design and communication
	Media and communication	Media and communication, other
	Humanities and theological	Business language, correspondents
	Humanities and theological	Business language, other educations
	Arts	Artistic without further specification
	Arts	Crafts
	Arts	Music
	Arts	Artistic, other
	Social Studies, Business and administration	Social Studies, Business and administration without further specification
	Social Studies, Business and administration	Administrative economist etc.
	Social Studies, Business and administration	Financial advisor etc.
	Social Studies, Business and administration	Trade and marketing economist, etc.
	Social Studies, Business and administration	Service economist etc.
	Social Studies, Business and administration	Transportation and logistics economist etc.
	Technical	Technical, without further specification
	Technical	Construction and construction engineering
	Technical	Electronics and IT
	Technical	Installer of high voltage and plumbing technology
	Technical	Energy and high voltage engineering
	Technical	Mechanical engineering
	Technical	Environmental engineering
	Technical	Health and care technology
	Technical	Technical, production and product development
	Technical	Technically, other
	Food, bio and laboratory technology	Food, bio- and laboratory technology without further specification
	Food, bio and laboratory technology	Laboratory, food and process technology
	Food, bio and laboratory technology	Laboratory Technician
	Agriculture, nature and environment	Agriculture, nature and environment without further specification
	Agriculture, nature and environment	Agricultural technology

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	Maritime	Shipmasters
	Maritime	Maritime education, other
	Healthcare professional	Health professional education without further specification
	Healthcare professional	Care work
	Healthcare professional	Dental hygienist and clinical dental technology
	Healthcare professional	Other healthcare educations
	Police and defense etc.	Police and defense without further specification
	Police and defense etc.	The prison service
	Higher education without further specification	Higher education without further specification
Medium-cycle higher education	Pedagogical	Pedagogical studies without further specification
	Pedagogical	Pedagogue
	Pedagogical	Elementary school teacher
	Pedagogical	Other teacher education
	Pedagogical	Pedagogy
	Media and communication	Communication and dissemination without further specification
	Media and communication	Graphical design and communication
	Media and communication	Journalist and journalistic work
	Media and communication	Photojournalist
	Media and communication	Sign language interpreter
	Media and communication	Media and communication, other
	Humanities and theological	Library and information courses
	Humanities and theological	Business language, correspondents
	Humanities and theological	Business language diploma, etc.
	Humanities and theological	Business language, other educations
	Arts	Artistic without further specification
	Arts	Film and Television
	Arts	Theater
	Arts	Dance
	Arts	Actor
	Arts	Design education
	Arts	Music
	Social Studies, Business and administration	Social Studies, Business and administration without further specification
	Social Studies, Business and administration	Administration etc.
	Social Studies, Business and administration	Finance etc.
	Social Studies, Business and administration	Trade and marketing, etc.
	Social Studies, Business and administration	Leisure, culture and sports etc.
	Social Studies, Business and administration	Social counseling and mediation, etc.
	Social Studies, Business and administration	Business Economics, HD
	Social science	Administration, management and management
	Technical	Technical, without further specification
	Technical	Construction and construction engineering
	Technical	Electronics and IT, technical
	Technical	Energy and high voltage engineering
	Technical	Mechanical engineering
	Technical	Health and care technology
	Technical	Technical-natural sciences, combined
	Technical	Technical, production and product development
	Technical	Technically, other
	Technical science	Technical science without further specification
	Technical science	Biotechnology and chemical technology
	Technical science	Building and construction engineering
	Technical science	Electronics and IT, technical science
	Technical science	Electronics, other
	Technical science	Energy technology
	Technical science	Mechanical engineering, technical science
	Technical science	Technology, production and product development
	Food, Bio- and laboratory engineering	Food, Bio- and laboratory technology without further specification
	Food, Bio- and laboratory engineering	Laboratory, food and process technology
	Food, Bio- and laboratory engineering	Nutrition and Health
	Agriculture, nature and environment	Agriculture, nature and environment without further specification
	Agriculture, nature and environment	Agricultural Science
	Agriculture, nature and environment	Landscape architecture and management
	Agriculture, nature and environment	Natural resources and environment
	Maritime	Shipmasters
	Maritime	Ship officers
	Maritime	Maritime education, other
	Healthcare professional	Health professional education without further specification
	Healthcare professional	Biomedical Laboratory Technologist
	Healthcare professional	Occupational and physiotherapist
	Healthcare professional	Midwife

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	Healthcare professional	Care work, further education
	Healthcare professional	Radiographer
	Healthcare professional	Nursing and health care
	Healthcare professional	Dental hygienist and clinical dental technology
	Healthcare professional	Other healthcare educations
	Police and defense etc.	Police and defense without further specification
	Police and defense etc.	Police
	Police and defense etc.	Officer in the defense
	Higher education without further specification	Higher education without further specification
Bachelor degree	Pedagogical	Pedagogy
	Humanities and theological	Humanities without further specification
	Humanities and theological	Archaeology
	Humanities and theological	Classical studies
	Humanities and theological	Ethnology
	Humanities and theological	Philosophy
	Humanities and theological	History
	Humanities and theological	History of ideas and science
	Humanities and theological	Film, theater and musicology, etc.
	Humanities and theological	Literary studies
	Humanities and theological	Communication and dissemination
	Humanities and theological	Journalism and rhetoric
	Humanities and theological	Information science, etc.
	Humanities and theological	Experience Design
	Humanities and theological	Music therapy
	Humanities and theological	Religion and religious studies
	Humanities and theological	Linguistics
	Humanities and theological	Danish-Nordic language and literature
	Humanities and theological	Languages of the Middle East
	Humanities and theological	Romance languages
	Humanities and theological	Slavic, Eastern Europe and the Balkans
	Humanities and theological	Southeast Asian languages
	Humanities and theological	East Asian languages
	Humanities and theological	West Germanic languages (English, German or Dutch)
	Humanities and theological	Classical languages and philology
	Humanities and theological	Ancient languages
	Humanities and theological	Other language educations
	Humanities and theological	Business language bachelors
	Humanities and theological	Business language, combined
	Humanities and theological	Humanities, other
	Humanities and theological	Theology
	Arts	Artistic without further specification
	Arts	Architecture
	Arts	Visual arts
	Arts	Theater
	Arts	Dance
	Arts	Actor
	Arts	Designer
	Arts	Conservator
	Arts	Music
	Natural science	Natural sciences without further specification
	Natural science	Biochemistry and molecular biology
	Natural science	Biology
	Natural science	Biomedicine and molecular medicine
	Natural science	Medicinal chemistry etc.
	Natural science	Nanoscience and nanobioscience
	Natural science	Physics and physical subjects
	Natural science	Geodesy and geoinformatics
	Natural science	Geography
	Natural science	Geology
	Natural science	Sports
	Natural science	Chemistry
	Natural science	Mathematics
	Natural science	statistics
	Natural science	Natural science IT educations
	Social science	Social sciences without further specification
	Social science	Administration, management and management
	Social science	Anthropology
	Social science	Globalization and international social studies
	Social science	Law
	Social science	Political science and other political science courses
	Social science	Psychology
	Social science	Sociology
	Social science	Economics

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	Social science	Business economics without further specification
	Social science	Business Economics, HA
	Social science	Business economics-business language
	Technical science	Technical science without further specification
	Technical science	Biotechnology and chemical technology
	Technical science	Nanotechnology and nanobiotechnology
	Technical science	Building and construction engineering
	Technical science	Electronics and IT
	Technical science	Electronics, other
	Technical science	Energy technology
	Technical science	Land surveying Science
	Technical science	Mechanical engineering
	Technical science	Environmental technology etc.
	Technical science	Health and welfare technology, etc.
	Technical science	Technology-natural sciences, combined
	Technical science	Technology, production and product development
	Technical science	Technical science, other
	Food, Bio- and laboratory engineering	Food and Nutrition Science
	Food, Bio- and laboratory engineering	Food and nutrition science, other
	Agriculture, nature and environment	Agriculture, nature and environment without further specification
	Agriculture, nature and environment	Agricultural Science
	Agriculture, nature and environment	Animal science
	Agriculture, nature and environment	Veterinary medicine
	Agriculture, nature and environment	Forestry Science
	Agriculture, nature and environment	Agriculture, food and environment combined
	Agriculture, nature and environment	Landscape architecture and management
	Agriculture, nature and environment	Natural resources and environment
	Health science	Health sciences without further specification
	Health science	Pharmacy and Pharmaceutical Sciences
	Health science	Public health science
	Health science	IT and health sciences
	Health science	Medicine
	Health science	Odontology
	Health science	Healthcare sciences, other
	Higher education without further specification	Higher education without further specification
Long-cycle higher education	Pedagogical	Pedagogical studies without further specification
	Pedagogical	Other teacher education
	Pedagogical	Pedagogy
	Humanities and theological	Humanities without further specification
	Humanities and theological	Archaeology
	Humanities and theological	Classical studies
	Humanities and theological	Ethnology
	Humanities and theological	Philosophy
	Humanities and theological	History
	Humanities and theological	History of ideas and science
	Humanities and theological	Film, theater and musicology, etc.
	Humanities and theological	Literary studies
	Humanities and theological	Communication and dissemination
	Humanities and theological	Journalism and rhetoric
	Humanities and theological	Information science, etc.
	Humanities and theological	Digital design and interaction design
	Humanities and theological	Experience Design
	Humanities and theological	Music therapy
	Humanities and theological	Religion and religious studies
	Humanities and theological	Linguistics
	Humanities and theological	Danish-Nordic language and literature
	Humanities and theological	Languages of the Middle East
	Humanities and theological	Romance languages
	Humanities and theological	Slavic, Eastern Europe and the Balkans
	Humanities and theological	Southeast Asian languages
	Humanities and theological	East Asian languages
	Humanities and theological	West Germanic languages (English, German or Dutch)
	Humanities and theological	Classical languages and philology
	Humanities and theological	Ancient languages
	Humanities and theological	Other language educations
	Humanities and theological	Business language, interpreter.
	Humanities and theological	Business language, ling.merc.
	Humanities and theological	Business language, combined
	Humanities and theological	Humanities, other
	Humanities and theological	Theology
	Arts	Artistic without further specification
	Arts	Architecture

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	Arts	Visual arts
	Arts	Dance
	Arts	Designer
	Arts	Conservator
	Arts	Music
	Arts	Performing arts
	Natural science	Natural sciences without further specification
	Natural science	Biochemistry and molecular biology
	Natural science	Biology
	Natural science	Biomedicine and molecular medicine
	Natural science	Medicinal chemistry etc.
	Natural science	Nanoscience and nanobioscience
	Natural science	Physics and physical subjects
	Natural science	Geodesy and geoinformatics
	Natural science	Geography
	Natural science	Geology
	Natural science	Sports
	Natural science	Chemistry
	Natural science	Mathematics
	Natural science	statistics
	Natural science	Natural science IT educations
	Natural science	Natural sciences, other
	Social science	Social sciences without further specification
	Social science	Administration, management and management
	Social science	Anthropology
	Social science	Globalization and international social studies
	Social science	Law
	Social science	Political science and other political science courses
	Social science	Psychology
	Social science	Sociology
	Social science	Economics
	Social science	Business economics without further specification
	Social science	Business Economics, MSc.
	Social science	Business economics-business language
	Technical science	Technical science without further specification
	Technical science	Biotechnology and chemical technology
	Technical science	Nanotechnology and nanobiotechnology
	Technical science	Building and construction engineering
	Technical science	Electronics and IT
	Technical science	Electronics, other
	Technical science	Energy technology
	Technical science	Land surveying Science
	Technical science	Mechanical engineering
	Technical science	Environmental technology etc.
	Technical science	Health and welfare technology, etc.
	Technical science	Technology-natural sciences, combined
	Technical science	Technology, production and product development
	Technical science	Technical science, other
	Food, Bio- and laboratory engineering	Food and Nutrition Science
	Food, Bio- and laboratory engineering	Food and nutrition science, other
	Agriculture, nature and environment	Agriculture, nature and environment without further specification
	Agriculture, nature and environment	Agricultural Science
	Agriculture, nature and environment	Animal science
	Agriculture, nature and environment	Veterinary medicine
	Agriculture, nature and environment	Forestry Science
	Agriculture, nature and environment	Agriculture, food and environment combined
	Agriculture, nature and environment	Landscape architecture and management
	Agriculture, nature and environment	Natural resources and environment
	Health science	Health sciences without further specification
	Health science	Occupational therapy
	Health science	Physiotherapy
	Health science	Pharmacy and Pharmaceutical Sciences
	Health science	Public health science
	Health science	IT and health sciences
	Health science	Midwifery Science
	Health science	Nursing Science
	Health science	Medicine
	Health science	Odontology
	Health science	Specialist doctor and specialist dentist
	Health science	Healthcare sciences, other
	Police and defense etc.	Police and defense without further specification
	Police and defense etc.	Defense, other

Continued on next page

Table G1: Translated names of presented educations

2 digit name	4 digit name	6 digit name
	Police and defense etc.	Officer in the defense
	Higher education without further specification	Higher education without further specification
PhD and research education	Pedagogical	Pedagogical studies without further specification
	Pedagogical	Pedagogy
	Humanities and theological	Humanities without further specification
	Humanities and theological	Philosophy
	Humanities and theological	Communication and dissemination
	Humanities and theological	Information science, etc.
	Humanities and theological	Business language, ling.merc.
	Humanities and theological	Theology
	Arts	Artistic without further specification
	Arts	Architecture
	Arts	Conservator
	Arts	Music
	Natural science	Natural sciences without further specification
	Natural science	Mathematics
	Social science	Social sciences without further specification
	Social science	Anthropology
	Social science	Law
	Social science	Political science and other political science courses
	Social science	Psychology
	Social science	Sociology
	Social science	Economics
	Social science	Business economics without further specification
	Technical science	Technical science without further specification
	Technical science	Land surveying Science
	Agriculture, nature and environment	Agriculture, nature and environment without further specification
	Agriculture, nature and environment	Agricultural Science
	Agriculture, nature and environment	Veterinary medicine
	Health science	Health sciences without further specification
	Health science	Pharmacy and Pharmaceutical Sciences
	Health science	Medicine
	Health science	Odontology
	Higher education without further specification	Higher education without further specification