What Works for the Unemployed?
Evidence from Quasi-Random Caseworker Assignments*

Anders Humlum† **Jakob R. Munch**‡ **Mette Rasmussen**§

March 23, 2023

Abstract

This paper examines if active labor market programs help unemployed job seekers find jobs using a novel random caseworker instrumental variable (IV) design. Leveraging administrative data from Denmark, our identification strategy exploits that (i) job seekers are quasi-randomly assigned to caseworkers, and (ii) caseworkers differ in their tendencies to assign similar job seekers to different programs. Using our IV strategy, we find assignment to classroom training increases employment rates by 25% two years after initial job loss. This finding contrasts with the conclusion reached by ordinary least squares (OLS), which suffers from a negative bias due to selection on unobservables. The employment effects are driven by job seekers who complete the programs (post-program effects) rather than job seekers who exit unemployment upon assignment (threat effects), and the programs help job seekers change occupations. We show that job seekers exposed to offshoring – who tend to experience larger and more persistent employment losses – also have higher employment gains from classroom training. By estimating marginal treatment effects, we conclude that total employment may be increased by targeting training toward job seekers exposed to offshoring.

*We thank the Danish Agency for Labor Market and Recruitment (STAR) for providing data and the many caseworkers who helped us understand the local organization of their job centers. Thanks to Manudeep Bhuller, Dan Black, Gordon Dahl, Peter Hull, Jack Mountjoy, and Lars Skipper for comments on the paper and to Mette Rasmussen's PhD committee (Søren Leth Petersen, Sally Sadoff, and Bas van der Klaauw) for comments on an early version of the paper. Financial support from the Economic Policy Research Network (EPRN), the Rockwool Foundation, and the Independent Research Fund Denmark (1027-00011A) is gratefully acknowledged.

†University of Chicago, Booth School of Business; anders.humlum@chicagobooth.edu

‡University of Copenhagen, Department of Economics; jakob.roland.munch@econ.ku.dk

§University of Copenhagen, Department of Economics; mette.rasmussen@econ.ku.dk

1 Introduction

Active labor market policies (ALMPs) are integral to modern welfare states, but the use of ALMPs varies widely between countries. For example, Denmark spends as much as 2% of its gross domestic product (GDP) on ALMPs, whereas US labor market policies rely more on passive measures such as unemployment and disability insurance (Bown and Freund, 2019; Kreiner and Svarer, 2022). ALMPs offer a diverse set of programs, including classroom training and private and public on-the-job training, but classroom training is increasingly receiving attention as a way to mitigate skill mismatches caused by structural changes in the labor market (World Economic Forum, 2020). Job seekers displaced by, for example, trade, technology, or automation may have accumulated specific human capital that is no longer in demand in the labor market, and as a result, these job seekers, in particular, may need retraining to find employment (Hummels et al., 2018; Costa et al., 2019; Humlum and Munch, 2019).

Do these classroom-training programs help unemployed job seekers find jobs? The literature suggests private on-the-job training is the more effective program, whereas the evidence for classroom training is at best mixed. McCall et al. (2016) review the evidence for six countries and conclude for Denmark that despite the large expenditures, "classroom training is largely ineffective in raising their participants' employment rates." They continue as follows: "Moreover, also in line with the evidence for other countries, programs that involve on-the-job training and are thus more similar to regular jobs, on average perform considerably better than vocational classroom training."[1]

Yet, prior evaluations of ALMPs faced three empirical challenges. The first challenge is that job seekers self-select into programs based on their preexisting job opportunities (Ashenfelter, 1978). For example, a job seeker who already has a pending job offer may not enroll in a long-term training course. Such information is typically unobserved to the researcher but likely plays a crucial role in the selection into classroom training. Second, job seekers may reap heterogeneous returns from training programs. For example, retraining may particularly benefit job seekers whose skills have become obsolete due to globalization or automation. The literature has mainly examined effect heterogeneity across age, gender, and unemployment-

[1] McCall et al. (2016) also conclude that "More broadly, the US literature suggests that zero is sometimes, but not always, a good summary of the impact of training programs." The meta-study by Card et al. (2018) finds that "Human capital programs have small (or in some cases even negative) short term impacts, coupled with larger impacts in the medium or longer run."

duration groups, whereas exposure to globalization has rarely been considered (Hyman, 2018). Finally, comparing studies of different ALMPs is difficult because research designs differ in their study populations and time horizons (Card et al., 2018), and any single treatment often pulls job seekers in from a multitude of alternative programs (Heckman et al., 2000).

In this paper, we develop theory and use data from the Danish labor market to construct caseworker-tendency instruments that overcome the three challenges. Our identification strategy exploits the facts that, in Denmark, (i) job seekers are quasi-randomly assigned to caseworkers according to their day of birth (1-31), and (ii) caseworkers differ in their tendencies to assign job seekers to different training programs. These instrumental variables (IVs) allow us to evaluate the heterogeneous causal effects of different ALMPs within a coherent framework. IV methods are uncommon in the vast literature evaluating ALMPs, due to data limitations and lack of exogenous variation determining selection into training.[2]

We structure our analysis within a generalized Roy model of how caseworkers assign job seekers to training programs. Our setting is theoretically challenging because caseworkers choose from a menu of training options, and the benefits to job seekers are heterogeneous across these programs. We show caseworker tendencies, defined as the leave-out means of program assignment rates by caseworkers, emerge as theory-consistent instruments that satisfy the conditions for non-parametric identification of potential outcomes along all treatment margins. In particular, we provide evidence that the instruments satisfy an *extended monotonicity* condition by which tendencies for classroom training do not affect the assignment to on-the-job training and vice versa. The extended monotonicity allows us to identify heterogeneous treatment effects in our setting with multiple unordered treatments.

We establish four headline findings. First, using the caseworker-tendency instruments, we estimate that classroom-training programs have persistent positive effects on employment. Attending a classroom-training program is associated with 25% more hours of employment two years after the initial job loss. By contrast, we do not find that on-the-job training programs, such as employment programs with wage subsidies, are associated with robust employment gains.

[2]McCall et al. (2016) identify only one published study using an IV approach: Frölich and Lechner (2010). Their instrument exploits variation in training propensities across regions in Switzerland. Hyman (2018) is a more recent example of a study using an IV approach (see below for more details). Cederlöf et al. (2021) develop a similar caseworker instrument for the Swedish labor market to examine the characteristics of effective caseworkers. They do not evaluate the impact of ALMP programs on the unemployed.

Second, we compare the IV estimates to ordinary least squares (OLS) estimates that assume "selection on observables only" – a conventional assumption in the literature. Strikingly, we find the IV and OLS estimates yield opposite conclusions regarding the benefits of ALMPs. In particular, using OLS implies all ALMPs have negative employment effects, and that classroom training is particularly detrimental to employment. The stark differences between the IV and OLS estimates (even with a rich set of controls) highlight the importance of accounting for selection based on unobserved characteristics of participants when evaluating ALMPs.

Third, we study the dynamics of how classroom training relates to employment at different time horizons. We decompose the classroom-training effects into threat effects, lock-in effects, and post-program effects and make two important findings. First, we find OLS and IV identify similar-sized and negative lock-in effects of classroom training. Hence, OLS and IV do not differ due to different lock-in effects; they differ because job seekers with worse job prospects are those who opt into training.[3] Second, we find the long-run IV estimates are primarily driven by post-program effects rather than threat effects. This finding suggests that classroom training increases employment by providing job seekers with new skills valued in the labor market.[4] This *reskilling* mechanism is consistent with our finding that assignment to classroom training especially increases employment outside the job seekers' original occupations; occupational changes likely require new skills. Furthermore, we find evidence suggesting that the employment effects of classroom training are driven by more successful job applications rather than by intensified job search. The higher success rate of job applications could reflect that classroom training provides skills demanded by employers.

Finally, we examine heterogeneity in the causal effects of training. We find job seekers initially employed in offshoring-exposed occupations – who tend to experience larger and more persistent employment losses – also have much higher gains from classroom training and that they drive the positive effects of these programs. Skill mismatches caused by globalization

[3] The bias is a prospective version of Ashenfelter's dip (Ashenfelter, 1978). Following the theoretical discussions in McCall et al. (2016), this selection bias may arise from optimal training assignments if caseworkers or job seekers take into account the opportunity costs of time spent on training (in terms of foregone earnings) and have latent information about these job options.

[4] This is consistent with van den Berg and Vikström (2022), who derive identification results in a dynamic setting and show in an application that classroom training in Sweden has positive long-run earnings effects. They also emphasize the importance of accounting for the fact that non-treated job seekers tend to have short unemployment durations and more favorable personal characteristics. The dynamic perspective of our approach bears a resemblance to the timing-of-events evaluation method (Abbring and van den Berg, 2003) that has been used to evaluate threat and lock-in program effects using duration models. Crépon and van den Berg (2016) provide a review of this literature with a focus on a dynamic perspective of program effects.

or technological change are an often-cited motivation for training programs (e.g., Braxton and Taska (2023)), and our results suggest globalization-exposed job seekers have higher employment gains and that total employment may be increased by targeting training toward these job seekers.[5] We also explore whether our local average treatment effects (LATE) for instrument compliers are informative for a broader set of job seekers. By the estimation of marginal treatment effects, we find positive effects of classroom training for job seekers with different underlying resistance to training. These findings suggest that our LATE estimate for classroom training is representative of the broader population of job seekers.

We make three contributions to measurement that are critical for our empirical analysis. First, we collect two new administrative data sets: caseworker meeting registrations and individual job plans. The meeting registrations allow us to link job seekers to caseworkers. The job plans are law-mandated logs of caseworker inventions and provide detailed information about all ALMP program assignments in Denmark. Second, we document that many job centers allocate job seekers to caseworkers based on their day of birth (1-31), establishing our source of quasi-random variation. Third, we link our data to registers at Statistics Denmark, allowing us to study how treatment effects vary along a wealth of job-seeker characteristics, such as exposure to offshoring.

We also make a methodological contribution by extending judge IV designs to settings with multiple treatments and heterogeneous treatment effects. We conduct a host of specification checks that allow us to interpret our IV estimates as local average treatment effects (LATE). First, we test the independence assumption by validating that caseworker assignments do not correlate with observable characteristics of job seekers. Second, we verify that caseworker tendencies are highly relevant for training assignments (our first-stage relationships) and that the effects likely are monotone across job seekers.[6] Finally, our identifying assumption is that caseworkers' tendencies for assigning training programs only affect job-seeker outcomes, including any caseworker value added through training assignment. We provide evidence that supports this exclusion restriction, showing that the training tendencies are uncorrelated with other fac-

[5]In Denmark, a stated goal of the classroom-training program is to "solve labor market restructuring and adaptation problems in accordance with the needs on the labor market in a short and a long term perspective." (Danish Ministry of Education, 2021), but classroom training is not targeted toward specific job seeker types. To the best of our knowledge, we are the first to show that training programs particularly benefit globalization-exposed job seekers.

[6]Important for our setting with multiple treatments, we show the program-specific instruments do not have "cross-effects" on different training programs; see Behaghel et al. (2013); Bhuller and Sigstad (2022).

tors potentially driving their value added, such as the experiences, meeting frequencies, and caseloads of caseworkers. Further, our explorations of the mechanisms behind the employment effects indirectly support the exclusion restriction by suggesting *reskilling* is the key mechanism. Indeed, job seeker reskilling is difficult to rationalize with other caseworker behaviors than training assignments.

The empirical ALMP program-evaluation literature is extensive. The paper closest to ours is Hyman (2018), who evaluates the Trade Adjustment Assistance (TAA) program that targets workers displaced by import competition in the US. He also constructs a caseworker-tendency instrument and finds significantly positive earnings effects for TAA-approved workers. We provide broader evidence for the impact of ALMP programs as we compare classroom training with on-the-job training programs, we decompose the effects into threat, lock-in, and post-program effects, and we examine heterogeneity across subpopulations including globalization-exposed job seekers. In addition, our instrument relies directly on randomness generated by the day-of-birth rules in allocating unemployed job seekers to caseworkers. A growing strand of the literature uses randomized controlled trials (RCTs) that also have the potential to address the identification challenges in program evaluations outlined above. Our natural-experiment strategy complements this literature in several ways. The natural setting of our study mitigates concerns about external validity. Also, RCTs are often limited in scale due to resource requirements, whereas we exploit the extent of our data to estimate effects by subpopulations. Further, we decompose the dynamics of our IV estimates and estimate marginal treatment effects to extrapolate to broader populations.[7]

The remainder of the paper is structured as follows. Sections 2 and 3 first describe our institutional setting and data. Section 4 presents a conceptual framework and derives our identification strategy based on caseworker tendencies. Section 5 performs instrument diagnostics. Section 6 presents our main empirical results, estimating how classroom training affects labor market outcomes. Section 7 decomposes the impacts into ex-ante, lock-in, and post-program effects, and investigates occupational mobility. Section 8 studies heterogeneity in the treatment effects across training programs and worker types. Section 9 concludes.

[7]In addition, RCTs sometimes include many initiatives that make isolating the effects of particular programs difficult. For example, the Danish labor market authorities have adopted a systematic use of RCTs, and in some experiments, the treatment consists of a mix of job-search assistance programs, caseworker meetings, on-the-job training, and classroom training, see, for example, Graversen and van Ours (2008), Vikström et al. (2013), and Gautier et al. (2018). McCall et al. (2016) provide a discussion of challenges faced by RCTs.

2 Institutional Setting

In this section, we describe key features of the unemployment insurance system in Denmark. We mainly focus on caseworkers and the assignment of training programs, and we document how job seekers are quasi-randomly allocated to caseworkers due to a day-of-birth rule.

2.1 Unemployment Insurance in Denmark

In Denmark, unemployed job seekers may receive unemployment insurance (UI) benefits for up to two years. The UI benefits are financed by membership contributions to UI funds, which the government subsidizes. The economic incentive to claim UI benefits is high: the UI benefits cover 90% of prior monthly earnings up to 3,075 USD.

The public employment services in Denmark build on a *right and duty* principle (Kreiner and Svarer, 2022). That is, unemployed job seekers have the right to receive UI benefits but also a duty to live up to specific requirements. The requirements involve regular meetings with a caseworker at the local job center, active job search, and participation in training programs assigned by the caseworker. If the job seeker does not comply with the caseworker assignments, she will lose her right to receive UI benefits.

Table 1 shows the average job seeker meets with her caseworker for the first time in week five of the UI spell. The meetings typically last 30-45 minutes, and the average job seeker has four meetings over her unemployment spell. It is during these caseworker meetings that job-search monitoring, counseling, and sanctioning take place. On top of this job-search assistance, some job seekers are also assigned to training programs.

Table 1: Timing and Duration of Activities Assigned within First 12 Months

	(1) Assignment rate (%)	(2) Timing (week no.)	(3) Total duration (days)
Caseworker Meetings	100	5	1
Classroom Training	39	18	52
On-the-Job Training	25	21	88

Notes: This table shows the timing and duration of activities assigned within the first 12 months after job loss. Column (1) reports the share of job seekers who have at least one caseworker meeting (row 1), and who are assigned to a training program within the first 12 months after job loss (row 2-3). Column (2) reports the average timing of the *first* assigned activity measured relative to job loss. Column (3) reports the total duration of *all* assigned activities. Row 2-3 report the total number of days job seekers are supposed to spend on classroom and on-the-job training. For comparison, row 1 reports our best estimate of how many days the average job seeker spends with the caseworker over the unemployment spell. It is based on (i) the observation that jobseekers on average have 4 caseworker meetings over the unemployment spell, (ii) each of which lasts about 30-45 minutes, and (iii) presuming job seekers spend at least five hours per day at training facilities.

2.2 Training Programs

The caseworker must prepare a *job plan* for the job seeker, specifying their assignments to training programs. Caseworkers have two classes of training programs at their disposal: *on-the-job training* and *classroom training*.

On-the-job training includes internships and wage subsidies for employment at a private or public firm for a pre-specified time period. An objective of these programs is to provide the job seeker with some hands-on workplace experience.

Classroom training includes ordinary education, vocational training, and job-search-assistance courses. *Ordinary education* includes basic coursework, for example, language training or primary school courses, and more advanced coursework, for example, a university class. *Vocational training* courses target specific occupations or industries, such as a forklift certificate or a pc course.[8] Finally, *job-search-assistance courses* include dedicated courses on CV or job-application writing, as well as courses targeting more "soft skills".[9,10] Across classroom training programs, the primary objective is to teach the job seeker new skills.

Table 2 reports the share of job seekers in our analysis sample who are assigned to training programs within the first 12 months of their UI spell. The table shows that 39% of all job seekers are assigned to classroom training, and 25% are assigned to on-the-job training. The two training programs are not mutually exclusive: 13% of job seekers are assigned to both programs. Finally, the table shows 48% of job seekers are in "passive UI"; they are neither assigned to classroom training nor to on-the-job training. Hence, while all job seekers in the sample receive job-search assistance facilitated through caseworker meetings, only about half the jobseekers are also assigned to some training program.

[8]A stated goal of the vocational training program in Denmark is to "solve labor market restructuring and adaptation problems in accordance with the needs on the labor market in a short and a long term perspective." (Danish Ministry of Education, 2021).

[9]We distinguish between job-search assistance provided through caseworker meetings, which all job seekers receive, and job search assistance provided in dedicated classroom courses if assigned.

[10]For example, "Job search" and "CV courses" as well as "business coaching", "preparation for education" and "information about the Danish labor maket" are among the job-search-assistance courses used by the Danish job centers (VIVE, 2022; Slotsholmen, 2013).

Table 2: Assignment to Classroom Training and On-The-Job Training

	Percent of Job Seekers	
	Classroom Training = 0	Classroom Training = 1
On-the-Job Training = 0	48	27
On-the-Job Training = 1	13	12

Notes: This table shows the share of job seekers in the analysis sample who are assigned to a given type of training program within the first 12 months after the UI-spell start.

Since the training programs have different objectives, “skills” vs. “experience”, they are likely *not* close substitutes for the caseworker. For example, newly graduated job seekers may have up-to-date skills, but lack labor market experience, and so would benefit from on-the-job training. Other more experienced job seekers may have outdated skills, and so may benefit from classroom training. Hence, the caseworker’s decision to assign a given training program (if any) is likely based on the needs of the individual job seeker and detached from the decision about the other training program. Indeed, Appendix Figure D.1 shows that caseworkers’ classroom training and on-the-job-training tendencies are uncorrelated. This supports the view that assignment to classroom training and on-the-job training are two separate decisions for the caseworker.

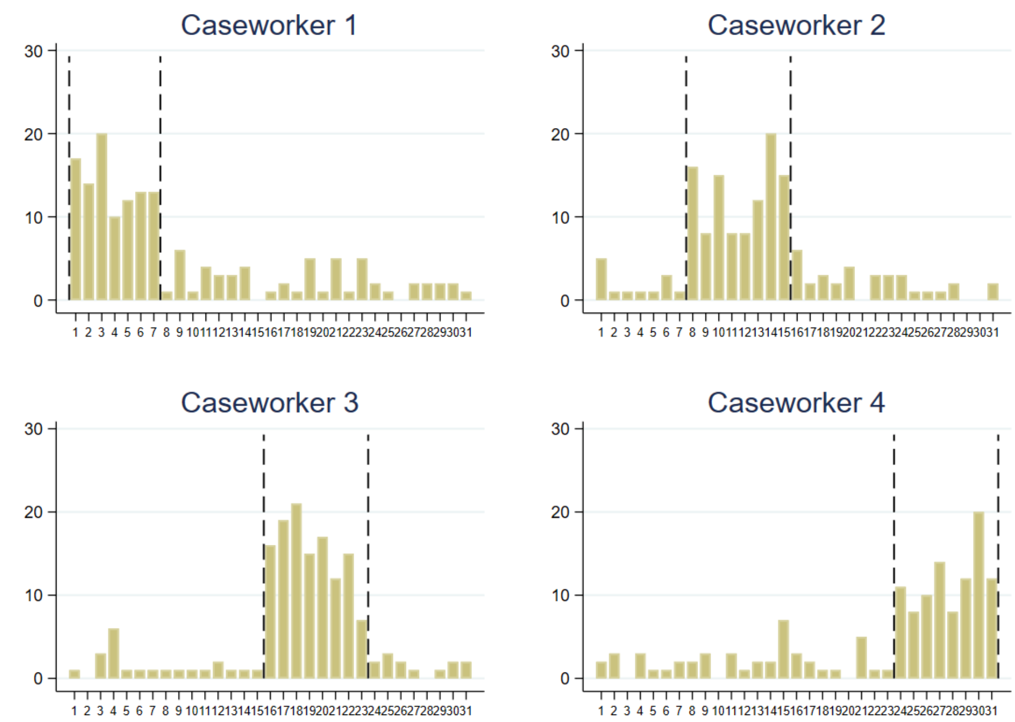
Table 1 compares the timing and duration of caseworker meetings and training programs. The average trainee spends 52 and 88 days in total on classroom and on-the-job training, respectively. In comparison, these job seekers spend the equivalent of almost a full day with their caseworker over their unemployment spell. Put differently, an essential way caseworkers affect job-seeker activities is by assigning training programs.

2.3 Assignment of Job Seekers to Caseworkers

The public employment services in Denmark are organized by local job centers, which allocate job seekers to caseworkers. Our identification strategy exploits that many job centers assign job seekers to caseworkers based on their monthly day-of-birth (1-31).

Figure 1 illustrates this allocation for a representative job center, plotting the day-of-birth distributions of job seekers allocated to different caseworkers in the job center. The figure reveals two features. First, job centers use *blocks* of days to allocate job seekers: caseworker 1 primarily handles job seekers born on the $1^{st} - 7^{th}$ of the month, caseworker 2 handles the $8^{th} - 15^{th}$, and so forth. The block structures enable us to easily detect whether a job center uses a day-of-birth allocation rule. Second, job centers occasionally deviate from the day-of-birth allocation rule. For example, caseworker 1 is also allocated a few job seekers outside the $1^{st} - 7^{th}$ interval. Deviations could happen for exogenous reasons, for example, caseworker illness, or endogenous reasons, for example, match effects. To circumvent these endogenous deviations, we base our identification strategy on the *day-of-birth-predicted* caseworkers, thus exclusively using the quasi-random variation arising from job seeker birthdays.

Figure 1: Allocation of Job seekers to Caseworkers over Day of Birth



Notes: This figure illustrates the day-of-birth distributions of job seekers allocated to different caseworkers in a representative job center. Due to data confidentially, we simulate a job center with a compliance rate of 58%, the median compliance in our analysis sample.

3 Data

Our empirical analysis relies on two new registers from the Danish Agency for Labor Market and Recruitment (STAR). The registers record meetings between caseworkers and job seekers and the assignments to training programs. We link these registers to several administrative registers at Statistics Denmark using unique person identifiers, providing detailed information on the characteristics and labor market outcomes of job seekers.

3.1 Data Sources

Caseworker Meetings

Our data on caseworker meetings record the date, time, and type of all meetings in the Danish job centers from 2012 to 2019. The dataset also includes identifiers for the job seeker and caseworker participating in each meeting. We use this dataset to create a linked job-seeker–caseworker data set.

First, we identify job seekers who became unemployed from 2012 to 2018. We link each of the job seekers to the caseworker who participated in the job seekers' first face-to-face meeting. We call this caseworker her *realized caseworker*.[11]

Second, we document the use of day-of-birth allocation rules in all 94 job centers in Denmark over time. Online Appendix OA1.4 details our procedure, which proceeds in two steps. First, we identify the job-center units in which job seekers are allocated to caseworkers.[12] Second, we link each caseworker to the birthday of their job seekers. For each job-center unit and year, we assign a *predicted caseworker* to each day in the month (1-31), defined as the caseworker with the most job seekers born on that day. This procedure is similar to Cederlöf et al. (2021), who identify the day-of-birth predicted caseworker for Swedish job seekers. Appendix Table B.1 shows we are able to link 90% of all job seekers (84% of all UI spells) who had at least one caseworker meeting to a *realized* and a *day-of-birth-predicted* caseworker.

Training Assignments

We measure the assignment of job seekers to classroom training and on-the-job training programs using a register on the individual *job plans* prepared by caseworkers. As noted in Section 2, the job plans are law-mandated registrations of all training assignments throughout the UI spells of job seekers. The job plans have two advantages over existing registers.

First, the job plans include *all* training assignments, regardless of whether the job seeker ends up participating in the programs. In the Danish context, the standard dataset for research on ALMPs is the *Danish Register for Evaluation of Marginalization* (DREAM). However, DREAM only records actual participation in programs. Furthermore, the data source for

[11]See Online Appendices OA1.2 and OA1.3 for details on how we identify the first face-to-face meeting in the UI spell and implement a crosswalk between caseworker identifiers over time.

[12]Some job centers are organized into smaller units (e.g., according to job seeker age), wherein caseworkers are allocated to job seekers. See further details in Online Appendix OA1.4.

DREAM is, in fact, the job-plan registrations. As discussed in McCall et al. (2016), measuring *assignment*, as opposed to only participation, may be more relevant for policy because "it corresponds to what policies, in most contexts, can actually do."

Second, the job plans record the exact *timing* of the training assignments, including the start and end dates of the programs. By linking the timing of training to data on when job seekers enter and exit unemployment, we can decompose the dynamic employment effects of training into threat, lock-in, and post-program effects; see Section 7.1.

Online Appendix OA1.8 provides a validation of the job plans and our two interventions of interest: assignment to classroom training and on-the-job training. Two insights merit note. First, a substantial part of the *classroom-training* assignments likely represents actual education (ordinary education or vocational training). We find that 44% of classroom training assignees are *registered* as enrolled in education within the first 12 months of unemployment. This corresponds to 50% of the total assigned classroom training days. Given that some assignments never lead to enrollment, this is likely a lower bound of the education content of classroom training. Second, the *on-the-job-training* assignments include both public and private internships and wage subsidies. For example, one-third of the internships are public sector programs. This observation is important because the existing literature has found very different employment effects of public and private programs (Card et al., 2010).

Labor Market Outcomes

Our data on employment outcomes come from the *Register for Employees* (BFL), recording the work hours, earnings, and occupational codes of all job spells in Denmark from 2008 to 2019.[13] A notable feature of the BFL register is that it contains a high-quality and continuous measure of employment hours, which stems directly from third-party reports to the Tax Authorities. We combine the register with DREAM to measure non-supported employment; see details in Online Appendix OA1.5.

Job Seeker Characteristics

Our data on *demographics* come from the *Population Register* (BEF), recording the gender, age, and country of origin for all Danish residents. We obtain information on job seekers' *education*

[13]We cap the outcome data at this horizon to focus on labor market outcomes before the onset of Covid-19 and related lockdowns in 2020

from the *Education Register* (UDDA). Finally, we measure the *labor market histories* of job seekers using DREAM, BFL, and the *Employment Classification Module* (AKM). The DREAM register records public transfers, including UI benefits and education subsidies. The BFL and AKM registers provide information on past employment of the job seekers, including earnings, hours, industry, and occupations.

3.2 Analysis Sample

We base our analysis sample on all job seekers (i) who became unemployed between 2012 and 2018 and (ii) who had at least one meeting with a caseworker from the local job center. Using our linked job-seeker–caseworker data, we obtain information about the job seekers' *realized* and *day-of-birth predicted* caseworker. Using the job plan data, we obtain information about the jobseekers' assignments to classroom and on-the-job training programs.

We apply five sample restrictions to support our identification strategy. First, based on a visual inspection of the day-of-birth allocation rules, we drop job-center-unit-years that do not use a clear (block) structure for the allocation of birthdays, or where the compliance to the block structure is very low (see Online Appendix OA1.4). Appendix Table B.1 shows the number of unique job centers drops from 94 to 51, while the day-of-birth compliance rate increases from 42% to 52% when we impose this restriction. Second, we exclude from the analysis sample all job seekers of non-western origin because an institutional feature of the Danish immigration system makes non-western job seekers over-represented among job seekers born on the 1st of the month.[14] Third, we exclude job seekers for whom we cannot identify their previous occupation. We make this restriction because we will distinguish job seekers by the offshorability of their previous occupation in Section 8.2. Fourth, to implement our identification strategy based on random caseworker allocations, we require that all job-center-unit-year cells contain at least two predicted caseworkers. Fifth, to estimate training tendencies reliably, we require that each caseworker in the sample must have at least 50 assigned job seekers over the sample period.

Appendix Table B.1 shows our analysis sample captures 51 of all job centers and about 20% of all job seekers who initiated a UI-spell from 2012-2018 and had at least one caseworker

[14] An immigrant who arrives in Denmark without a birth certificate is automatically assigned January 1st as their birthday. As a result, non-western make up 10% of job seekers born on the 1st of the month while only constituting 3.5% of the job seeker sample.

meeting (the linked job-seeker–caseworker data). Appendix Figure A.1 shows these job centers are spread out across Denmark. We develop our caseworker-tendency instruments based on this *full* sample of job seekers. In order to study the dynamic effects of assignment to training programs, we base all estimations on a *balanced* sample of job seekers for whom we observe labor market outcomes in the first two years after job loss.[15] In Section E.6, we show that we obtain similar results if we used the full (but unbalanced) sample of job seekers.

4 Identification Strategy

The goal of this paper is to estimate the effects of assignment to training programs, as opposed to passive UI, on the employment of job seekers. Identifying these effects is challenging because job seekers may opt into training programs based on their job opportunities. This selection bias could come through self-selection of the job seekers or from the caseworker assignments. For example, a job seeker with a pending job offer may be more resistant to start training than a job seeker without immediate job prospects. Similarly, caseworkers may be hesitant to assign a long-term training course to job seekers who already have good job options. This type of information may be revealed to the caseworker during meetings, but it is unobserved to the econometrician. Furthermore, whether controlling for observables of the job seekers eliminates this selection bias is unclear because job seekers with similar work *histories* could face different job *prospects* that are not recorded in our administrative data.

To overcome this identification challenge, we follow the judge IV literature and develop caseworker-tendency instruments for assignment to training programs. Our setting departs from the canonical judge IV setup in two aspects. First, caseworkers may assign job seekers to *multiple* training programs. Second, job seekers are allocated to caseworkers based on an *observed rule* (day of birth).

In Appendix C, we develop a generalized Roy model that addresses these methodological novelties. In the model, caseworkers differ in their preferences for the training programs but rank individual job seekers similarly in their resistance to participate in each program.[16,17]

[15] Since we observe outcomes up until 2019, we essentially restrict the sample to job seekers who became unemployed from 2012 to 2017.

[16] For example, a job seeker with a pending job offer may be more resistant to start training than a job seeker with no immediate job prospects.

[17] The role of caseworkers' preferences is motivated by our qualitative interviews at the job centers, during which a caseworker, for example, ascribed differences in training tendencies to reflect differences in "values" ("*værdisæt*"

First, we show our setting collapses to the canonical single-treatment case once we compare caseworkers with similar tendencies for the other training programs. Second, we state the identifying assumptions explicitly in terms of the day-of-birth-predicted caseworkers.

In particular, we instrument the assignment of job seeker i to training program k with the training program assignment rate among other job seekers with the same day-of-birth predicted caseworker. Let $c(b)$ denote the rule that allocates job seekers with day of birth $b \in \{1, 31\}$ to caseworkers c, and let the indicator D_{ki} equal one if job seeker i is assigned to program k within the first 12 months after job loss. Our training tendency instruments are defined as follows:

$$Z_{ki} = \frac{1}{(J(i)-1)} \sum_{j \neq i} \mathbf{1}[c(b_j) = c(b_i)] \times D_{kj}, \quad J(i) = \sum_{j} \mathbf{1}[c(b_j) = c(b_i)], \tag{1}$$

where $J(i)$ is the number of job seekers with the same day-of-birth predicted caseworker as job seeker i. We denote caseworkers with a low Z_k as k-*restrained* and caseworkers with a high Z_k as k-*inclined*.

The instruments must satisfy the usual assumptions of relevance, independence, exclusion, and monotonicity. In the canonical binary-treatment case, monotonicity requires that the instrument shift all job seekers toward or away from the treatment in consideration. However, because job seekers face *multiple* training options, identification in our setting requires an *extended* monotonicity assumption about how instruments affect multiple training programs. In particular, the assignment of job seeker i to training k must solely depend on the k-tendency of her caseworker. Hence, comparing two otherwise similar caseworkers, a more k-inclined caseworker will shift all job seekers toward training program k but not alter the participation in other programs l. This *extended monotonicity* assumption plays a key role in IV analysis with multiple treatments (Behaghel et al., 2013; Lee and Salanié, 2018, 2020; Bhuller and Sigstad, 2022). See theory and further discussion in Appendix C.

The theoretical extension in Appendix C motivates instrument diagnostics in Section 5 and facilitates the estimation of marginal treatment effects in Section 8.3. The theory highlights the importance of controlling non-parametrically for the other training-program instruments (Blandhol et al., 2022). In practice, we first estimate a standard two-stage least squares (TSLS) specification to facilitate comparison with how prior studies have handled multiple treatments in judge IV setups (Autor et al., 2015; Bhuller et al., 2020). In a second step, we show robust-

in Danish) of caseworkers.

ness to estimating the specification around an evaluation point for the other training-program instruments (Mountjoy, 2022).

We estimate the dynamic effects of assignment to classroom training on the employment of jobseekers. Let Y_{it} denote the employment of job seeker i in period t relative to the start of the unemployment spell. The TSLS specification reads

$$D_{ki} = \delta_{q(i)k} + \delta_{k1} Z_{1i} + \delta_{k2} Z_{2i} + \varepsilon_{1i} \tag{2}$$

$$Y_{it} = \beta_{q(i)t} + \beta_{1t} D_{1i} + \beta_{2t} D_{2i} + \varepsilon_{2it}, \tag{3}$$

where $q(i)$ are job-center-unit-start-year combinations, the units wherein our randomization takes place. Hence, we compare job seekers from the same job-center unit and year who, due to their day-of-birth, receive different treatments. We cluster standard errors on the job seeker and predicted caseworker levels.

We abstract from interaction terms between classroom and on-the-job training in Equations (2)-(3) since our focus is on the causal effects of the separate training interventions. Furthermore, we do not have the statistical power to identify the interaction terms.[18] Leaving out interaction effects implies the β_{1t} coefficient represents the effect of shifting compliers into classroom training from either passive UI or on-the-job training only. Appendix E.2 shows that our results are robust to focusing exclusively on compliers shifted from the passive margin and into classroom training.

Importantly, our identification strategy does *not* preclude that caseworkers differ in their "value added" to the outcomes of job seekers as documented by, e.g., Behncke et al. (2010) and Cederlöf et al. (2021). Instead, our identifying assumption is that caseworkers' tendencies for assigning training programs only affect job seeker outcomes, including any caseworker value added, through training assignments. In Section 5.4, we provide evidence that supports this exclusion restriction, showing that the training tendencies are uncorrelated with other factors potentially driving their value added, such as the experiences, meeting frequencies, and caseloads of caseworkers.

[18] To estimate interaction effects, we would need an additional first-stage equation with an interaction of the indicators (tendencies) for classroom and on-the-job training on the LHS (RHS):

$$D_{1i} D_{2i} = \delta_{q(i)} + \delta_1 Z_{1i} + \delta_2 Z_{2i} + \delta_3 Z_{1i} Z_{2i}.$$

When estimating this equation, the coefficient on $Z_1 Z_2$ is barely significant (p-value of 0.067).

5 Instrument Diagnostics

In this section, we assess our caseworker-tendency instruments. We provide evidence that the instruments satisfy the relevance, independence, exclusion, and monotonicity conditions for interpreting our IV estimates as local average treatment effects (LATE).

Our identification strategy makes two departures from earlier judge-IV designs. First, we directly exploit the *source* of quasi-random assignment coming from the day-of-birth allocation of job seekers to caseworkers.[19] We find exploiting the source of randomization is crucial for instrument independence in our context. Second, we have a setting with multiple treatments, which motivates new instrument diagnostics. In particular, we show the program-specific tendencies do *not* alter the assignment of other programs (Behaghel et al., 2013; Bhuller and Sigstad, 2022).

Finally, we note that our exclusion restriction is milder than the *strict* version proposed by Frandsen et al. (2023) because we allow caseworkers to differ in value-added unrelated to training tendencies; see Appendix C.1 for theoretical discussions. Indeed, our instrument diagnostics follow the Frandsen et al. (2023) guidelines for testing average exclusion and average monotonicity.

5.1 Relevance

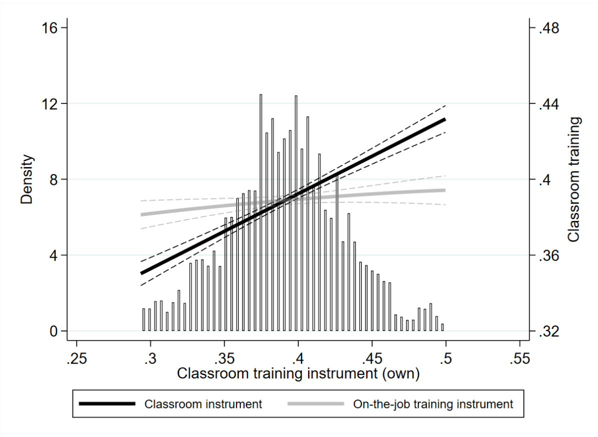
Figure 2 shows how our caseworker-tendency instruments affect the assignment of job seekers to training programs after initial job loss. The figure delivers two takeaways.

First, the caseworkers' training-program tendencies strongly affect the job seekers' assignment to training programs. Figures 2a-2b corroborate this conclusion by combining the first-stage relationships with the distribution of caseworker training tendencies. For example, shifting from the most classroom-training *restrained* to the most classroom-training *inclined* caseworker within a job-center-unit-year corresponds to an 8 percentage point increase in the probability of being assigned to classroom training within the first 12 months. This increase is a 20% increase relative to the mean classroom-training assignment rate.[20] As Figures 2c-2d show, most of these assignments occur in the first quarter after the initial job loss. In Sections 7.1 and 6.1, we study how the assignments affect the dynamics of training and employment of job seekers.

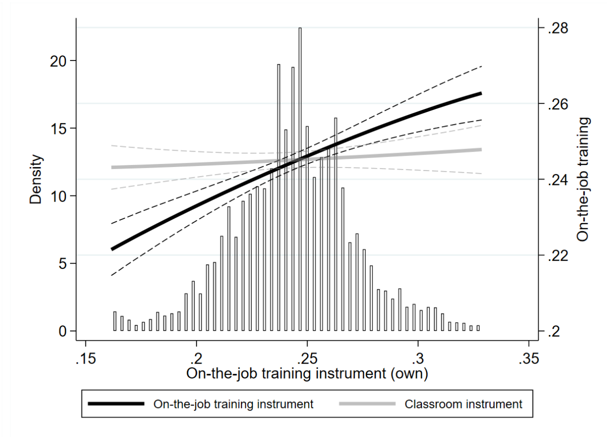
[19] Existing studies either rely on the law mandating some random allocation or anecdotal evidence on why the realized allocation may be quasi-random.

[20] Appendix Table D.1 shows the program-specific instruments have power for each training program.

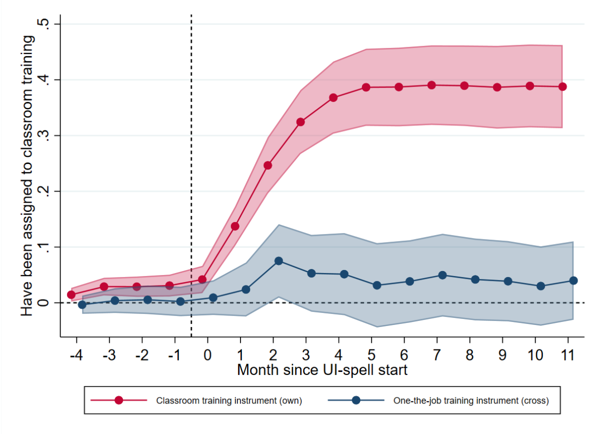
Figure 2: Assignment to Training Programs



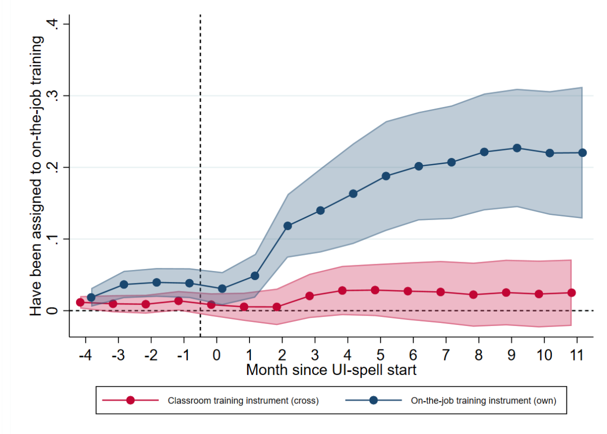
(a) Classroom Training



(b) On-The-Job Training



(c) Classroom Training



(d) On-The-Job Training

Notes: This figure plots the first stage estimates for assignment to training programs. Panels (a) and (b) represent the first stage of assignment to classroom training and on-the-job training within the first 12 months on the two caseworker-tendency instruments. The bars represent the distribution of the own-instrument demeaned by job-center-unit-year fixed effects and the cross-instrument (excluding top and bottom 1%). The black line represents the coefficients from a local linear regression of the training program on the own-instrument, both demeaned by job-center-unit-year fixed effects and the cross-instrument. The gray line represents the coefficients from a local linear regression of the training program on the cross-instrument, both demeaned by job-center-unit-year fixed effects and the own-instrument. The local linear regressions are based on an Epanechnikov kernel (with bandwidth 0.1). Panels (c) and (d) represent a dynamic version of the first stage. They represent the first stage of having been assigned to classroom training and on-the-job training in or prior to a given month on the two instruments. Shaded areas represent 95% confidence intervals.

Second, the program-specific instruments do *not* alter the assignment to other programs. For example, Figure 2a shows the assignment to classroom training only is affected by the caseworkers' classroom-training tendencies, not by their on-the-job-training tendencies. This absence of "cross effects" is not surprising: Theoretically, the two programs are unlikely to be close substitutes (see discussion in Section 2), and empirically, caseworkers' training tendencies are uncorrelated across programs (see Appendix Figure D.1). As discussed in Section 4, the absence of "cross-effects" is crucial for identification in our setting with multiple treatments.

In particular, it allows us to collapse our setting to the canonical single-treatment case by comparing caseworkers with similar tendencies for the other training programs. Importantly, this requires us to find otherwise similar caseworkers who differ in their tendency for each training program. Fortunately, as already mentioned, caseworker tendencies are indeed uncorrelated across programs in our setting, see Appendix Figure D.1.

5.2 Independence

In Table 3, we test the independence of the caseworker-tendency instruments. The table is based on the following logic: if job seekers are allocated to caseworkers in a quasi-random fashion, we should not be able to predict caseworkers' training tendencies based on the characteristics of job seekers measured before job loss.

The independence test yields three takeaways. First, the assignment of job seekers to training programs is highly endogenous (Columns (1) and (2)), confirming the common finding that job seekers select into training. Second, instrumenting with the training tendencies of job seekers' *realized* caseworker does not solve the endogeneity issue (Columns (3) and (4)), because job centers deviate from the day-of-birth allocation rule in a non-random fashion. Finally, using the *day-of-birth-predicted* caseworkers, caseworker training tendencies and job seekers' characteristics are uncorrelated (Columns (5) and (6)). The evidence suggests our day-of-birth-predicted training tendencies are indeed exogenous shifters to the assignment of job seekers to training programs. These findings highlight the importance of using the explicit quasi-random variation arising from the day-of-birth allocation rules.

Table 3: Testing for Random Assignment of Job Seekers to Caseworkers

	Actual assignment		Realized Caseworker		Predicted Caseworker		Covariates	
	(1) Classroom training	(2) On-the-job training	(3) Classroom training	(4) On-the-job training	(5) Classroom training	(6) On-the-job training	(7) mean	(8) sd
Demographics								
Age	0.002***	0.000	-0.001**	-0.001***	0.000*	-0.000	41.915	11.980
Male	0.014**	-0.074***	-0.001	0.001	-0.001	-0.002	0.480	0.500
Immigrant	0.128***	0.017	0.005	-0.003	-0.001	-0.001	0.043	0.202
Descendant	0.090	-0.081	0.031	-0.049	0.004	0.002	0.002	0.041
Married	-0.066***	-0.022***	-0.006*	-0.007*	-0.005***	-0.001	0.423	0.494
Number of children	-0.014***	-0.019***	-0.003*	-0.004**	-0.001	-0.002	0.757	1.011
Education								
Primary	0.014	0.115	0.064	0.041	0.043	0.001	0.001	0.037
Lower secondary	-0.056**	-0.024	0.015	0.004	0.011	0.004	0.197	0.398
Upper secondary	-0.071***	-0.029	0.025	0.008	0.010	0.001	0.525	0.499
Short cycle tertiary	-0.026	-0.033	0.028	0.011	0.010	0.008	0.051	0.221
Bachelor	-0.123***	-0.037	0.012	-0.004	0.013	0.007	0.154	0.361
Master	-0.094***	0.009	0.027	-0.023	0.005	-0.004	0.058	0.235
Doctoral	-0.187***	0.008	0.041	0.029	0.027	0.024	0.002	0.048
Labor market history								
UI-benefits in year t-1	-0.064***	-0.083***	-0.005	0.007	-0.002	0.005**	0.448	0.497
UI-benefits in year t-2	-0.074***	-0.059***	-0.008	0.002	0.001	-0.004	0.458	0.498
Cash benefits previous 5 years	0.098***	0.046***	0.004	0.021**	-0.001	0.001	0.076	0.265
Parental leave in year t-1	-0.018*	0.008	-0.020***	-0.034***	0.002	0.002	0.073	0.260
Education subsidy in year t-1	-0.099***	-0.022*	-0.004	-0.016	-0.000	0.003	0.100	0.300
Public transfers in year t-1	0.096***	0.109***	0.005	0.034***	0.001	0.001	0.364	0.481
Employed in year t-1	-0.086***	-0.139***	-0.005	0.000	-0.003	0.002	0.889	0.315
Employed in year t-2	-0.057***	-0.027**	-0.020**	-0.031***	-0.002	-0.003	0.883	0.322
Hours in year t-1	-0.001***	-0.001***	-0.000***	-0.001***	-0.000	0.000	93.101	57.831
Hours in year t-2	0.001***	0.001***	0.000	0.000	-0.000	-0.000	94.650	59.802
Earnings in year t-1 (1,000 DKK)	-0.001	-0.003***	0.000	0.000	0.000	0.000	17.403	13.937
Earnings in year t-2 (1,000 DKK)	-0.003***	-0.005***	-0.000	-0.001	0.000	0.000	16.804	13.346
Previous industry								
C Manufacturing	0.083***	0.018*	0.004	-0.029***	0.001	0.001	0.117	0.322
F Construction	-0.162***	-0.131***	-0.004	-0.024***	-0.000	0.001	0.102	0.302
G Wholesale and repair	0.023**	0.091***	0.008	-0.003	0.001	0.001	0.134	0.341
H Transportation and storage	0.050***	-0.032**	0.005	-0.003	-0.001	-0.002	0.039	0.194
I Accomodation and food service	-0.055***	0.009	-0.013	-0.039***	0.001	0.006	0.037	0.189
N Administrative and support service	0.031**	-0.031***	0.014**	-0.022***	0.001	-0.001	0.085	0.279
P Education	-0.021*	-0.010	0.001	0.016*	-0.005	-0.002	0.055	0.229
Q Human health and social work	-0.056***	-0.022**	0.004	0.005	-0.003	0.001	0.186	0.389
Previous occupation								
High risk of offshoring	0.071***	0.070***	0.008	0.013**	-0.004	0.007**	0.250	0.433
Professionals	-0.050***	-0.045**	0.015	0.006	0.010	-0.006	0.142	0.349
Technicians and associate professionals	-0.042**	-0.055***	0.015	0.014	0.008	-0.013*	0.088	0.284
Clerical support workers	-0.040*	-0.039**	0.012	0.009	0.011	-0.006	0.099	0.299
Service and sales workers	-0.059***	-0.013	0.013	0.006	0.009	-0.005	0.240	0.427
Skilled agricultural, forestry amd fishery workers	-0.167***	-0.161***	0.005	-0.045***	0.020**	-0.017*	0.022	0.148
Craft and related trade workers	-0.106***	-0.084***	0.004	-0.022*	0.007	-0.002	0.163	0.369
Plant and machine operators, and assemblers	-0.004	-0.045**	0.012	-0.007	0.010	-0.005	0.090	0.286
Elementary occupations	-0.075***	-0.063***	-0.005	-0.018	0.006	-0.010	0.133	0.339
UI-fund								
Academics Association	0.099***	-0.028***	0.032***	0.017	0.004	-0.000	0.095	0.293
Danish Trade Union Association	0.041***	0.025***	-0.003	-0.019***	0.004	-0.003	0.713	0.453
Obs	167,222	167,222	166,666	166,666	167,222	167,222	167,222	
Number of FE's	189	189	189	189	189	189		
F-stat	76.247	87.495	2.628	6.402	1.167	1.106		
P-value	0.000	0.000	0.000	0.000	0.219	0.301		

Notes: This table implements a randomization test of the caseworker-tendency instruments. Columns (1) and (2) regress the assignment to training programs on job seekers' covariates. Columns (3) and (4) regress the training tendency of job seekers' *realized* caseworker on the job seekers' covariates. Columns (5) and (6) regress the training tendency of job seekers' *day-of-birth-predicted* caseworker (defined in Equation (1)) on the job seekers' covariates. The training tendencies are leave-out means. To ease comparison across columns, the dependent variable has been standardized to have mean zero and standard deviation one. Education levels are defined according to the 9 sections of the 1-digit ISCED classification. The omitted category is "Early childhood education". Previous industries are defined as the 21 sections of the NACE level-1 classification. Only dummies for the eight largest industries are included in the regression. Previous occupations are defined as the 10 sections of the 1-digit ISCO08 classification. The omitted category is "Managers". "High risk of offshoring" is a dummy indicating the offshoring risk of the job seeker's previous occupation; see details in Section 8.2. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job seeker level, but only significance levels are reported: *$p<0.10$ ** $p<0.05$ *** $p<0.01$.

5.3 Monotonicity

The monotonicity assumption states that a job seeker assigned to training by a training-restrained caseworker also should be assigned by a training-inclined caseworker. Furthermore, in our setting with multiple treatments, monotonicity requires that the program-specific caseworker training tendencies do not alter the assignment to other programs. Behaghel et al. (2013) call this assumption *extended monotonicity*; see Appendix C for discussions. In other words, we require that the "own-instruments" are significant and the "cross-instruments" are insignificant among all subsamples of job seekers.

To test the monotonicity assumption, we adapt two specification tests from Bhuller et al. (2020) to our setting with multiple treatments. First, in Appendix Tables D.4 and D.5, we split the job seekers into quartiles based on their propensities for assignment to training. We then show that the "own-instruments" are significant and the "cross-instruments" are insignificant in all quartiles.

Second, we assess whether caseworkers who are more training-inclined toward one job seeker are also more training-inclined toward other job seekers. To test this, we rely on the same quartiles but now measure caseworker training tendencies using job seekers from *other* subgroups. Appendix Tables D.6 and D.7 show that, in all subgroups, the coefficient on the "reversed" own-instrument is positive, whereas the "reversed" cross-instrument is insignificant (except in one subgroup). Taken together, these results support the monotonicity assumption of our identification strategy.

5.4 Exclusion

To interpret our IV estimates as treatment effects of training, we require the *exclusion restriction* that caseworker training tendencies only affect job-seeker outcomes through the assignment to training. An obvious threat to the exclusion restriction is that caseworkers serve multiple purposes; aside from assigning training programs, caseworkers meet with job seekers and give advice on their job search. Therefore, one concern may be that caseworkers who are very training-inclined also meet more frequently with their job seekers or provide them with better job-search advice. We conduct two tests to address these concerns. We focus on the classroom training instrument in this section and refer to appendix D.3, where we perform the same tests for the on-the-job training instrument.

First, we test whether caseworkers' classroom-training tendencies correlate with proxies for the quality or quantity of their *job-search advice*. For example, the quality of advice could depend on the caseworker's experience, and the quantity of advice could depend on her caseload. Following this logic, Table 4 shows caseworkers' classroom-training tendencies are uncorrelated with their experience and caseloads.

Table 4: Experience, Caseload Size, and Classroom-Training Tendency

	Summary			Classroom Training Tendency	
	obs	mean	sd	coef×100	se×100
Caseload size					
- Meetings/year	576	430.6	235.6	0.0027**	(0.0012)
- Assigned jobseekers/year	576	117.7	64.1	0.0055	(0.0038)
Experience					
- Years	576	1.2	1.2	0.0374	(0.2164)
- Meetings	576	868.9	783.8	0.0003	(0.0003)
- Jobseeker assignments	576	255.0	237.2	0.0008	(0.0011)
Obs	576	1221		1221	

Notes: This table shows correlations between caseworkers' classroom-training tendencies and caseworkers' characteristics, e.g. their experience and caseload size, Columns (1)-(3) report the number of unique caseworkers along with the mean and standard deviation of experience and caseload size across caseworkers and years. Columns (4)-(5) report the coefficient and standard errors from a regression of caseworker experience or caseload size in a given year on the caseworker's classroom-training tendency in the same year, while controlling for job-center-unit-year fixed effects as well as the caseworker's on-the-job-training tendency. Standard errors (in parentheses) are two-way clustered on the predicted caseworker and job seeker level. *p<0.10 ** p<0.05 *** p<0.01.

Second, we assess whether more classroom-training-inclined caseworkers *meet* earlier or more frequently with the job seekers. Panel A of Table 5 shows the caseworker's classroom-training tendency is unable to predict the timing of the first meeting as well as the frequency of meetings with the job seeker. That is, job seekers with more classroom-training inclined caseworkers neither meet sooner nor more frequently with their caseworker. To further support the exclusion restriction, we re-estimate our main IV specification while controlling for the frequency and timing of meetings between the job seeker and caseworker; see Appendix E.3. Our results are robust to these controls and suggest meeting frequency and timing do not explain our estimates for the effects of classroom training.

Table 5: Meetings, and Classroom-Training Tendency

		Within j-u-y		Classroom Training Tendency	
	obs	mean	sd	coef	se
Timing of first meeting	167,222	4.99	3.70	-0.34	(0.39)
Meetings/week	167,222	0.15	0.10	0.00	(0.01)

Notes: This table shows correlations between caseworkers' classroom-training tendencies and meeting behaviors with individual job seekers. "Timing of first meeting" refers to the first week the job seeker meets with a caseworker (first week relative to UI-spell start). "Meetings/week" refers to the number of meetings per week during the first six months of unemployment. Columns (1)-(3) report the number of observations along with the mean and standard deviation of the meeting behavior across job seekers within job-center-unit-years. Columns (4)-(5) report the coefficient and standard errors from separate regressions of a given meeting behavior on the caseworker's classroom-training tendency. All regressions include job-center-unit-year fixed effects as well as a control for the caseworker's on-the-job-training tendency. Standard errors (in parentheses) are two-way clustered on the predicted caseworker and job seeker level. *p<0.10 ** p<0.05 *** p<0.01.

In addition to the direct tests, our investigation of mechanisms in Section 7 also lends support to the exclusion restriction. First, in Section 7.1, we show that the effect of assignment to classroom training is driven by job seekers who leave unemployment after the assigned training program has ended (post-program effects). This timing is difficult to rationalize with other caseworker behaviors than training assignments. Second, in Section 7.2, we show that assignment to classroom training increases occupational mobility. Changing occupation likely requires new skills obtained through classroom training. Third, in Section 7.3, we present evidence that the employment effects of classroom training are driven by more successful job applications rather than by intensified job search. The higher success rate of job applications could reflect that classroom training provides jobseekers with skills demanded by employers.

Overall, we conclude that our caseworker-tendency instruments are relevant, and we find no violations of independence, (extended) monotonicity, or the exclusion restriction.

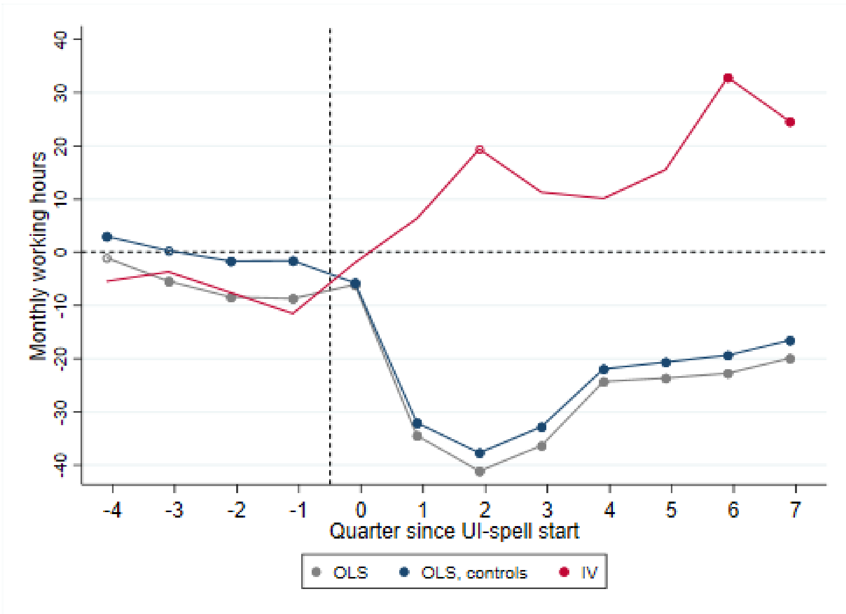
6 Effects of Assignment to Classroom Training

In the following sections, we use our caseworker-tendency instruments to estimate the effects of assignment to training on the employment of job seekers. We benchmark our IV estimates to OLS estimates that assume "selection on observables only." In Sections 6 and 7, we first focus on classroom training. In Section 8, we then compare the effects of classroom training to the effects of on-the-job training.

6.1 Employment Effects

Figure 3 shows the effect of assignment to classroom training on the average monthly work hours in a given quarter relative to job loss ("UI-spell start"). The figure includes our IV estimates, instrumenting assignment to training with caseworker training tendencies, and conventional OLS estimates, with and without controls. These controls include all socio-demographics and labor market history information from Table 3 as well as education levels (1-digit ISCED; 9 sections), previous industry (NACE level-1; 21 sections), and the typical occupation over the career (3-digit ISCO08). [21] The figure reveals three insights.

Figure 3: Effect of Assignment to Classroom Training



Notes: This figure shows the effect of assignment to classroom training on average monthly working hours in a given quarter relative to job loss. The gray and blue lines represent the effect obtained with a simple OLS regression and OLS that controls for job seeker pre-determined characteristics. These controls include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 sections), previous industry (NACE level-1; 21 sections), and the typical occupation over the career (3-digit ISCO08). The red line represents the effect obtained by IV estimation. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

First, the IV estimates show persistent positive effects of assignment to classroom training on employment. The employment gains grow steadily over time, stabilizing at about 25 hours per month two years after the initial job loss, equivalent to a 25% increase relative to hours

[21] Biewen et al. (2014) and van den Berg and Vikström (2022) propose a dynamic control strategy, combining the "selection on observables" assumption with the timing of program assignments. We adopt a simple control strategy based on pre-determined characteristics to facilitate comparison to the IV estimates, which do not use information about the timing of program assignment. All job seekers in our analysis sample have at least one meeting with their caseworkers and are thus exposed to their caseworkers' tendencies to assign training programs.

worked before job loss.[22]

Second, OLS reaches the *opposite* conclusion regarding the benefits of assignment to classroom training. In particular, the OLS estimates suggest training is associated with reduced employment, especially when job seekers are most engaged in training. The OLS conclusions align with the prior literature summarized by McCall et al. (2016).[23]

Third, controlling for pre-determined characteristics of the job seekers ("OLS, controls") does not change the strong lock-in effects in the post-period. The spurious lock-in effects of OLS highlight that job seekers with worse job *prospects* are the individuals who opt into training, revealing a *prospective* version of the Ashenfelter dip (Ashenfelter, 1978).

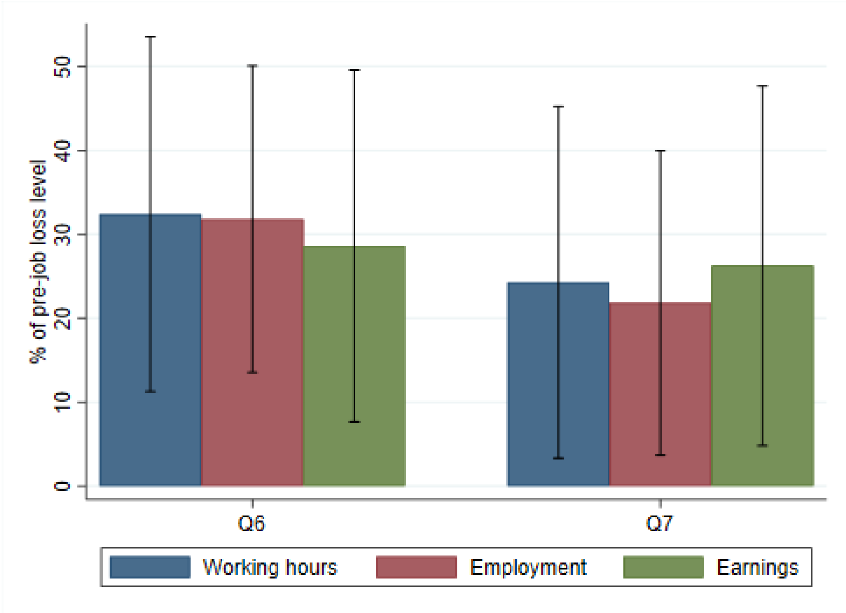
Appendix Figure A.3a and A.3b show the benefits of assignment to classroom training are consistent across alternative labor market outcomes. The figures show assignment to classroom training causes a positive and steadily growing effect on extensive margin employment and earnings. Figure 4 summarizes the effects two years after initial job loss. At this point, assignment to classroom training increases the extensive margin of employment by about 20 percentage points and the monthly earnings by about 25% of the job seekers' earnings before job loss. Hence, the effects on extensive margin employment and earnings are very similar to the effects on working hours.

The similarity of the effects across outcomes indicates *how* assignment to classroom training affects job seekers. First, the similar effects on earnings and working hours suggest assignment to classroom training primarily increases job seekers' earnings by boosting hours and not individual productivity (hourly wages). Second, the fact that the effects on extensive-margin employment and working hours are close (20% and 25%) suggests assignment to classroom training primarily increases the job seekers' ability to find jobs (extensive-margin employment) as opposed to the number of working hours conditional on employment (intensive-margin employment).

[22] The brief dip in the IV estimates in quarter three after job loss is likely statistical noise. Appendix Figure A.2 unpacks the IV estimates by plotting the employment outcomes with and without classroom training for compliers in a given quarter relative to job loss. This figure shows that the employment of compliers assigned to training flattens somewhat from quarter two (solid line), whereas the employment for compliers *not* assigned to training increases up until quarter four and only thereafter flattens (dashed line).

[23] See quote in Footnote 1.

Figure 4: Effect of Assignment to Classroom Training



Notes: This figure shows IV-estimates of the effect of assignment to classroom training on labor market outcomes in quarters six and seven relative to initial job loss. All outcomes are measured in percent of the pre-job-loss level. They include (non-supported) working hours, extensive margin employment, and earnings. All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job-seeker levels. Black lines represent 95% confidence intervals.

Appendix E presents various robustness tests of our IV estimates. First, our baseline TSLS specification controls linearly for assignment to on-the-job training, yet Blandhol et al. (2022) highlight the importance of allowing for *flexible* controls to interpret TSLS estimates as LATEs. Our theoretical framework in Appendix C also formalizes this point. In Appendix E.1, we show our results are robust to estimating our TSLS specification around an evaluation point z_2' for the on-the-job training instrument (Mountjoy, 2022). Second, as explained in Appendix C, our caseworker-tendency instrument shifts job seekers into classroom training from two margins: From "passive UI" to "classroom training only", and from "on-the-job training only" into "both on-the-job training and classroom training." Appendix E.2 shows most job seekers are shifted along the passive margin, and our results are robust to focusing exclusively on these compliers.

6.2 Differences between OLS and IV

In this section, we discuss the differences between the IV and OLS estimates. We conclude the differences are driven by negative *selection bias* (job seekers with worse job prospects opting into training) and not *effect heterogeneity* (instrument compliers gaining more from training than the average trainee).

First, examining effect heterogeneity on *observables*, Appendix Figure E.4 shows the OLS estimates are robust to reweighing job seekers to match the characteristics of compliers. Second, investigating the role of *unobserved* heterogeneity, Section 8.3 shows the marginal treatment effects are positive across job seekers with different latent resistances to training. These findings indicate the differences between OLS and IV are not driven by effect heterogeneity on unobservables. Finally, Appendix Figure A.4 highlights the negative selection into training by unpacking the OLS and IV estimates into their implied employment outcomes with and without training. The figure shows that job seekers who opt out of training have better employment prospects than job seekers whose training decisions depend on the training tendency of their caseworker. In Section 7.1, we further investigate the difference between OLS and IV by decomposing the effects into the underlying threat, lock-in, and post-program effects of assignment to classroom training.

7 Mechanisms

This section investigates the mechanisms through which assignment to classroom training affects employment. In Section 7.1, we first decompose the effects into threat, lock-in, and post-program effects of assignment to classroom training. In Section 7.2, we examine the role of occupational mobility in generating the employment gains from classroom training. Finally, in Section 7.3, we investigate whether assignment to classroom training alters individual job search behaviors.

7.1 Threat, Lock-in, and Post-program Effects

The OLS estimates in Figure 3 align with the existing literature (McCall et al., 2016; Jespersen et al., 2008; Munch and Skipper, 2008). In the literature, the short-run drop in employment has been ascribed to negative *lock-in* effects, whereas the long-run losses have led to the conclusion that the *post-program* effects of training are at best zero. The IV estimates, however, are difficult to rationalize with the same underlying dynamics. To investigate the difference between OLS and IV, we now decompose the effect of assignment to classroom training into the underlying threat, lock-in, and post-program effects.

Our decomposition relies on splitting all job seekers who get assigned to classroom training (assignees, hereafter) into four mutually exclusive training states in a given period, $s \in$

$\{a, b, c, d\}$. The state of an assignee depends on whether she (*a*) has not yet been assigned to training, (*b*) has not yet started her training assignment, (*c*) is undergoing training, or (*d*) is done with training. We can thus decompose the effects of assignment to classroom training on employment in a given period (β_{1t} in Equation (3)) into contributions from the assignees in each of the states:

$$\beta_{1t} = \sum_{s \in \{a,b,c,d\}} \gamma_{1t}^{s} \times (\beta_{1t}^{1s} - \beta_{1t}^{0s}), \tag{4}$$

where γ_{1t}^{s} denotes the share of assignees present in state s in period t, and $(\beta_{1t}^{0s}, \beta_{1t}^{1s})$ denotes their potential employment outcomes with and without the training assignment. Appendix F.1 describes how we assign job seekers to training states in our data.

Importantly, each of the components in Equation (4) corresponds to an effect discussed by the existing literature: job seekers who (*b*) have not yet started training are subject to the so-called *threat* effect of training. Job seekers who (*c*) are undergoing training are subject to the *lock-in* effect. Job seekers who (*d*) are done with training are subject to the *post-program* effect. Finally, job seekers who (*a*) have not yet been assigned to training are subject to a *placebo* effect.

Our estimation of Equation (4) proceeds in three steps. First, we estimate the training state probabilities γ_{1t}^{s} by regressing indicators D_{it}^{s} for whether a job seeker is in training state s on our treatment variables:

$$D_{it}^{s} = \gamma_{1t}^{s} D_{1i} + \gamma_{2t}^{s} D_{2i} + \Omega X_{it} + \varepsilon_{it}. \tag{5}$$

Second, we estimate the employment outcomes in each training state β_{1t}^{s}. To do so, we follow Abadie (2002) and interact the outcome variable Y with the state indicators D_{t}^{s} and then regress these interactions on our treatment-status indicators:

$$Y_{it} \times D_{it}^{s} = \beta_{1t}^{s} D_{it}^{s} + \beta_{2t}^{s} D_{2i} + \Omega X_{it} + \varepsilon_{it}. \tag{6}$$

Our third and final step is to estimate the state-specific counterfactual outcomes β_{1t}^{0s}. We adopt two approaches that turn out to yield similar results. In our first approach, we assume that the average counterfactual outcomes in a given period are homogeneous across states:

$$\beta_{1t}^{0s} = \beta_{1t}^{0}. \tag{7}$$

Then, again following Abadie (2002), we can estimate the counterfactual employment outcome, β^0_{1t}, by running the regression:

$$Y_{it} \times (1 - D_{1i}) = \beta^0_{1t}(1 - D_{1i}) + \beta^0_{2t} D_{2i} + \Omega X_{it} + \varepsilon_{it}, \tag{8}$$

Appendix Figure A.4 plots the estimated counterfactual employment outcomes, β^0_{1t}, according to OLS and IV.

In our second approach, we allow for heterogeneous state-specific counterfactuals, $\beta^{0s}_{1t} \neq \beta^0_{1t}$. In particular, we assume time-invariant selection premiums relative to never-takers. As the second approach is more technically involved, and turns out to deliver similar results, we show the first approach in the main text and relegate the second approach to Appendix F.2.

We use Equations (4)-(8) to decompose the OLS and IV estimates in Equation (3). For the IV estimates, we instrument the treatment variables (D_{1i}, D_{2i}) with our instruments (Z_{1i}, Z_{2i}), using the first-stage Equation (2). The vector of controls, X_{it}, includes job-center-unit-year fixed effects. For the OLS estimations, the vector also includes the rich set of job seeker predetermined characteristics used for the baseline OLS estimates in Figure 3.

7.1.1 Training-State Probabilities

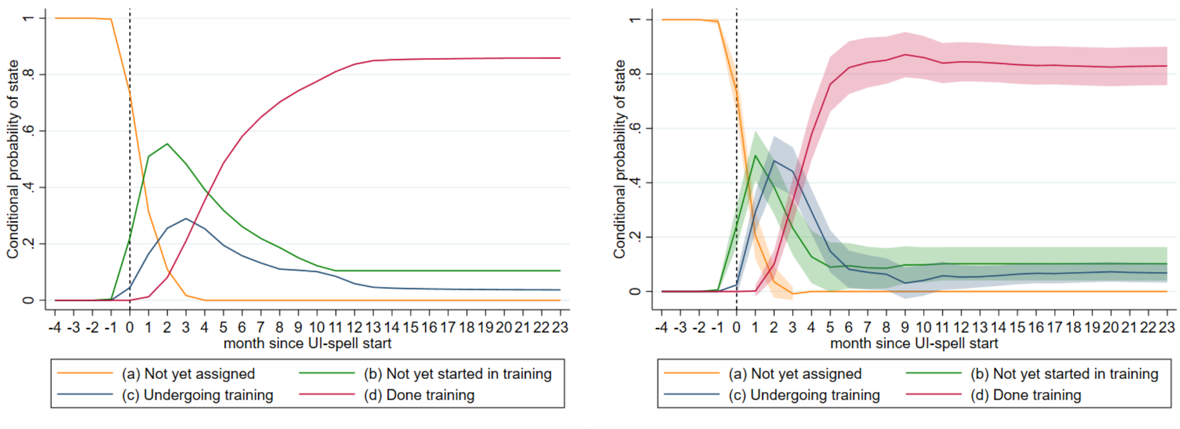
Figure 5 plots the training-state probabilities γ^s_{1t}, estimated by OLS (Panel (a)) and IV (Panel (b)). These probabilities represent the probability of being in a given state in a given month, conditional on being assigned to classroom training within the first 12 months of job loss, for the full population and the subpopulation of compliers, respectively.

Comparing Panels (a) and (b), the training dynamics are similar in the full population and the subpopulation of compliers. In both populations, most job seekers are waiting for their training program to begin (green) or undergoing the training program (blue) in the first couple of months after job loss. After the first year, about 90% of job seekers are done with their assigned training program (red).[24] This finding suggests the differences between the OLS and IV are *not* driven by heterogeneous training dynamics. For example, the OLS and IV are not different because the probability of being "locked in" to training is larger for the full population than for the complier population. If anything, compliers are more likely to be locked into training in the first couple of months.

[24]Note the finding that 90% of job seekers progress into state (d) is *not* mechanical; if a job seeker exits unemployment before completing the training program (d), she will remain in her latest state (b or c) in all future periods. See Online Appendix F.1 for details.

Because the two populations experience similar training dynamics, the difference between OLS and IV must be due to a difference in potential outcomes for assignees. Appendix Figure A.5 supports this hypothesis by showing that the employment outcomes of assignees (OLS) are lower than the corresponding counterfactual outcomes for compliers (IV). To further support this claim, Online Appendix E.5 shows the baseline OLS and the OLS re-weighted by the IV training dynamics are very similar.

Figure 5: Training-State Probabilities



(a) OLS (b) IV

Notes: This figure shows estimated state probabilities conditional on assignment to classroom training (γ^s_{1t}). The estimates are obtained by estimation of Equation (5). Panel (a) shows OLS estimates obtained by OLS regression that controls for job seeker covariates. The job seeker covariates include the socio-demographics and labor market history information from Table 3, education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08). Panel (b) shows IV estimates obtained by instrumenting assignments to classroom training and on-the-job training by the predicted caseworker tendencies (according to the first-stage Equation (2)). All regressions include fully interacted job-center-unit-year fixed effects. Shaded areas represent 95% confidence intervals.

7.1.2 Decomposition of Employment Effects

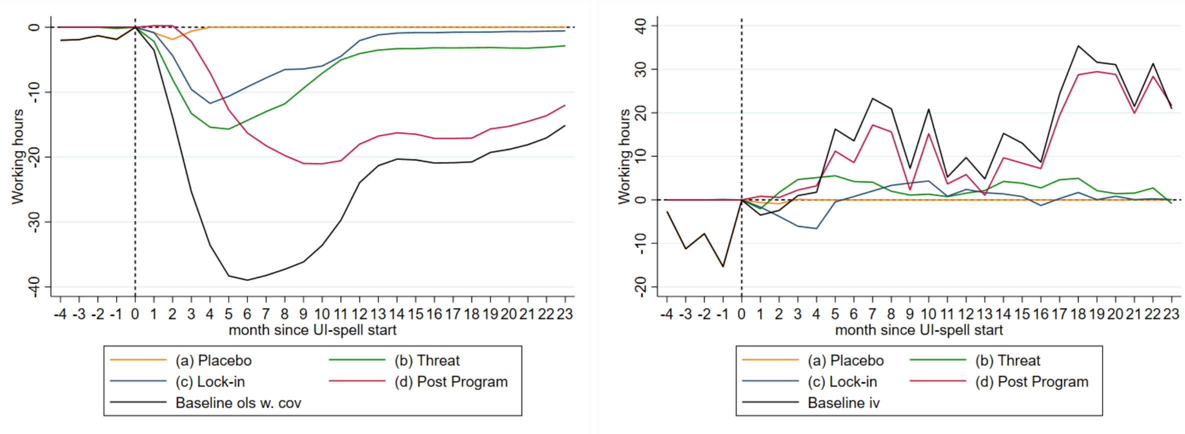
Figure 6 decomposes the employment effect of assignment to classroom training into a placebo, threat, lock-in, and post-program effect. Panels (a) and (b) represent a decomposition of the OLS and IV estimate, respectively.

Comparing the panels yields four findings. First, OLS and IV identify a similar-sized negative lock-in effect of classroom training. Hence, lock-in effects are not causing the difference between OLS and IV.

Second, OLS and IV are different in the short run because they identify very different threat effects. In particular, the IV estimates reveal a positive threat effect of assignment to classroom training, confirming prior experimental estimates (Black et al., 2003).[25] By contrast, the OLS suggests a strong negative threat effect. The difference between OLS and IV in estimated threat effects reflects that job seekers have worse job options while waiting for their assigned training program to begin.[26]

Third, OLS and IV are different in the long run because they identify very different post-program effects. The IV estimates show completion of classroom training helps job seekers find jobs, whereas the OLS estimates suggest classroom training hurts their long-run employment potential. The differences between OLS and IV show job seekers who opt into training have persistently lower employment potentials.[27]

Figure 6: Decomposition of the Effect on Working Hours



(a) OLS (b) IV

Notes: This figure shows a decomposition of the baseline estimate of the effect of assignment to classroom training on working hours in a given month relative to job loss. Panel (a) represents a decomposition of the baseline OLS estimate; panel (b) represents a decomposition of the baseline IV estimate. The baseline estimate is decomposed into contributions from each of the four training-states, $\gamma^s_{1t} \times (\beta^{1s}_{1t} - \beta^{0s}_{1t})$, estimated using Equations (4)-(8). For the IV estimates, we instrument assignments to classroom training and on-the-job training by the predicted caseworker tendencies (according to the first-stage Equation (2)). For the OLS estimates, we control for job seeker covariates. These include the socio-demographics and labor market history information from Table 3, education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08). This figure does not include indications of statistical significance.

Finally, the IV estimates show the positive long-run employment effects of classroom training are driven by post-program effects rather than threat effects.[28] This finding suggests classroom training increases employment by *reskilling* job seekers and that classroom training could

[25] Black et al. (2003) shows the threat of training makes job seekers exit unemployment; we show job seekers exit unemployment for employment.

[26] Appendix Figure A.5 shows average employment in state (c) is lower for the full population than in the subpopulation of compliers.

[27] Appendix Figure A.5 shows average employment in state (*d*) is lower for the full population than in the subpopulation of compliers.

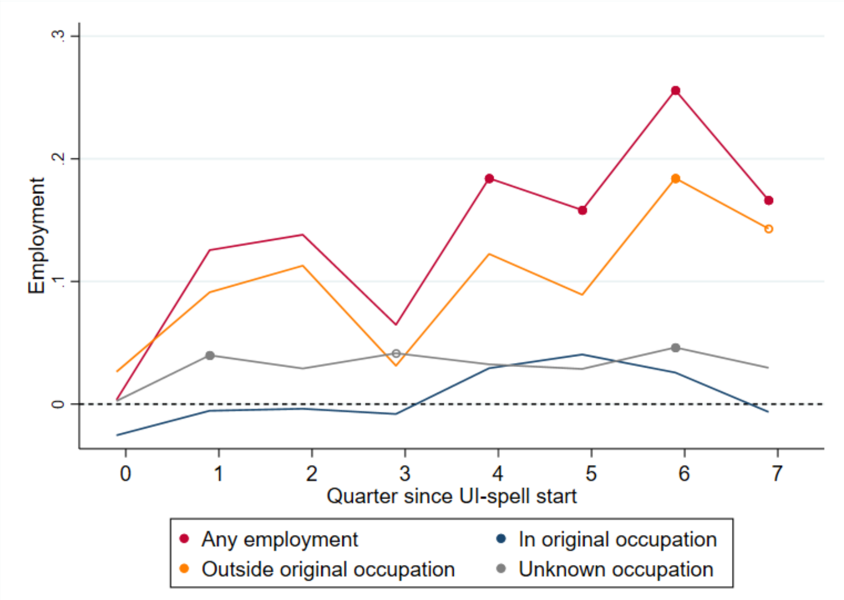
[28] To ease comparison with the baseline IV estimate, Appendix Figure A.6 aggregates the monthly contributions to the quarterly level.

help mitigate structural challenges in the labor market. In Section 8.2, we investigate whether job seekers exposed to offshoring have higher gains from classroom training.

7.2 Occupational Mobility

A core motivation of vocational training programs is to enhance the occupational mobility of job seekers.[29] Figure 7 evaluates this goal by decomposing the total employment effects of classroom training into employment inside and outside job seekers' previous occupations. In particular, for each job seeker, we identify the occupation that the job seeker most frequently held prior to job loss, which we denote as her *original occupation*. We then decompose total employment in a given period into three mutually exclusive categories: employment in the original occupation, outside the original occupation, and in an unknown occupation.[30] The figure is based on three-digit ISCO08 occupational codes, but Appendix Figure A.7 shows our results are qualitatively robust to using more or less disaggregated occupational codes.

Figure 7: Occupational Mobility



Notes: This figure shows the effect of assignment to classroom training on any employment (red) as well as employment in the job seeker's original occupation (blue), outside the original occupation (orange), or in an unknown occupation (gray). The job seeker's "original occupation" refers to the typical occupation she held prior to job loss. The comparison between the original occupation and any occupations held *after* job loss is based on 3-digit isco08 codes. The effects are obtained by IV-estimation of regressions that include job-center-unit-year fixed effects, and where training assignments are instrumented by caseworker training tendencies. Standard errors are two-way clustered on the predicted caseworker and job-seeker levels. Full (hollow) dots indicate statistical significance at the 5% (10%) level.

[29]The stated goal of the vocational training program in Denmark is to "solve labor market restructuring and adaptation problems in accordance with the needs on the labor market in a short and a long term perspective." (Danish Ministry of Education, 2021).

[30]We categorize the individual as employed in an unknown occupation if she has positive working hours but a missing occupation code.

Figure 7 shows the gains in employment from classroom training are primarily driven by employment *outside* job seekers' original occupations. This observation suggests the *post-program* effects of classroom training, which are dominant in the long run, help job seekers by enhancing their occupational mobility.

7.3 Job Search

Are the employment effects of assignment to classroom training driven by altered job search? Assignment to classroom training could in principle increase employment through intensified job search, e.g., by inducing job seekers to send more applications in response to the "threat" of training. Alternatively, employment prospects may improve due to more successful job applications, e.g., due to new skills that can be added to the CV upon completion of a training program. Our decomposition in Section 7.1, showing the long-run effects of classroom training are driven by post-program rather than threat effects, suggests *more successful job applications* is the primary mechanism. This mechanism would also be consistent with the occupational mobility results from section 7.2 as sending more job applications may not be sufficient to land a job outside the job seeker's original occupation. This likely requires new skills on the CV and/or contacts in the target sector.

To explore this further, we exploit that unemployed job seekers have been required to register applied-for jobs in an online joblog since September 2015.[31] Using data on the individual job logs within the first two years following job loss, we construct proxies for individual job search behaviors. For example, the number of job logs per month or the variety of job search channels or submission methods can be used as proxies for individual job search intensity. Furthermore, we have some information about the characteristics of the applied-for job, e.g. whether it was a full-time job, and whether the job was found through the job seeker's network.[32]

Table 6 shows that the classroom-training tendency of the predicted caseworker does not predict proxies for job search intensity or the probability to apply for full-time jobs. This finding suggests that any effect of classroom training on hours of employment is *not* driven by a more intense search for full-time jobs. Instead, the effect must be driven by more successful job applications, e.g., because classroom training provides new skills that make the job seeker better

[31] See Fluchtmann et al. (2022) for a validation of the data quality.

[32] Unfortunately, we do not have information about the occupation of the applied-for job. Hence, we cannot investigate whether job seekers assigned to classroom training direct their job search toward new occupations.

suited for the applied-for jobs.

In addition to new skills, classroom training may provide job seekers with a network and informal ties to the target sector, which could create new job opportunities.[33] Indeed, Table 6 shows a positive correlation between the caseworker's classroom-training tendency and the probability of finding jobs through one's network ("Network job"). These findings are consistent with Katz et al. (2022), who show that classroom courses combined with "wraparound" placement services is effective in placing marginalized workers in new jobs.

Table 6: Joblogs and Classroom-Training Tendency

		Within j-u-y		Classroom Training Tendency	
	obs	mean	sd	coef	se
Job search intensity					
- #Joblogs/month	62,412	5.61	2.86	-0.14	(0.24)
- #Distinct search channels	62,412	2.46	0.95	-0.05	(0.09)
- #Distinct submission types	62,412	3.04	1.19	-0.11	(0.11)
Characteristic of applied-for jobs					
-Unsolicited job	62,412	0.24	0.27	-0.00	(0.02)
-Network job	62,412	0.07	0.14	0.04***	(0.01)
-Full-time job	62,412	0.87	0.20	0.02	(0.02)
-Distance to job (km)	19,960	32.35	24.73	3.91	(4.20)
Submission of the application					
-Days applied before deadline	49,468	12.64	8.71	-2.20**	(0.94)
-Submit by email/letter	62,412	0.40	0.31	-0.02	(0.03)
-Submit by phone	62,412	0.15	0.23	0.02	(0.02)
-Submit in person	62,412	0.17	0.24	0.00	(0.02)
-Submit online (web)	62,412	0.26	0.29	-0.02	(0.03)

Notes: This table shows correlations between individual job seekers' job search behaviors and the classroom-training tendencies of caseworkers. The job search behaviors are approximated by job seekers' registrations in an online "Job log". "#Joblogs/month" represents the average number of applied-for jobs per month. "#Distinct search channels" and "#Distinct submission types" represent the number of distinct search channels and submission types ever registered, "Unsolicited job" represents the share of applied-for jobs that were unsolicited, "Network job" represents the share of applied-for jobs found through the job seeker's network, "Full-time job" represents the share of applied-for jobs that were full-time. "Distance to job (km)" represents the distance to the firm posting the applied-for job. "Days applied before deadline" represents the number of days between the deadline and submission of the application. "Submit by ..." indicates how the job seeker submitted the job application. The table is based on job seekers who became unemployed from September 2015 to December 2017; a period in which registration in the "Job log" was mandatory. The number of observations is lower for some variables, e.g. "Distance to job (km)", because not all information about the applied-for job was mandatory to register. Columns (1)-(3) report the number of observations along with the mean and standard deviation of the job-search behaviors across job seekers within job-center-unit-years. Columns (4)-(5) report the coefficients and standard errors from separate regressions of a given job-search behavior on the caseworker's classroom-training tendency. All regressions include job-center-unit-year fixed effects as well as a control for the caseworker's on-the-job-training tendency. Standard errors (in parentheses) are two-way clustered on the predicted caseworker and job seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Overall, the evidence for job search behaviors suggests assignment to classroom training increases employment through more successful job applications, e.g., due to new skills and contacts, rather than through intensified job search. This finding is consistent with our decomposition results in Section 7.1, showing that the employment gains from classroom training are

[33] As an example, a large provider of truck and bus driver certificates in Denmark, DEKRA, promises to help the job seeker find a job in the industry upon completing a course. These so-called "job guarantees" can be interpreted as job searches through the network.

driven by post-program effects. It is also consistent with our occupational mobility results in Section 7.2, showing that classroom training helps job seekers into jobs outside their original occupation.

8 Heterogeneity

This section studies heterogeneity in the effects of training. In Section 8.1, we first investigate the effect by *program* types, distinguishing between on-the-job training and classroom training. In Section 8.2, we explore whether classroom training is more beneficial to job seekers *exposed to offshoring*. In Section 8.3, we explore heterogeneity in treatment effects across job seekers' unobserved resistance to training by estimation of *marginal treatment effects* (MTE). Finally, we evaluate whether classroom training can close the employment gap between job seekers with different exposures to offshoring.

8.1 Classroom Training vs. On-The-Job Training

How does the effect of assignment to classroom training compare to the effect of assignment to on-the-job training? Since the quarterly estimates for assignment to on-the-job training are noisy, we focus on the yearly effects when comparing the two training programs.[34]

Table 7 summarizes the effects of assignment to classroom training and on-the-job training on average monthly hours of employment in a given year relative to job loss. In order to study longer-term effects of the two training programs, the table includes an additional year relative to Figure 3. The point estimates in this table suggest the effects of assignment to classroom and on-the-job training are different.

The estimates for classroom training are large and persistent. In accordance with the main results in Figure 3, the estimate in Column (2) shows assignment to classroom training increases average monthly employment by about 20 hours in year two after job loss. Although based on fewer observations, the estimate in Column (3) suggests these gains extend into year three.

The estimates for on-the-job training are statistically insignificant, yet interesting to compare to the estimates for classroom training.[35] The point estimate in Column (1) suggests on-the-job training causes a large – yet short-lived – increase in monthly working hours in the first year

[34] See Appendix Figure A.8 for a comparison of the quarterly effects of the two training programs.

[35] The lack of statistical significance is not too surprising; we have (five-fold) lower statistical power for the on-the-job training instrument (see Table D.1).

after initial job loss. Columns (2)-(3) suggest the temporary boost to employment quickly dies out. At best, there is a null effect of assignment to on-the-job training two (and three) years after initial job loss. Appendix Tables B.3-B.4 show the conclusions are similar for extensive-margin employment and earnings.

Table 7: Employment Effects of Classroom vs. On-The-Job Training

	Monthly Working Hours		
	(1)	(2)	(3)
	Year 1	Year 2	Year 3
Classroom Training	8.78	20.78**	18.45*
	(7.84)	(10.04)	(10.36)
On-The-Job Training	18.52	8.11	-21.39
	(19.20)	(21.42)	(20.69)
Obs	167,222	167,222	150,311

Notes: This table shows IV estimates for the effect of assignment to classroom training and on-the-job training on average monthly working hours in a given year relative to job loss. The number of observations drop in Column (3) since outcomes in year 3 are unobserved for jobseekers who lost their job more recently (in 2017). All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and jobseeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Overall, our IV estimates indicate assignment to classroom training is *more* effective in raising employment rates in the long run than on-the-job training. Our MTE estimations in Section 8.3 suggest that this conclusion holds if we extrapolate to the effect for a broader population (ATE). This finding challenges the conventional wisdom that on-the-job training is more effective, see McCall et al. (2016). However, this wisdom is mainly based on observational data and a "selection-on-observables" approach. Our finding is consistent with recent experimental evidence, showing that, in Uganda, vocational training programs are much more effective in combating long-term youth unemployment than firm-provided training (Alfonsi et al., 2020). Our results suggest similar differences in the effectiveness of classroom and on-the-job training in the case of Denmark.

8.2 Exposure to Offshoring

Training programs are often motivated by skill mismatches that arise from globalization (Hyman, 2018; Braxton and Taska, 2023). For example, offshoring of production could make some job seeker skills obsolete in the local labor market. We now investigate whether job seekers with higher exposure to offshoring benefit more from classroom training.

We adopt the offshorability index from Autor and Dorn (2013) and characterize all job seekers according to the offshorability of their most frequently held occupation before job loss. The offshorability index is based on the task content of occupations and, as stated by Autor and Dorn (2013), it "captures the degree to which an occupation requires either direct interpersonal interaction or proximity to a specific work location." The index is defined such that occupations requiring *less* interpersonal interaction and proximity to the workplace location are *more* exposed to offshoring.[36]

For simplicity, we distinguish between occupations with “high” and “low” risk of offshorability by defining all job seekers above the 75th percentile in the job seeker distribution to be high-risk. Using this definition, many high-risk job seekers in our sample were employed as *office clerks* prior to job loss. For comparison, many low-risk job seekers were employed as *child-care workers*. See Online Appendix Table OA2 for additional examples.

Appendix Figure A.9 plots the employment of job seekers by their exposure to offshoring. The figure shows job seekers at high risk of offshoring have depressed employment rates after job loss compared with job seekers at low risk. Strikingly, this employment gap persists three years after the initial job loss.

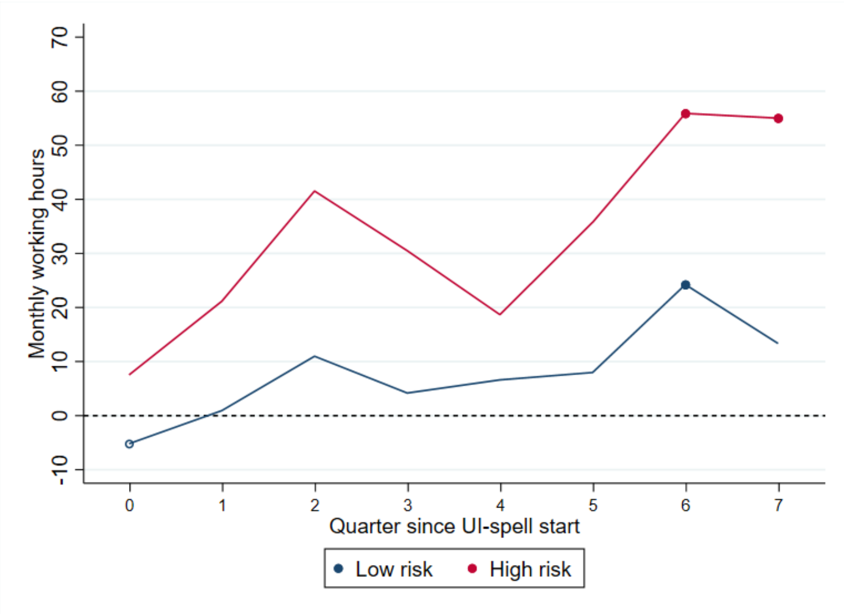
Figure 8 plots the employment effects of assignment to classroom training by job seeker exposure to offshoring. The figure shows job seekers at high risk of offshoring have large gains from assignment to classroom training. These employment gains increase over time. By quarter seven after the initial job loss, high-risk job seekers gain 55 hours of employment per month from assignment to classroom training. This gain corresponds to 50% of their pre-job loss level of employment.

In contrast, job seekers at low risk of offshoring derive much lower employment gains from assignment to classroom training. By quarter seven after the initial job loss, the gains for low-risk job seekers are not statistically significantly different from zero. Although the difference in employment gains from classroom training between job seekers at high and low risk of offshoring is *not* statistically significant (p-value=0.12), the mean difference in employment gains is economically significant. Appendix Figures A.10-A.11 show the conclusions are similar when considering the effects of classroom training on extensive-margin employment or earnings.

[36]Appendix OA1.6 describes how we use the O*NET database to construct an occupation-based measure of exposure to offshorability.

Why do job seekers at high risk of offshoring gain more from classroom training? In Section 7.1, we showed that the long-run gains from classroom training are driven by the post-program effects as job seekers complete the programs. Furthermore, in Section 7.2, we showed classroom training enhances occupational mobility. These pieces of evidence suggest classroom training helps job seekers reskill toward new occupations. Occupational mobility may be particularly helpful for job seekers exposed to offshoring, who may face obstacles reattaching to the labor market with their current skill set and may be hit by future offshoring shocks if they stay in their original occupations (Humlum and Munch, 2019).

Figure 8: Employment Effects of Classroom Training by Exposure to Offshoring



Notes: This figure shows IV estimates of the effect of assignment to classroom training on average monthly working hours in a given quarter relative to job loss, for job seekers at high and low risk of offshoring. The estimates are obtained by separately estimating Equations (2)-(3) for job seekers at high- and low risk. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

Table 8 compares the yearly employment effects of assignment to classroom and on-the-job training for job seekers in high and low risk of offshoring.[37] Overall, the table suggests the relative effectiveness of classroom and on-the-job training, as documented in Section 8.1, is driven by job seekers in high risk of offshoring. First, in accordance with Figure 8, the estimates suggest high-risk job seekers derive large and very persistent employment gains from assignment to classroom training. Even three years after initial job loss, the high risk job seekers still enjoy

[37] Appendix Tables B.5-B.6 show the yearly effects on extensive margin employment and earnings.

employment gains of about 50 hours per month from assignment to classroom training. Second, although statistically insignificant, the estimates suggest high-risk job seekers have less robust long-term gains from on-the-job training compared with classroom training. At best, on-the-job training has a null employment effect for high-risk job seekers three years after initial job loss. At worst, on-the-job training *reduces* employment by almost 100 hours per month for these job seekers. The lack of long-term positive gains could arise if on-the-job training reinforces the job seekers' ties to their original occupations where offshoring is likely to occur again (Hummels et al., 2012). In summary, the table suggests that high-risk job seekers enjoy both larger gains if assigned to classroom training, *and* losses if assigned to on-the-job training.

Table 8: Employment Effects of ALMPs by Exposure to Offshoring

	Average Monthly Working Hours		
	(1) year 1	(2) year 2	(3) year 3
Low risk			
Classroom Training	2.73	13.05	7.10
	(7.80)	(10.62)	(11.42)
On-The-Job Training	7.74	0.88	-7.99
	(18.52)	(21.83)	(20.77)
High risk			
Classroom Training	25.19	41.34*	50.71*
	(21.74)	(24.15)	(26.65)
On-The-Job Training	97.10	66.21	-98.04
	(90.15)	(92.24)	(92.82)
Obs low risk	125,413	125,413	113,079
Obs high risk	41,809	41,809	37,232

Notes: This table shows IV estimates for the effect of assignment to classroom training and on-the-job training on average monthly working hours in a given year relative to job loss, for job seekers at low and high risk of offshoring. The estimates are obtained by separately estimating Equations (2)-(3) for job seekers at high- and low risk. The number of observations drop in Column (3), since outcomes in year 3 are unobserved for jobseekers who lost their job more recently (in 2017). All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and jobseeker levels. *p<0.10 ** p<0.05 *** p<0.01.

8.3 Marginal Treatment Effects

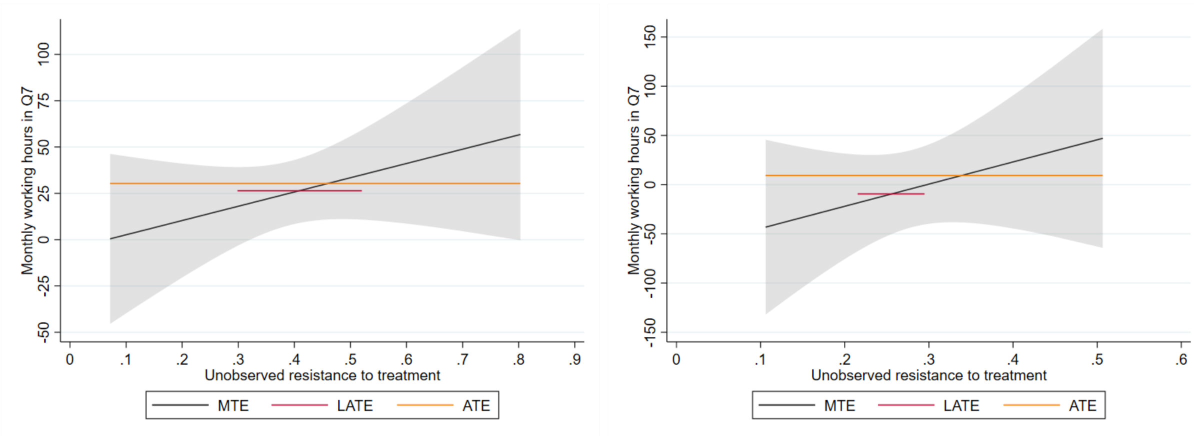
In this section, we investigate the selection patterns into classroom and on-the-job training by estimating marginal treatment effects (MTEs). Building on the theoretical framework in Section 4 (Appendix C), MTEs correspond to the average treatment effect (ATE) among job seekers with a particular resistance to training. To estimate the ATE in the broader population, we follow Mogstad et al. (2018) and impose shape restrictions on the MTE functions. In particular, we assume an additively separable and linear specification in the quantiles of resistances to training. Appendix C.4 details our estimation approach.

8.3.1 MTE, LATE, and ATE

Figure 9 plots the MTE functions for classroom and on-the-job training. Although imprecisely estimated, the MTE functions reveal important insights.

First, the MTE of classroom training (Panel (a)) is non-negative for all job seekers, varying between 0 and 56 additional working hours per month for job seekers with the lowest and highest resistance to training. In terms of mean resistance to training, the instrument compliers are representative for the broader population. Therefore, our LATE estimate for instrument compliers is informative about the effect of classroom training in the broader population (ATE).

Figure 9: Marginal Treatment Effects



(a) Classroom Training

(b) On-The-Job Training

Notes: This Figure plots the MTE, LATE and ATE on monthly working hours in quarter 7 after initial job loss. Panel (a) plots the effects of assignment to classroom training, and Panel (b), the effects of assignment to on-the-job training. As detailed in Appendix C.4, the *MTE* estimates are obtained from a specification with second-order polynomials in the propensity scores for classroom and on-the-job training. The estimations are based on job seekers within the common support of the relevant propensity score (see panels (a) and (b) in Appendix Figure A.12). Note the x-axis in this figure differs slightly from that in Appendix Figure A.12, because the propensity scores are re-estimated based on the common support sample). The shaded area represents 90% confidence intervals for the MTE function (these do not account for generated regressors). The ATE estimate is obtained by integrating the MTE function over the common support of the propensity score. The *LATE* estimate is obtained by integrating the MTE function from the propensity score for job seekers assigned to the least training-inclined caseworker to the propensity score for job seekers assigned to the most training-inclined caseworker (as approximated by percentile 1 and 99 on the relevant caseworker tendency instrument).

Second, the MTE of on-the-job training (Panel (b)) increases steeply with job seekers' resistance to treatment, varying from an employment *decrease* of 43 working hours per month to an *increase* of 47 working hours per month. Put differently, job seekers with the *lowest* gains from on-the-job training are the ones who have a *lower* resistance to treatment. For example, finding a wage-subsidized job for a job seeker who already has good job opportunities may be easier. The stark effect heterogeneity implies our LATE estimates are more negative than the average effects of on-the-job training in the broader population (ATE). These differences may help reconcile our findings with estimates in the literature, showing more positive effects for on-the-job-training interventions (McCall et al., 2016).

8.3.2 Policy Experiment

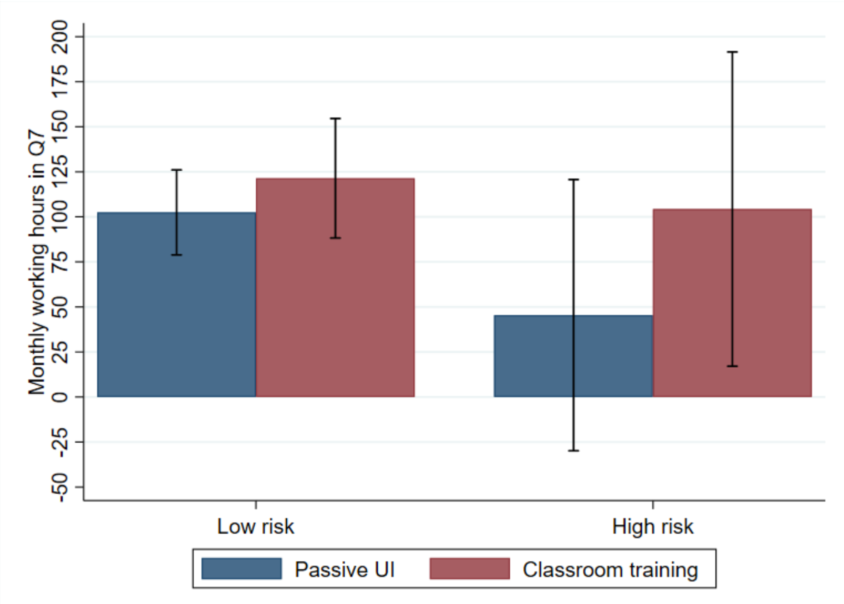
In Section 8.2, we found a persistent employment gap following job loss between job seekers at high and low risk of offshoring. In this section, we evaluate how much of the employment gap can be closed by targeting classroom training to high-risk job seekers. To evaluate this counterfactual, we leverage the MTE functions (estimated separately by risk groups) to evaluate employment gaps under different training scenarios.[38]

Figure 10 plots employment levels (two years after initial job loss) by risk group with and without assignment to classroom training. The plot delivers three takeaways. First, high-risk job seekers suffer depressed employment rates relative to low-risk job seekers in the scenario without training (blue bars). Second, providing all job seekers with classroom training (red bars) closes the employment gap between high- and low-risk job seekers. Third, because the employment gains from classroom training for low-risk job seekers are relatively low, a cost-effective way to close the employment gap is to redistribute classroom training from low-risk to high-risk job seekers. Note that this counterfactual scenario corresponds to assigning 25% of all job seekers to classroom training. In contrast, 39% of job seekers are assigned to classroom training in the status quo (see Table 1). Hence, total spending on classroom training programs would be lower in this counterfactual scenario.[39]

[38] Appendix Figure A.13 shows that, across risk groups, the MTE, LATE, and ATE estimates of the effect of classroom training are very similar.

[39] This exercise is micro-founded and abstracts from general equilibrium effects. Nevertheless, evaluating the effects of scaling *down* the use of classroom training programs – as opposed to up – makes general equilibrium effects less concerning. Furthermore, by reorienting workers toward in-demand occupations less exposed to offshoring, classroom training may have smaller displacement effects in the labor market (Katz et al., 2022).

Figure 10: Policy Counterfactuals



Notes: This figure shows the counterfactual employment rate in quarter 7 after an initial job loss, for low- and high-risk job seekers. The blue bars represent the counterfactual employment rate in a scenario in which *no* job seekers in the group are assigned to training ("passive UI"). The red bars represent the counterfactual employment rate in a scenario in which all job seekers in the group are assigned to classroom training. The counterfactual employment rates are estimated separately for low- and high-risk job seekers (see the MTE and ATE estimates by group in Appendix Figure A.13). Black lines represent 90% confidence intervals (these do not account for generated regressors).

9 Conclusion

This paper investigates the employment effects of assigning unemployed job seekers to training programs. Using caseworker-tendency instruments, we find large employment effects of assignment to classroom training: two years after job loss, the job seekers assigned to a classroom training program have 25% higher employment rates. By contrast, we do not find robust employment effects of on-the-job-training programs.

Interestingly, we find a stark difference between our IV estimates and OLS estimates that assume "selection on observables" only. The latter approach is widely used in the literature and suggests classroom training is associated with detrimental employment effects. The differences between OLS and IV highlight the importance of controlling for unobserved job-seeker characteristics.

Studying the dynamics of training programs, we show the large negative lock-in effects suggested by OLS, to a large extent, are driven by job seekers with *worse job prospects* opting into training. Further, we show the long-run employment effects detected by our IV estimates are driven by post-program effects rather than threat effects. That is, the increase in employment is mainly generated by job seekers who actually complete the training programs. We further show classroom training enhances employment *outside* the job seekers' original occupation, which suggests that participation in classroom training helps job seekers reskill.

Finally, we study heterogeneity in the causal effects of training programs across job seeker exposure to offshoring. We show job seekers at high risk of offshoring, who face depressed employment prospects after job loss, have larger gains from classroom training.

We estimate MTEs and use them to recover the ATE of training and evaluate counterfactual policies. We show our LATE estimate for classroom training can be extrapolated to the full population of job seekers and that classroom training can close the employment gap of job seekers exposed to offshoring.

References

Abadie, A. (2002), 'Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models', *Journal of the American statistical Association* **97**(457), 284–292.

Abbring, J. H. and van den Berg, G. J. (2003), 'The Nonparametric Identification of Treatment Effects in Duration Models', *Econometrica* **71**, 1491–1517.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M. and Vitali, A. (2020), 'Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda', *Econometrica* **88**(6), 2369–2414.

Ashenfelter, O. (1978), 'Estimating the Effect of Training Programs on Earnings', *The Review of Economics and Statistics* **60**, 47–57.

Autor, D. H. and Dorn, D. (2013), 'The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market', *American Economic Review* **103**(5), 1553–97.

Autor, D. H., Maestas, N., Mullen, K. J. and Strand, A. (2015), Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants, NBER Working Paper No. 20840.

Behaghel, L., Crépon, B. and Gurgand, M. (2013), Robustness of the Encouragement Design in a Two-Treatment Randomized Control Trial, IZA Discussion Paper No. 7447.

Behncke, S., Frölich, M. and Lechner, M. (2010), 'Unemployed and Their Caseworkers: Should They Be Friends or Foes?', *Journal of the Royal Statistical Society: Series A* **173**(1), 67–92.

Bhuller, M., Dahl, G. B., Løken, K. V. and Mogstad, M. (2020), 'Incarceration, Recidivism, and Employment', *Journal of Political Economy* **128**(4), 1269–1324.

Bhuller, M. and Sigstad, H. (2022), 2SLS with Multiple Treatments, Working paper, University of Oslo.

Biewen, M., Fitzenberger, B., Osikominu, A. and Paul, M. (2014), 'The Effectiveness of Public-Sponsored Training Revisited: The Importance of Data and Methodological Choices', *Journal of Labor Economics* **32**(4), 837–897.

Black, D. A., Smith, J. A., Berger, M. C. and Noel, B. J. (2003), 'Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system', *American Economic Review* **93**(4), 1313–1327.

Blandhol, C., Bonney, J., Mogstad, M. and Torgovitsky, A. (2022), 'When is 2SLS actually LATE?', *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2022-16).

Bown, C. and Freund, C. (2019), Active labor market policies: Lessons from other countries for the United States, Peterson Institute for International Economics Working Paper No. 19/2.

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

Card, D., Kluve, J. and Weber, A. (2010), 'Active Labour Market Policy Evaluations: A Meta-Analysis', *The Economic Journal* **120**(548), F452–F477.

Card, D., Kluve, J. and Weber, A. (2018), 'What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations', *Journal of the European Economic Association* **16**(3), 894–931.

Cederlöf, J., Söderström, M. and Vikström, J. (2021), What Makes a Good Caseworker?, IFAU Working Paper No. 2021:9.

Costa, R., Dhingra, S. and Machin, S. (2019), Trade and Worker Deskilling, NBER Working Paper No. 25919.

Crépon, B. and van den Berg, G. (2016), 'Active Labor Market Policies', *Annual Review of Economics* **8**, 521–546.

Dahl, G. B., Kostøl, A. R. and Mogstad, M. (2014), 'Family Welfare Cultures', *The Quarterly Journal of Economics* **129**(4), 1711–1752.

Danish Ministry of Education (2021), 'Adult vocational training | ministry of children and education', `https://eng.uvm.dk/adult-education-and-continuing-training/adult-vocational-training`. (Accessed on 02/11/2021).

Fluchtmann, J., Glenny, A. M., Harmon, N. and Maibom, J. (2022), Unemployed Job Search across People and over Time: Evidence from Applied-for Jobs, Working Paper, University of Copenhagen.

Frandsen, B., Lefgren, L. and Leslie, E. (2023), 'Judging judge fixed effects', *American Economic Review* **113**(1), 253–77.

Frölich, M. and Lechner, M. (2010), 'Exploiting Regional Treatment Intensity for the Evaluation of Labor Market', *Journal of the American Statistical Association* **105**, 1014–1029.

Gautier, P., van der Klaauw, B., Mueller, P., Rosholm, M. and Svarer, M. (2018), 'Estimating Equilibirum Effects of Job Search Assistance', *Journal of Labor Economics* **36**, 1073–1125.

Graversen, B. K. and van Ours, J. C. (2008), 'How to Help Unemployed Find Jobs Quickly: Experimental Evidence from a Mandatory Activation Program', *Journal of Public Economics* **92**, 2020–2035.

Heckman, J., Hohmann, N., Smith, J. and Khoo, M. (2000), 'Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment', *The Quarterly Journal of Economics* **115**(2), 651–694.

Heckman, J. J., Urzua, S. and Vytlacil, E. (2006), 'Understanding Instrumental Variables in Models with Essential Heterogeneity', *The Review of Economics and Statistics* **88**(3), 389–432.

Humlum, A. and Munch, J. R. (2019), Globalization, Flexicurity and Adult Vocational Training in Denmark, *in* 'Making Globalization More Inclusive - Lessons from Experience with Adjustment Policies', WTO.

Hummels, D., Jørgensen, R., Munch, J. and Xiang, C. (2014), 'The Wage Effects of Offshoring: Evidence from DAnish Matched Worker-Firm Data', *American Economic Review* **104**(6), 1597–1629.

Hummels, D., Munch, J. R. and Xiang, C. (2018), 'Offshoring and Labor Markets', *Journal of Economic Literature* **56**, 981–1028.

Hummels, D., Munch, J., Skipper, L. and Xiang, C. (2012), 'Offshoring, Transition and Training: Evidence from Danish Matched Worker-Firm Data', *American Economic Review Papers and Proceedings* **102**, 424–428.

Hyman, B. G. (2018), Can Displaced Labor Be Retrained? Evidence from Quasi-Random Assignment to Trade Adjustment Assistance, Working paper, University of Chicago.

Imbens, G. W. and Angrist, J. D. (1994), 'Identification and Estimation of Local Average Treatment Effects', *Econometrica* pp. 467–475.

Jespersen, S. T., Munch, J. R. and Skipper, L. (2008), 'Costs and Benefits of DAnish Active Labour Market Programmes', *Labour Economics* **15**(5), 859–884.

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

Kirkeboen, L. J., Leuven, E. and Mogstad, M. (2016), 'Field of Study, Earnings, and Self-Selection', *The Quarterly Journal of Economics* **131**(3), 1057–1111.

Kreiner, C. T. and Svarer, M. (2022), 'Danish Flexicurity: Rights and Duties', *Journal of Economic Perspectives* **36**(4), 81–102.

Lee, S. and Salanié, B. (2018), 'Identifying Effects of Multivalued Treatments', *Econometrica* **86**(6), 1939–1963.

Lee, S. and Salanié, B. (2020), Filtered and Unfiltered Treatment Effects with Targeting Instruments, CEPR Discussion Paper No. 15092.

McCall, B., Smith, J. and Wunsch, C. (2016), Government-Sponsored Vocational Education for Adults, *in* 'Handbook of the Economics of Education', Vol. 5, Elsevier, pp. 479–652.

Mogstad, M., Santos, A. and Torgovitsky, A. (2018), 'Using Instrumental Variables for Inference about Policy Relevant Treatment Parameters', *Econometrica* **86**(5), 1589–1619.

Mountjoy, J. (2022), 'Community Colleges and Upward Mobility', *American Economic Review* **112**(8), 2580–2630.

Munch, J. R. and Skipper, L. (2008), 'Program Participation, Labor Force Dynamics, and Accepted Wage Rates', *Advances in Econometrics* **21**, 197–262.

Slotsholmen (2013), Kortlægning af indholdet i øvrig vejledning og opkvalificering, Technical report.

van den Berg, G. and Vikström, J. (2022), 'Long-Run Effects of Dynamically Assigned Treatments: A New Methodology and an Evaluation of Training Effects on Earnings', *Econometrica* **90**(3), 1337–1354.

Vikström, J., Rosholm, M. and Svarer, M. (2013), 'The Relative Efficiency of Active Labour Market Policies: Evidence from a Social Experiment and Non-parametric Methods', *Labour Economics* **24**, 58–67.

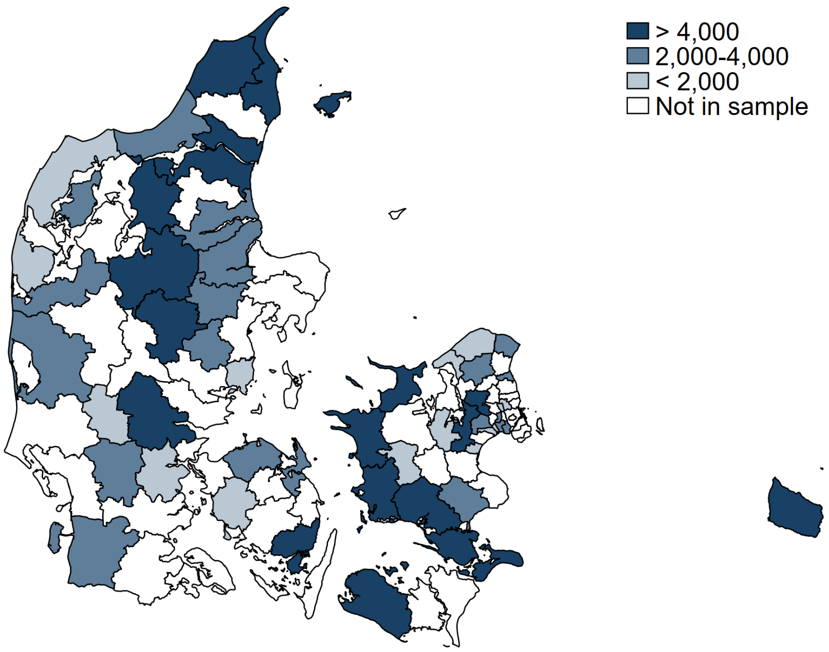
VIVE (2022), Jobcentrenes Beskæftigelses-Indsats, Technical report.

World Economic Forum (2020), 'Toward a Reskilling Revolution', www.weforum.org/projects/reskilling-revolution-platform. Accessed 2020-04-29.

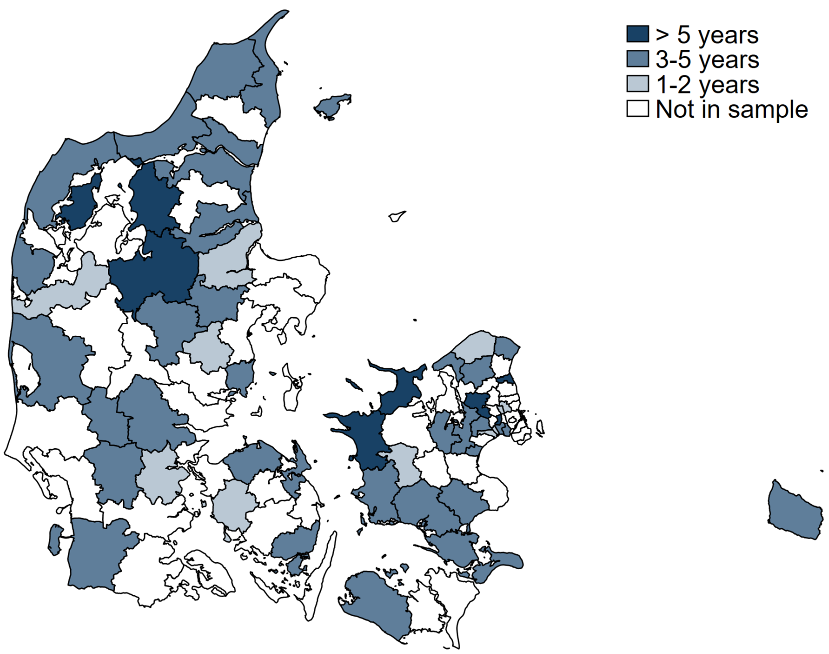
Appendix

A Appendix Figures

Figure A.1: Analysis Sample Geographical Coverage



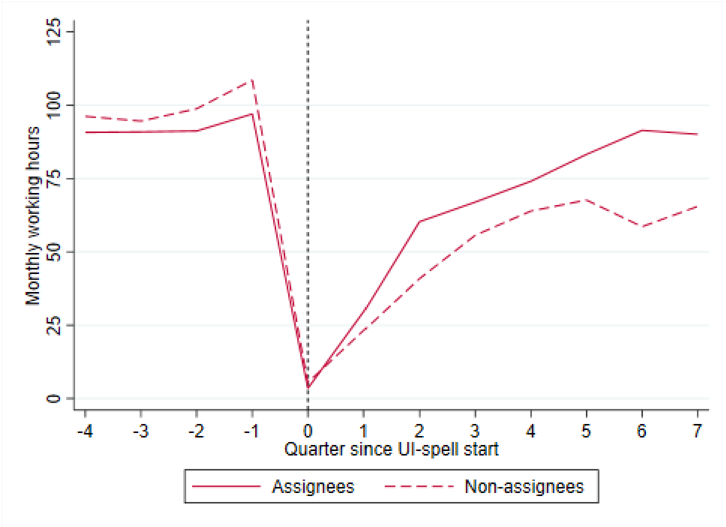
(a) Number of job seekers



(b) Number of years

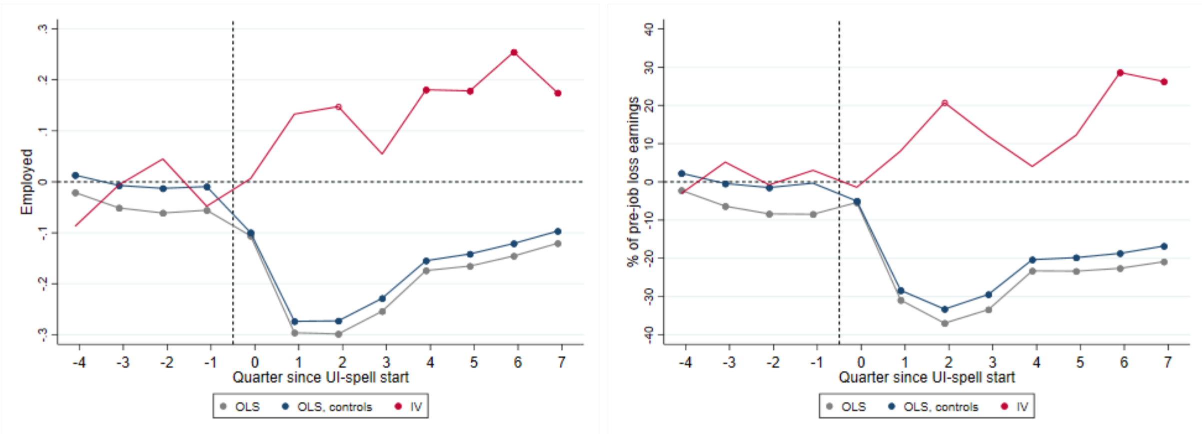
Notes: This figure breaks the full analysis sample by job centers in Denmark (i.e. UI spells from 2012 to 2018). Panel (a) shows the number of job seekers from a given job center. Panel (b) shows the total number of years that a given job center is part of our analysis sample. With four exceptions, a job center corresponds to a municipality (Esbjerg-Fanø, Ishøj-Vallensbæk, Læsø-Frederikshavn and Dragør-Tårnby).

Figure A.2: Average Working Hours



Notes: This figure plots average monthly working hours by training assignment status for individuals in the IV population (compliers). The outcome for compliers (not) assigned to classroom training is obtained by running a regression with the outcome interacted by an indicator for (non-) assignment to classroom training on the left-hand side, and indicators for (non) assignment to classroom training and assignment to on-the-job training on the right-hand side. The training indicators on the right-hand side are instrumented by the caseworker training tendencies. All regressions include job-center-unit-year fixed effects. No indication of statistical significance in this figure.

Figure A.3: Effect of Assignment to Classroom Training

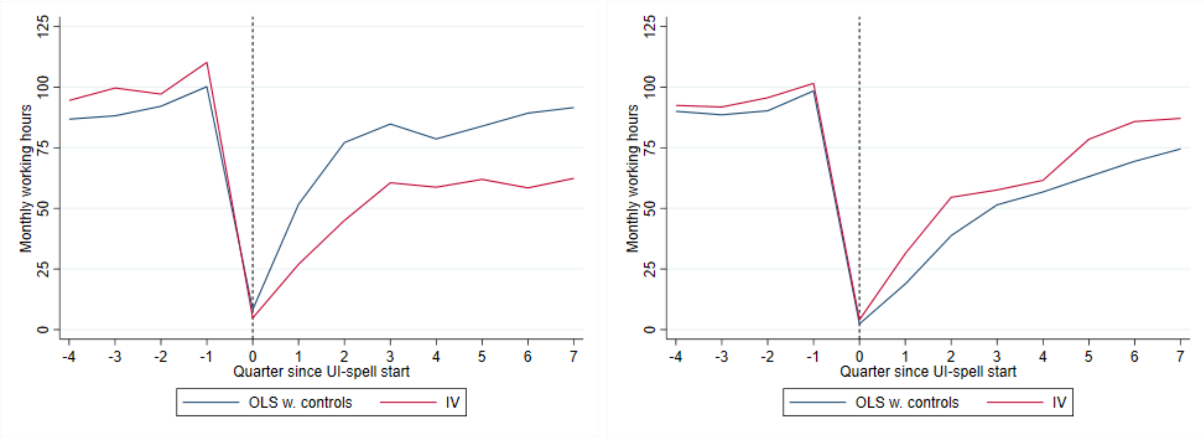


(a) Extensive-Margin Employment

(b) Earnings (percent of pre-job loss level)

Notes: This figure plots the effect of assignment to classroom training on extensive margin employment (Panel (a)) and monthly earnings in percent of the pre-job loss level (Panel (b)) in a given quarter relative to job loss. The gray and blue lines represent the effect obtained with a simple OLS regression and OLS that controls for job seeker predetermined characteristics. These controls include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08).The red line represents the effect obtained by IV-estimation. All regressions include fully interacted job-center unit and year fixed effects. Standard errors are two-way clustered on predicted caseworker and job-seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

Figure A.4: Average Working Hours - OLS and IV

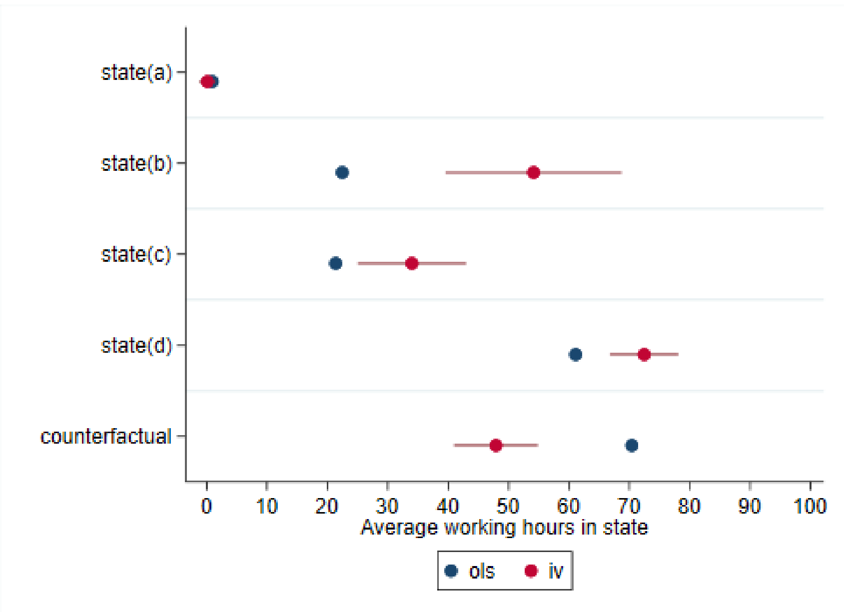


(a) Not assigned to training

(b) Assigned to training

Notes: This figure plots average monthly working hours by training assignment status for individuals in the OLS (blue) and IV (red) population. Panel (a) plots the outcomes for job seekers *not* assigned to classroom training, and Panel (b) plots the outcomes for job seekers *assigned* to classroom training. The outcome for job seekers (not) assigned to classroom training is obtained by running a regression with the outcome interacted by an indicator for (non-) assignment to classroom training on the left-hand side, and indicators for (non) assignment to classroom training and assignment to on-the-job training on the right-hand side. The OLS regressions also include a rich set of predetermined job seeker characteristics: socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08). The IV regressions instrument training (non) assignments by the caseworker training tendencies. All regressions include job-center-unit-year fixed effects. No indication of statistical significance in this figure.

Figure A.5: Average Employment by Training State

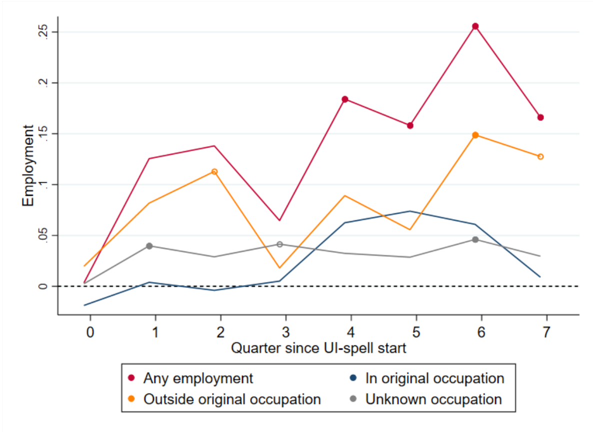


Notes: This figure shows average monthly working hours in months 0-23 following job loss and 95% confidence intervals for a given training state. First, by estimation of Equation (6), we obtain estimates for the monthly working hours in a given training state s and month t. Second, using the state probabilities as weights (see Figure 5), we compute a weighted average of the monthly working in hours in months 0-23 for a given state. Blue dots represent estimates obtained by OLS regressions, and red dots represent estimates obtained by IV regressions. Standard errors are constructed based on 100 bootstrap repetitions (see Appendix OA2.1).

Figure A.6: Threat, Lock-in, and Post-program Effects by Quarter after Job Loss

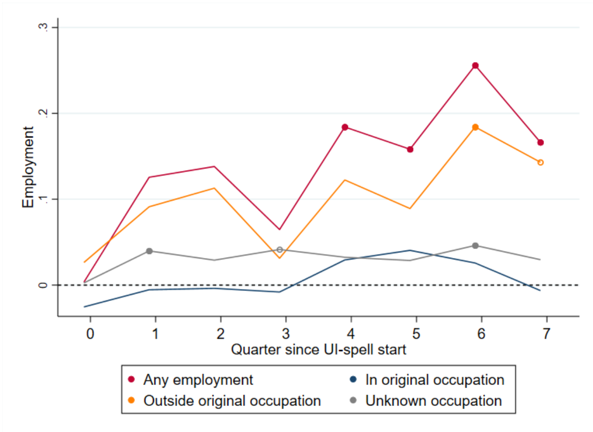
Notes: This figure shows the IV placebo, threat, lock-in, and post-program effects of assignment to classroom training on average monthly working hours in a given quarter relative to job loss. The quarterly effects are simple averages of the monthly effects plotted in Figure 6. No indications of statistical significance in this figure.

Figure A.7: Occupational Mobility

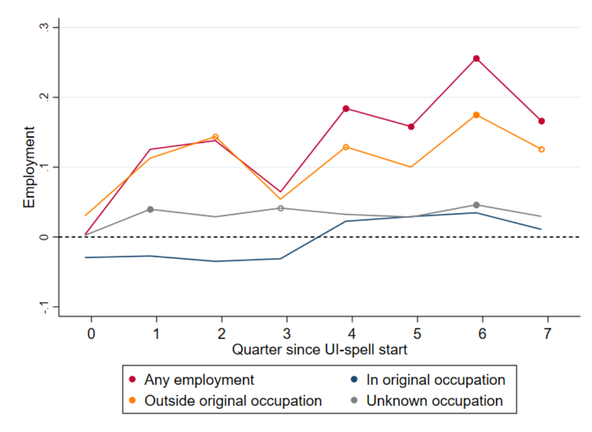


(a) 1-digit isco08

(b) 2-digit isco08



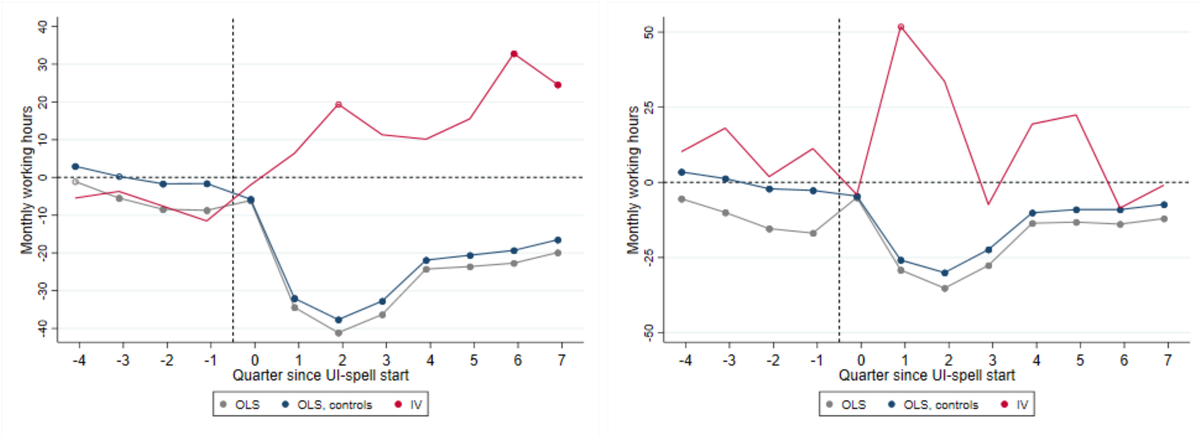
(c) 3-digit isco08



(d) 4-digit isco08

Notes: This figure shows the effect of assignment to classroom training on any employment (red) as well as employment in the job seeker's original occupation (blue), outside the original occupation (orange), or in an unknown occupation (gray). The job seeker's "original occupation" refers to the typical occupation she held prior to job loss. The difference between panels (a)-(d) is whether the comparison of the original occupation and any occupations held *after* job loss is based on 1-, 2-, 3-, or 4-digit isco08 codes. The effects are obtained by IV-estimation of regressions that include job-center-unit-year fixed effects, and where training assignments are instrumented by caseworker training tendencies. Standard errors are two-way clustered on the predicted caseworker and job-seeker levels. Full (hollow) dots indicate statistical significance at the 5% (10%) level.

Figure A.8: Effect of Training on Hours of Employment

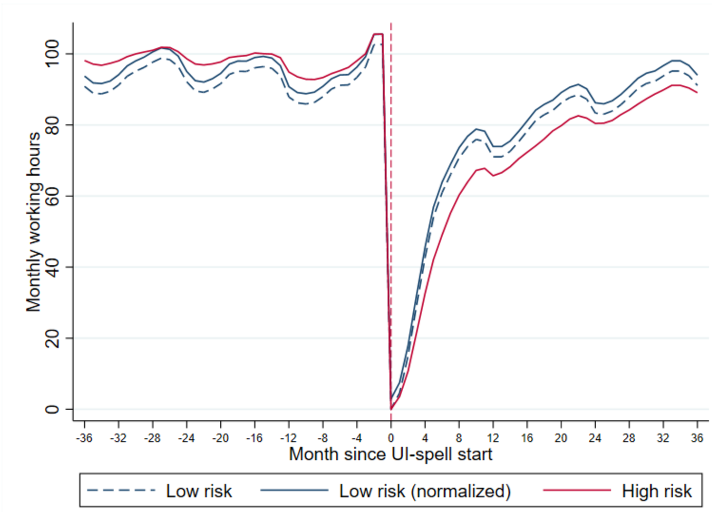


(a) Classroom Training

(b) On-The-Job Training

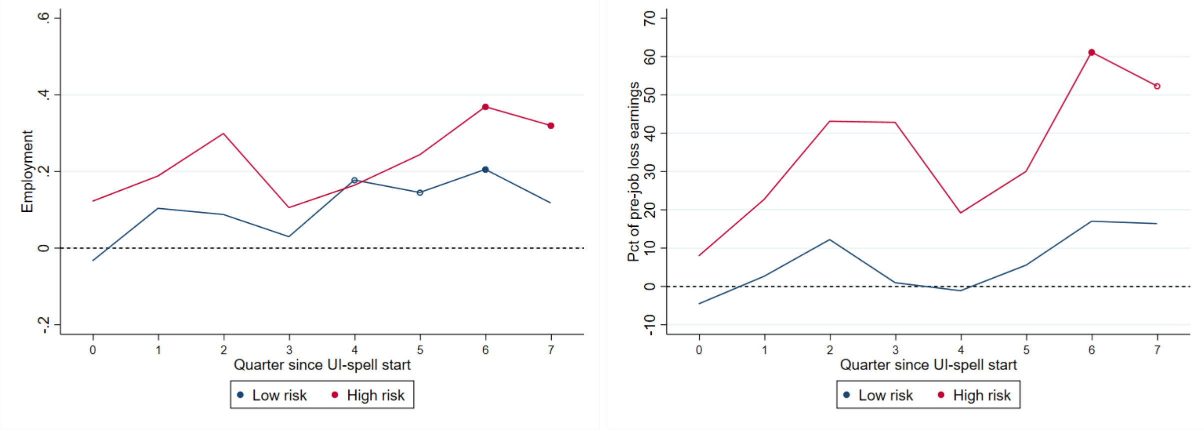
Notes: This figure shows the effect of assignment to classroom training (Panel (a)) and on-the-job training (Panel (b)) on average monthly working hours in a given quarter relative to job loss. The gray and blue lines represent the effect obtained with a simple OLS regression and OLS that controls for predetermined job seeker characteristics. These include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08). The red line represents the effect obtained by IV estimation. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

Figure A.9: Employment Prospects by Exposure to Offshoring



Notes: This figure shows employment by job seeker exposure to offshoring. The offshorability index is based on Autor and Dorn (2013), and we use the 75th percentile in the job-seeker distribution to distinguish between job seekers at high and low risk. To facilitate comparison, we normalize the employment level for the low-risk job seekers to match the employment of high-risk job seekers in the month before job loss. No indication of statistical significance in this figure.

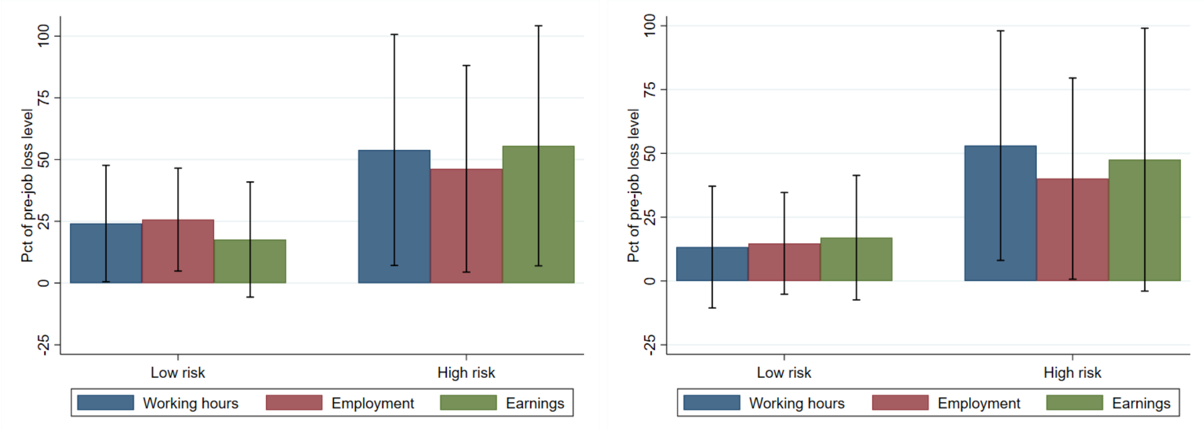
Figure A.10: Effects of Classroom Training by Exposure to Offshoring



(a) Extensive Margin Employment (b) Earnings

Notes: This figure shows IV estimates of the effect of assignment to classroom training on extensive-margin employment and earnings in a given quarter relative to job loss, for job seekers at high and low risk of offshoring. The estimates are obtained by separately estimating Equations (2)-(3) for job seekers at high- and low risk. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job-seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

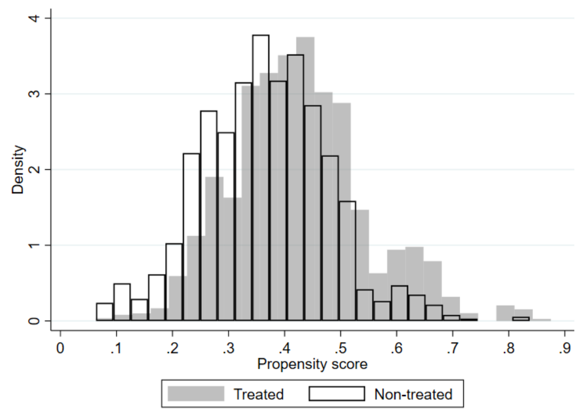
Figure A.11: Effect of Assignment to Classroom Training by Exposure to Offshoring



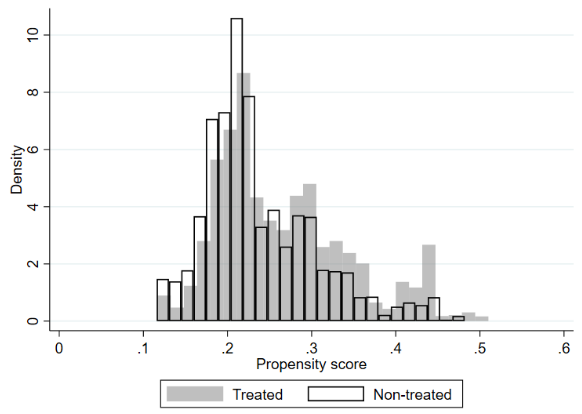
(a) Quarter 6 after job loss (b) Quarter 7 after job loss

Notes: This figure shows IV-estimates of the effect of assignment to classroom training on labor market outcomes in quarter six (Panel (a)) and quarter seven (Panel (b)) relative to initial job loss, for job seekers at high and low risk of offshoring. All outcomes (non-supported working hours, extensive margin employment, and earnings) are measured in percent of the pre-job-loss level. The estimates are obtained by separately estimating Equations (2)-(3) for job seekers at high- and low risk. All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job-seeker levels. Black lines represent 95% confidence intervals.

Figure A.12: Common Support



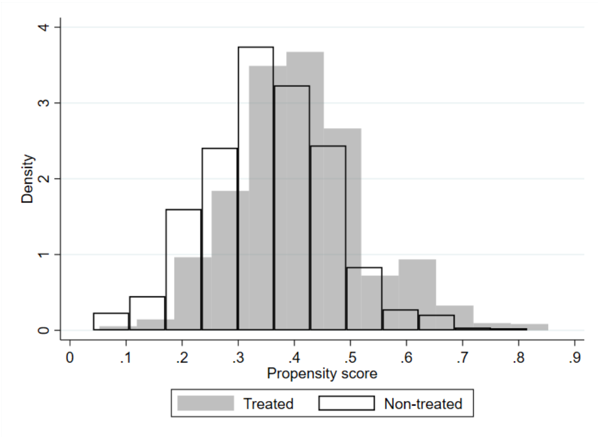
(a) Classroom Training



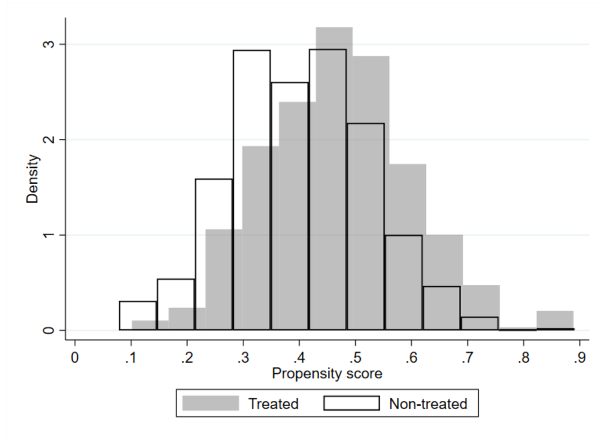
(b) On-The-Job Training

Notes: This figure shows the common support of the propensity score for treated and non-treated job seekers. Panel (a) represents the propensity score for classroom training and Panel (b) represents the propensity score for on-the-job training. The propensity scores are obtained by estimating and saving the predicted values from the first-stage equations (2).

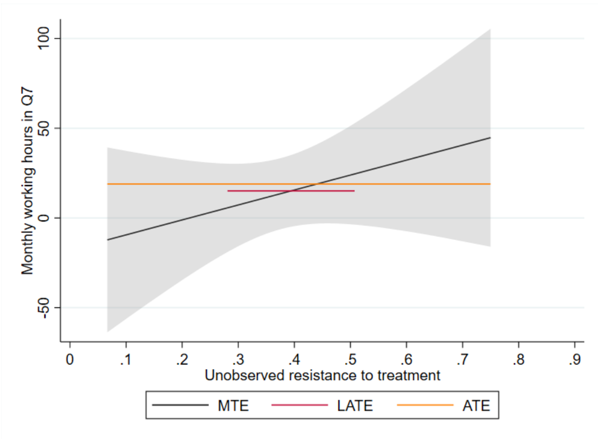
Figure A.13: Common Support and Marginal Treatment Effect of Classroom Training by Offshorability Risk



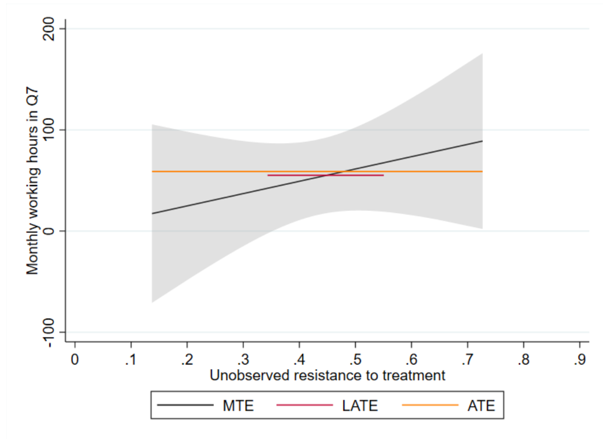
(a) Common support, low risk



(b) Common support, high risk



(c) MTE, low risk



(d) MTE, high risk

Notes: This figure shows the MTE estimations of the effect of assignment to classroom training on monthly working hours in quarter 7, conducted separately for job seekers at low and high risk of offshoring. Panels (a) - (b) plot the common support of the propensity score for treated and non-treated job seekers. The propensity scores are obtained by estimating and saving the predicted values from the first-stage equations (2). Panels (c) - (d) plot the MTE, LATE, and ATE of assignment to classroom training on monthly working hours in quarter 7 relative to job loss. As detailed in Appendix C.4, the *MTE* estimates are obtained from a specification with second-order polynomials in the propensity scores for classroom and on-the-job training. The estimations are based on job seekers within the common support of the relevant propensity score. The shaded area represents 90% confidence intervals for the MTE function (these do not account for generated regressors). The ATE estimate is obtained by integrating the MTE function over the common support of the relevant propensity score. The *LATE* estimate is obtained by integrating the MTE function from the propensity score for job seekers assigned to the least training-inclined caseworker to the propensity score for job seekers assigned to the most training-inclined caseworker (as approximated by percentile 1 and 99 on the relevant caseworker tendency instrument).

B Appendix Tables

Table B.1: Analysis Sample Restrictions

	UI-spells	Jobseekers	Jobcenters	Caseworkers Realized	Caseworkers Predicted	Average Weeks of UI	Average Compliance
Linked jobseeker-caseworker data	**934,922**	**628,352**	**94**	**7,910**	**1,949**	**33.35**	**0.42**
- Only j-u-y's using day-of-birth	203,101	152,057	51	2,330	680	32	0.52
- No non-western imigrants	196,522	146,964	51	2,289	680	32	0.52
- Previous occupation observed	192,813	144,046	51	2,270	680	32	0.52
- Min two caseworkers per j-u-y	186,604	139,342	51	2,245	670	32	0.52
- Caseload size $\geq$ 50	183,778	137,477	51	2,225	577	32	0.52
Full sample (2012-2018)	**183,778**	**137,477**	**51**	**2,225**	**577**	**31.88**	**0.52**
Balanced sample (2012-2017)	**167,222**	**127,713**	**51**	**2,060**	**536**	**31.92**	**0.52**

Notes: This table shows the number of units (UI-spells, job seekers, job centers, realized and predicted caseworkers) retained after each sample-selection step along with sample descriptives (average UI-spell length and the compliance between realized and predicted caseworker). Row (1) reports the statistics for all job seekers, who i) initiated a UI spell from 2012-2018 and ii) can be linked to a caseworker (i.e. all job seekers in the linked job-seeker–caseworker data set; see Online Appendix Table OA1). Row (2) restricts the sample to job-center-unit-years that use a clear (block) structure for the allocation of birthdays. Row (3) restricts the sample to job seekers with a non-western origin. Row (4) restricts the sample to job seekers for whom we observe their previous occupation. Row (5) restricts the sample to job-center-unit-years with minimum two (predicted) caseworkers. Row (6) restricts the sample to (predicted) caseworkers who were assigned at least 50 job seekers. Row (7) is identical to row (6) and summarizes the final (full) analysis sample. Row (8) reports statistics for a balanced analysis sample; job seekers for whom we observe their labor market outcomes throughout the first two years after job loss. These are job seekers who initiated a UI spell from 2012 to 2017.

Table B.2: Summary of Caseworker-Tendency Instruments

	count	mean	sd	1	10	25	50	75	90	99
Classroom-Training Instrument										
Z_1	167,222	0.39	0.12	0.15	0.24	0.31	0.38	0.45	0.55	0.73
Z_1, residualized by $q(i)$	167,222	0.39	0.04	0.29	0.34	0.37	0.39	0.42	0.44	0.50
Z_1, residualized by $q(i)$, Z_2	167,222	0.39	0.04	0.29	0.34	0.37	0.39	0.42	0.44	0.50
On-The-Job-Training Instrument										
Z_2	167,222	0.24	0.07	0.12	0.17	0.19	0.23	0.28	0.35	0.46
Z_2, residualized by $q(i)$	167,222	0.24	0.03	0.16	0.21	0.23	0.24	0.26	0.28	0.33
Z_2, residualized by $q(i)$, Z_1	167,222	0.24	0.03	0.16	0.21	0.23	0.24	0.26	0.28	0.33

Notes: This table reports pseudo-percentiles for the caseworker-tendency instruments. Due to data confidentiality, the data has been collapsed such that a pseudo-percentile is based on the five job seekers closest to the actual percentile. Z_{1i} and Z_{2i} represent the classroom- and on-the-job-training instruments. $q(i)$ represent fully interacted job-center-unit-year fixed effects.

Table B.3: Employment Effects of Classroom vs. On-The-Job Training
Extensive Margin Employment

	Extensive-Margin Employment		
	(1)	(2)	(3)
	year 1	year 2	year 3
Classroom Training	0.07	0.11*	0.04
	(0.08)	(0.06)	(0.06)
On-The-Job Training	0.31	0.05	-0.28*
	(0.21)	(0.15)	(0.14)
Obs	167,222	167,222	150,311

Notes: The table reports the IV estimates of the effect of assignment to classroom training and on-the-job training on extensive-margin employment in a given year relative to job loss. The number of observations is reduced in Column (3) since outcomes in year 3 are unobserved for jobseekers who lost their job more recently (in 2017). All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and jobseeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table B.4: Employment Effects of Classroom vs. On-The-Job Training:
Earnings in Percent of Pre-Job-Loss

	Monthly Earnings (% of Pre-Job-Loss)		
	(1)	(2)	(3)
	year 1	year 2	year 3
Classroom Training	11.61	20.98*	16.20
	(8.60)	(11.24)	(12.22)
On-The-Job Training	12.30	4.37	-30.95
	(19.98)	(24.86)	(25.04)
Obs	167,222	167,222	150,311

Notes: The table reports the IV estimates of the effect of assignment to classroom training and on-the-job training on average monthly earnings (in % of pre-job loss earnings) in a given year relative to job loss. The number of observations is reduced in Column (3) since outcomes in year 3 are unobserved for jobseekers who lost their job more recently (in 2017). All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and jobseeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table B.5: Effect of ALMP by Exposure to Offshoring:
Extensive-Margin Employment

	Extensive Margin Employment		
	(1)	(2)	(3)
	year 1	year 2	year 3
Low risk			
Classroom Training	0.03	0.10	0.01
	(0.09)	(0.07)	(0.07)
On-The-Job Training	0.30	0.08	-0.21
	(0.21)	(0.15)	(0.14)
High risk			
Classroom Training	0.17	0.13	0.13
	(0.19)	(0.14)	(0.16)
On-The-Job Training	0.56	-0.01	-0.62
	(0.79)	(0.53)	(0.54)
Obs low risk	125,413	125,413	113,079
Obs high risk	41,809	41,809	37,232

Notes: The table reports IV estimates of the effect of assignment to classroom training and on-the-job training on extensive-margin employment in a given year relative to job loss, for job seekers at low and high risk of offshoring. Estimations are run separately for job seekers at high and low risk of offshoring. The number of observations is reduced in Column (3) since outcomes in year 3 are unobserved for jobseekers who lost their job more recently (in 2017). All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table B.6: Effect of ALMP by Exposure to Offshoring
Earnings in Percent of Pre-Job-Loss Level

	Monthly Earnings (% of Pre-Job Loss)		
	(1)	(2)	(3)
	year 1	year 2	year 3
Low risk			
Classroom Training	3.35	11.18	0.20
	(8.28)	(11.47)	(13.37)
On-The-Job Training	-3.20	-16.21	-28.55
	(19.04)	(24.65)	(25.70)
High risk			
Classroom Training	34.43	48.02	61.15**
	(26.71)	(33.95)	(28.73)
On-The-Job Training	122.21	146.03	-37.16
	(108.86)	(139.94)	(97.18)
Obs low risk	125,413	125,413	113,079
Obs high risk	41,809	41,809	37,232

Notes: The table reports IV estimates of the effect of assignment to classroom training and on-the-job training on earnings (in % of pre-job-loss earnings) in a given year relative to job loss, for job seekers at low and high risk of offshoring. Estimations are run separately for job seekers at high and low risk of offshoring. The number of observations is reduced in Column (3) since outcomes in year 3 are unobserved for jobseekers who lost their job more recently (in 2017). All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered on predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

C Identification Strategy

In this section, we present a generalized Roy model of the assignment of job seekers to training programs. We use the model to discuss the bias in estimators that assume "selection on observables only" and derive our IV strategy based on caseworker tendencies. Our model deviates from the canonical judge IV setup in two aspects. First, caseworkers may assign job seekers to *multiple* training programs. Second, job seekers are allocated to caseworkers based on an *observed rule* (day of birth).

C.1 Setup

A job seeker i may be assigned to classroom training $D_1 \in \{0,1\}$ and on-the-job training $D_2 \in \{0,1\}$. Her potential employment outcomes are

$$Y_i(D_i) = \beta_{0i} + \beta_{1i}D_{1i} + \beta_{2i}D_{2i}. \tag{9}$$

Job seekers are allocated to caseworkers who assign training programs. Let $c(b)$ denote the default rule that allocates job seekers with day-of-birth $b \in \{1,31\}$ to a caseworker c.

Caseworkers differ in their preferences for the programs (V_1, V_2) but rank individual job seekers similarly in their resistance to participate in each program (U_1, U_2). For example, a job seeker with a pending job offer may be more resistant to start training than a job seeker with no immediate job prospects. The role of caseworkers' preferences is motivated by our qualitative interviews at the job centers, during which a caseworker, for example, ascribed differences in training tendencies to reflect differences in "values" ("*værdisæt*" in Danish) of caseworkers.

Let c_i denote the caseworker assigned to job seeker i. The job seeker is assigned to training program $k \in \{1,2\}$ if

$$D_{ki} = \mathbf{1}[V_{kc_i} \geq U_{ki}], \tag{10}$$

where we normalize the marginal distributions of the resistances to be uniform, $U_{ki} \sim U[0,1]$.

We assume the preferences of a job seeker's *day-of-birth-predicted* caseworker are independent of the job seeker's training resistances and potential outcomes:

$$V_{c(b_i)} \perp\!\!\!\perp (U_i, \beta_i) \tag{11}$$

A sufficient condition for Equation (11) is that job seekers' training and employment potential are unrelated to their day of birth. Note Equation (11) allows for both differences in the general ability of caseworkers ("value added") and match effects between job seekers and caseworkers, as long as these are orthogonal to caseworker preferences for training.[40]

C.1.1 Selection into Training

The selection patterns into training programs are governed by how the job seekers' resistance to training U_i correlates with their employment potential β_i in Equation (9). For example, job seekers with worse job opportunities (low β_{i0}) may be less resistant to training (low U_{ik}). Furthermore, controlling for observables of the job seekers may not necessarily eliminate this selection bias. For example, two job seekers with identical work histories might face different job *prospects* that are not recorded in our administrative data. For example, one of the job seekers could have a pending job offer. If caseworkers learn about these *latent* job prospects during the meetings at the job center, controlling for observables will not alleviate the selection bias. These concerns motivate developing our identification strategy based on caseworker tendencies that control for unobservables of the job seekers.

C.2 Caseworker-Tendency Instruments

We instrument the assignment of job seeker i to training program k with the tendency of her *day-of-birth-predicted* caseworker to assign program k. For job seeker i, we measure the program tendencies using the assignment rates among other job seekers with the same day-of-birth predicted caseworker:

$$Z_{ki} = \frac{1}{(J(i)-1)} \sum_{j \neq i} \mathbf{1}[c(b_j) = c(b_i)] \times D_{kj}, \quad J(i) = \sum_{j} \mathbf{1}[c(b_j) = c(b_i)]. \tag{12}$$

Because job seekers are quasi-randomly allocated to caseworkers (Equation (11)), caseworkers' preferences are revealed by her observed k-tendency as she handles a large number of cases:

$$Z_{ki} \xrightarrow{p} \mathbb{P}(V_{kc(b_i)} \geq U_{kj}) = V_{kc(b_i)} \quad \text{as} \quad N_{c(b_i)} \to \infty. \tag{13}$$

We now show the caseworker-tendency instruments satisfy the *independence*, *exclusion*, and *monotonicity* conditions for the identification of local average treatment effects (LATE) of training programs (Imbens and Angrist, 1994). To ease the exposition, we first assume job centers

[40]For example, if caseworkers differ in their value added β_c, we would redefine $\beta_{0i} = \beta_{0i} + \beta_{c_i}$.

perfectly comply with the day-of-birth rule to allocate job seekers to caseworkers, $c_i = c(b_i)$. In Section C.5, we extend our identification results to allow for non-compliance with the allocation rule.

Independence and Exclusion

Because caseworker training tendencies recover preferences (V_1, V_2), it follows from Equation (11) that they satisfy the independence and exclusion restrictions:

$$Z_i \perp\!\!\!\perp \beta_i. \tag{14}$$

We purposely exclude the job seeker from her own training-tendency instruments to ensure this independence restriction holds.

Extended Monotonicity

In the canonical binary-treatment case, monotonicity requires that the instrument shifts all job seekers toward or away from the treatment in consideration. However, because job seekers face *multiple* training options, identification in our setting requires an *extended* monotonicity assumption about how instruments affect multiple training programs.

Using Equation (13), we can restate Equation (10) in terms of caseworker tendencies:

$$D_{ki} = \mathbf{1}[Z_{ki} \geq U_{ki}]. \tag{15}$$

Let $D_i(Z)$ denote the potential training assignments of job seeker i depending on the training tendencies Z of his caseworker. Equation (15) implies the assignment to classroom training and on-the-job training are two separate decisions. In particular, the assignment of job seeker i to training k depends solely on the k-tendency of her caseworker. Hence, comparing two otherwise similar caseworkers, a more k-inclined caseworker will shift all job seekers toward training program k but not alter the participation in other programs l:

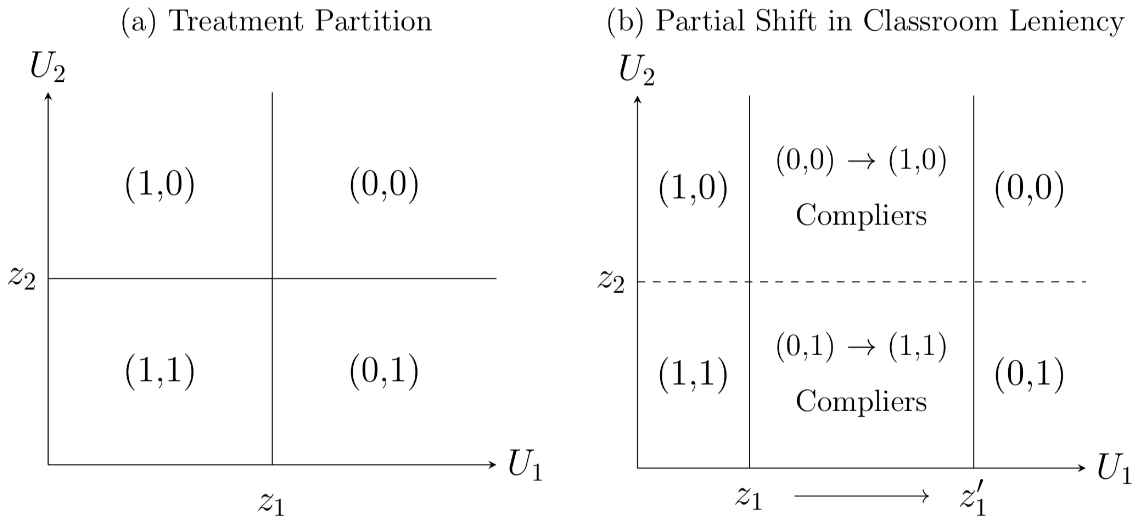
$$z'_k > z_k \implies D_{ki}(z'_k, z_l) \geq D_{ki}(z_k, z_l) \tag{16}$$

$$D_{li}(z'_k, z_l) = D_{li}(z_k, z_l), \quad k \neq l, \forall i. \tag{17}$$

The property in Equations (16)-(17) is labeled *extended monotonicity* in the literature and plays a key role in IV analysis with multiple treatments (Behaghel et al., 2013; Lee and Salanié, 2018, 2020; Bhuller and Sigstad, 2022). In particular, when evaluating treatment k, extended monotonicity allows us to collapse the analysis to the canonical single-treatment case by comparing caseworkers with similar tendencies for the other training programs l.[41]

Figure C.1 illustrates the concept of extended monotonicity. Panel (a) first illustrates how the assignment to training programs depends on job seeker resistance U and caseworker preference thresholds Z. The horizontal and vertical axes represent job seekers' resistance to classroom training and on-the-job training, respectively. The two lines represent the caseworker's preference for each training program. A job seeker is assigned to training k if her resistance is below the caseworker threshold, $U_{ki} \leq Z_{ki}$. The caseworker preference thresholds thus partition job seekers into four training combinations: $(D_1, D_2) \in \{(0,0), (1,0), (0,1), (1,1)\}$.

Figure C.1: Assignment to Training Program



Notes: This figure illustrates the concept of extended monotonicity. The horizontal and vertical axes measure job seekers' resistance to classroom training and on-the-job training, respectively. Panel (a) illustrates how job seekers are assigned to program k if their resistance U_k is lower than the caseworker preference thresholds z_k. Panel (b) shows how a partial increase in classroom training tendency $z_1 \to z_1'$, holding on-the-job training tendency fixed at z_2, shifts some job seekers toward classroom training but does *not* alter the participation in on-the-job training.

Panel (b) illustrates how a partial increase in classroom training tendency $z_1 \to z_1'$, holding on-the-job training tendency fixed at z_2, shifts some job seekers toward classroom training but does *not* alter the participation in on-the-job training. The area between z_1 and z_1' represents job

[41]In many settings with multiple treatments, we have economic theory and empirical evidence that IVs should not satisfy extended monotonicity. Mountjoy (2022) shows shorter commuting distances to community colleges divert some students from enrolling in four-year colleges. Kirkeboen et al. (2016) show crossing the admission threshold between a preferred and a next-best major changes the likelihood that students enroll in alternative majors to the next-best option.

seekers who are shifted (horizontally) into classroom training. The new trainees come from two margins: passive UI $(0,0) \rightarrow (1,0)$ and on-the-job training only $(0,1) \rightarrow (1,1)$. Importantly, the shift in classroom training tendency does *not* induce any (vertical or diagonal) shifts into or away from on-the-job training.

C.3 Local Average Treatment Effects

Proposition 1 formalizes how the extended monotonicity property is sufficient to identify LATEs of training program for jobseekers whose training assignments depend on the preferences of their caseworker. In a nutshell, extended monotonicity implies the standard single-treatment LATE analysis applies for each own-instrument once we condition on an evaluation point for the cross-instruments.

Proposition 1 (Mean Potential Outcomes of Instrument Compliers)**.**

1. *Denote job seekers who shift from treatment d to d' if assigned a caseworker with tendencies z' instead of z by*

$$\{d \rightarrow d'\}_{(z \rightarrow z')} = \{u \in [0,1]^2 \quad \textit{s/t} \quad d' = \mathbf{1}[z' \geq u],\, d = \mathbf{1}[z \geq u]\}. \tag{18}$$

2. *Define potential outcomes Y and training assignments D as in Section C.1.*

3. *The caseworker-tendency instruments developed in Section C.2 identify the mean potential outcomes for instrument compliers along all training margins as*

$$\mathbb{E}\left[Y(d'')|\{d \rightarrow d'\}_{(z_k \rightarrow z'_k, z_l)}\right] = \frac{\mathbb{E}\left[Y\mathbf{1}_{[D=d'']}|z'_k, z_l\right] - \mathbb{E}\left[Y\mathbf{1}_{[D=d'']}|z_k, z_l\right]}{\mathbb{E}\left[\mathbf{1}_{[D=d'']}|z'_k, z_l\right] - \mathbb{E}\left[\mathbf{1}_{[D=d'']}|z_k, z_l\right]}. \tag{19}$$

Having recovered the mean potential outcomes, we can calculate the treatment effects of compliers around each separate treatment margin. For example, the LATE for job seekers who are shifted from passive unemployment to classroom training is

$$\mathbb{E}\left[\beta_1|\{(0,0) \rightarrow (1,0)\}_{(z_k \rightarrow z'_k, z_l)}\right] = \mathbb{E}\left[Y(1,0) - Y(0,0)|\{(0,0) \rightarrow (1,0)\}_{(z_k \rightarrow z'_k, z_l)}\right]. \tag{20}$$

Proof. Consider caseworkers A and B, who have the same tendency to assign on-the-job training but differ in their tendency for classroom training.[42] Comparing the share of job seekers

[42]The example focuses on shifts in the classroom training instrument, keeping the on-the-job training instruments fixed. All arguments apply symmetrically to shifts in the on-the-job training instrument, keeping the classroom training instruments fixed.

assigned to classroom training by caseworkers A and B, we can estimate the share of compliers with the $(z_1^A \to z_1^B)$ shift at z_2. By splitting the counts by on-the-job-training status, we can calculate the share of compliers along each separate margin:

$$\mathbb{E}\left[D_1(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[D_1(1-D_2)|z_1^A, z_2\right] = \mathbb{P}\left[\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \tag{21}$$

$$\mathbb{E}\left[D_1 D_2|z_1^B, z_2\right] - \mathbb{E}\left[D_1 D_2|z_1^A, z_2\right] = \mathbb{P}\left[\{(0,1) \to (1,1)\}_{(z_1^A \to z_1^B, z_2)}\right]. \tag{22}$$

Second, by studying how total employment shifts across treatment cells, we can estimate the total potential outcomes of the instrument compliers:

$$\begin{aligned} &\mathbb{E}\left[YD_1(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[YD_1(1-D_2)|z_1^A, z_2\right] \\ &= \mathbb{P}\left[\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \times \mathbb{E}\left[Y(1,0)|\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \end{aligned} \tag{23}$$

$$\begin{aligned} &\mathbb{E}\left[Y(1-D_1)(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[YD_1(1-D_2)|z_1^A, z_2\right] \\ &= -\mathbb{P}\left[\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \times \mathbb{E}\left[Y(0,0)|\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \end{aligned} \tag{24}$$

$$\begin{aligned} &\mathbb{E}\left[Y(1-D_1)(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[Y(1-D_1)(1-D_2)|z_1^A, z_2\right] \\ &= \mathbb{P}\left[\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \times \mathbb{E}\left[Y(1,1)|\{(0,1) \to (1,1)\}_{(z_1^A \to z_1^B, z_2)}\right] \end{aligned} \tag{25}$$

$$\begin{aligned} &\mathbb{E}\left[Y(1-D_1)D_2|z_1^B, z_2\right] - \mathbb{E}\left[YD_1(1-D_2)|z_1^A, z_2\right] \\ &= -\mathbb{P}\left[\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] \times \mathbb{E}\left[Y(0,1)|\{(0,1) \to (1,1)\}_{(z_1^A \to z_1^B, z_2)}\right]. \end{aligned} \tag{26}$$

Finally, relating the shifts in total employment (Equations (23)-(24) and (25)-(26), respectively) to the shifts in count shares (Equations (21) and (22), respectively), we can isolate the mean potential outcomes of compliers who are shifted into classroom training if assigned to caseworker B instead of caseworker A:

$$\mathbb{E}\left[Y(1,0)|\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] = \frac{\mathbb{E}\left[YD_1(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[D_1(1-D_2)|z_1^A, z_2\right]}{\mathbb{E}\left[D_1(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[D_1(1-D_2)|z_1^A, z_2\right]} \tag{27}$$

$$\mathbb{E}\left[Y(0,0)|\{(0,0) \to (1,0)\}_{(z_1^A \to z_1^B, z_2)}\right] = \frac{\mathbb{E}\left[Y(1-D_1)(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[Y(1-D_1)(1-D_2)|z_1^A, z_2\right]}{\mathbb{E}\left[(1-D_1)(1-D_2)|z_1^B, z_2\right] - \mathbb{E}\left[(1-D_1)(1-D_2)|z_1^A, z_2\right]} \tag{28}$$

$$\mathbb{E}\left[Y(1,1)|\{(0,1) \to (1,1)\}_{(z_1^A \to z_1^B, z_2)}\right] = \frac{\mathbb{E}\left[YD_1D_2|z_1^B, z_2\right] - \mathbb{E}\left[YD_1D_2|z_1^A, z_2\right]}{\mathbb{E}\left[D_1D_2|z_1^B, z_2\right] - \mathbb{E}\left[D_1D_2|z_1^A, z_2\right]} \tag{29}$$

$$\mathbb{E}\left[Y(0,1)|\{(0,1) \to (1,1)\}_{(z_1^A \to z_1^B, z_2)}\right] = \frac{\mathbb{E}\left[Y(1-D_1)D_2|z_1^B, z_2\right] - \mathbb{E}\left[Y(1-D_1)D_2|z_1^A, z_2\right]}{\mathbb{E}\left[(1-D_1)D_2|z_1^B, z_2\right] - \mathbb{E}\left[(1-D_1)D_2|z_1^A, z_2\right]}. \tag{30}$$

□

C.3.1 Econometric Implementation

Following Proposition 1, we may identify the causal effects of classroom training by regressing the outcome variables

$$T = \{D_1 D_2, D_1(1 - D_2), (1 - D_1)D_2, (1 - D_1)(1 - D_2), .. \tag{31}$$

$$Y D_1 D_2, Y D_1(1 - D_2), Y(1 - D_1)D_2, Y(1 - D_1)(1 - D_2)\} \tag{32}$$

on the classroom-training instrument Z_1, holding the on-the-job training instrument fixed at some evaluation point z'_2:

$$T_i = \beta_0^T + \beta_1^T Z_{i1} \quad \text{for} \quad z_{i2} \in [z'_2 - \epsilon_2, z'_2 + \epsilon_2]. \tag{33}$$

Having estimated Equation (33), we may recover mean potential outcomes for classroom training $d \in \{0, 1\}$ along each margin of on-the-job training using the Wald ratios:

$$\mathbb{E}\left[Y(d, 0)|U_1^{(c,0)}(z'_2)\right] = \frac{\hat{\beta}_1^{Y(D_1=d)(1-D_2)}}{\hat{\beta}_1^{(D_1=d)(1-D_2)}} \tag{34}$$

$$\mathbb{E}\left[Y(d, 1)|U_1^{(c,1)}(z'_2)\right] = \frac{\hat{\beta}_1^{Y(D_1=d)D_2}}{\hat{\beta}_1^{(D_1=d)D_2}}. \tag{35}$$

To increase power, we can stack the point-specific evaluations in Equation (33) into a single regression, controlling flexibly for the on-the-job training instrument (e.g., using bins of bandwidth ϵ_2),

$$T_i = \beta_0^T + \beta_1^T Z_{i1} + g_{\epsilon_2}(Z_{i2}; \beta_2^T), \tag{36}$$

Note Equations (34)-(36) simplify to the standard TSLS specification if the control function $g(\cdot)$ is linear in Z_{i2}, i.e. $g(Z_{i2}; \beta_2^T) = \beta_2^T Z_{i2}$. Blandhol et al. (2022) discuss the importance of allowing for flexible controls in order to interpret TSLS estimates as LATEs. Section C.4 presents a marginal treatment effects framework that shows that the linear TSLS specification is valid if the marginal treatment effects of classroom training β_{1i} are uncorrelated with the job seekers' resistance to job training U_{2i}. In this case of no "essential heterogeneity" (Heckman et al., 2006), we only need to control for Z_{i2} to the extent that the instruments are correlated. Appendix Figure D.1 shows that Z_1 and Z_2 are largely orthogonal in our data, alleviating this concern for identification.

In practice, we first estimate the standard TSLS specification to facilitate comparisons to how prior papers have handled multiple treatments in judge IV setups (Bhuller et al., 2020;

Autor et al., 2015). In a second step, we follow Mountjoy (2022) and show robustness to estimating the specification around an evaluation point z_2' for the on-the-job-training instrument.

C.4 Marginal Treatment Effects

The patterns of selection into training depend on the correlation between job seekers' resistance to training U_i and their potential employment outcomes β_i. To make inferences about these selection patterns, we follow Mogstad et al. (2018) and impose shape restrictions on the marginal treatment response (MTR) and marginal treatment effect (MTE) functions. In practice, we assume an additively separable and linear specification in the quantile of the distribution of suitability for training:

$$\mathbb{E}[\beta_{i0}|U_i] = \alpha_{00} + \alpha_{01}U_1 + \alpha_{02}U_2 \tag{37}$$

$$\mathbb{E}[\beta_{i1}|U_i] = \alpha_{10} + \alpha_{11}U_{1i} \tag{38}$$

$$\mathbb{E}[\beta_{i2}|U_i] = \alpha_{20} + \alpha_{21}U_{2i}. \tag{39}$$

A positive value of the α_{0k} parameter captures if job seekers with worse job opportunities select into training (negative "selection on levels"). A negative value of the α_{1k} parameter captures if caseworkers prioritize training for job seekers who have the most to gain from the programs (positive "selection on gains").[43]

C.4.1 Estimation

Given the shape restrictions specified in (37)-(39), we can write the employment outcome for job seekers assigned to a caseworker with leniency Z, as a function of second-order polynomials in the propensity score with respect to classroom and on-the-job training. Namely, plugging

[43]That is, we allow potential employment outcomes to depend linearly on unobserved resistance to training. We have good reasons to believe job seekers opt into training based on their potential outcomes. For instance, a job seeker with high potential employment, for example, due to a pending job offer, may be resistant to start in classroom training, regardless of the training tendency of her caseworker. This outcome would be consistent with Figure A.4.(a), showing that never-takers of training have higher employment rates than compliers not assigned to training.

(37)-(39) into (9), we get

$$\begin{aligned}\mathbb{E}[Y_i|Z_i] &= \int_0^1 \int_0^1 (\alpha_{00} + \alpha_{01}U_1 + \alpha_{02}U_2) dU_1 dU_2 \\ &+ \int_0^{\hat{D}_{i1}} (\alpha_{10} + \alpha_{11}U_1) dU_1 + \int_0^{\hat{D}_{i2}} (\alpha_{20} + \alpha_{21}U_2) dU_2 \\ &= \beta_0 + \beta_{11}\hat{D}_{i1} + \beta_{21}\hat{D}_{i2} + \beta_{12}\hat{D}_{i1}^2 + \beta_{22}\hat{D}_{i2}^2 \end{aligned} \tag{40}$$

with

$$\beta_0 = \alpha_{00} + \frac{\alpha_{01} + \alpha_{02}}{2}, \quad \beta_{11} = \alpha_{10}, \quad \beta_{12} = \frac{\alpha_{11}}{2}, \quad \beta_{21} = \alpha_{20}, \quad \beta_{22} = \frac{\alpha_{21}}{2}, \tag{41}$$

where we have replaced caseworker preferences Z by the propensity score by normalizing the marginal distributions of the resistances to be uniform, $U_{ki} \sim U[0,1]$. We estimate the MTE functions based on the common support of the propensity scores for treated and non-treated job seekers. Note Equation (41) simplifies to the standard TSLS estimator in Equations (2)-(3) if MTEs are constant ($\alpha_{11} = \alpha_{21} = 0$).

C.4.2 Recovering Target Parameters

Following Mogstad et al. (2018), we can convert the estimated MTR functions into a host of parameters of interest. For example, the average treatment effect (ATE) for classroom training is

$$\text{ATE}_1 = \int_0^1 (\hat{\alpha}_{10} + \hat{\alpha}_{11}U_1) dU_1 = \hat{\alpha}_{10} + \frac{\hat{\alpha}_{11}}{2}. \tag{42}$$

C.5 Non-Compliance with Caseworker Allocation Rule

In this extension, we allow job centers to deviate from their predicted quasi-random rule for allocating job seekers to caseworkers, $c_i \neq c(b_i)$. We clarify the conditions under which our birthday-predicted caseworker-tendency instruments satisfy the relevance, independence, exclusion, and monotonicity conditions for the identification of LATEs.

Independence and Exclusion

To assess independence and exclusion, we note the variation in the caseworker-tendency instrument (Equation (12)) comes solely from the birthdays of job seekers. Hence, the instruments satisfy the independence and exclusion criteria if job seekers' training and employment potentials are unrelated to their birthday in the month:

$$b_i \perp\!\!\!\perp (U_i, \beta_i). \tag{43}$$

Relevance and Monotonicity

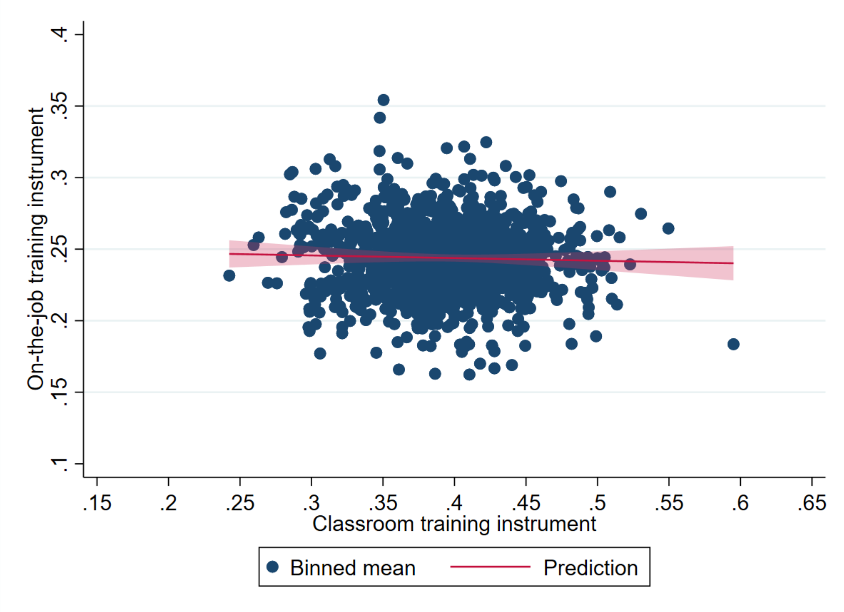
Monotonicity requires that workers with a more training-inclined *predicted* caseworker also end with a (weakly) more training-inclined *realized* caseworker:

$$V_{kc(b_i)} > V_{kc(b_j)} \implies V_{kc_i} \geq V_{kc_j},\ V_{lc_i} = V_{lc_j}, \quad k \neq l. \tag{44}$$

The "monotonic compliance" condition in Equation (44) implies the tendency instruments are also relevant. In Section 5, we provide empirical support for monotonic compliance. First, we show no correlation exists between the training tendency of a job seeker's predicted caseworker and the rate of compliance with the default allocation rule. Second, we show that if deviating from the allocation rule, job seekers with more training-inclined predicted caseworkers are *not* reassigned to less training-inclined realized caseworkers.

D Instrument Diagnostics

Figure D.1: Correlation between Caseworker Tendencies



Notes: This figure shows the correlation of caseworker training tendencies. Each bin represents the caseworker tendency for 100 job seekers in the sample demeaned by job-center-unit-year fixed effects. The red line represents the linear prediction obtained by OLS regression of the demeaned on-the-job-training instrument on the demeaned classroom-training instrument. The shaded areas represent 95% confidence intervals based on standard errors clustered at the level of the predicted caseworker. The regression line has a slope of -0.018 (with a t-stat of -0.61).

D.1 Relevance

Table D.1: First-Stage Estimates

	(1) Classroom Training	(2) On-The-Job Training
Classroom Training	0.38***	0.01
	(0.04)	(0.02)
On-The-Job Training	0.05	0.21***
	(0.03)	(0.05)
Obs	167,222	167,222
F-stat	55	10
Own-instrument pct 1	0.29	0.16
Own-instrument pct 99	0.50	0.33
Complier share	0.08	0.04

Notes: This table reports the first stage coefficients from estimations of Equation (2). All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. The complier share with respect to treatment k is obtained in two steps. First, we approximate a k-restrained and k-inclined caseworker by percentile 1 and 99 of the own-instrument demeaned by job-center-unit-year fixed effects (and the cross-instrument). Second, the first-stage coefficient (on the own-instrument) is multiplied by the difference between the two percentiles. See additional details in Online Appendix OA2.2. *p<0.10 ** p<0.05 *** p<0.01.

Table D.2: First-Stage Estimates of Classroom-Training Assignment

	Assignment to Classroom Training						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Classroom Training Instrument	1.00***	1.00***	0.95***	0.67***	0.66***	0.38***	0.38***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.04)
On-The-Job Training Instrument	-0.00	-0.00	-0.02	-0.09***	-0.06*	0.05	0.05
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)
Obs	167,222	167,222	167,222	167,222	167,222	167,222	167,222
F-stat	7697.08	5318.16	4728.76	402.17	372.08	97.64	54.64
Caseworker[a]	Realized	Predicted	Predicted	Predicted	Predicted	Predicted	Predicted
Leave-out[b]	-	-	spell	spell	spell	spell	spell
FE's[c]	-	-	-	j × u	j × u, y	j × u × y	j × u × y
Clustering[d]	-	-	-	-	-	-	yes

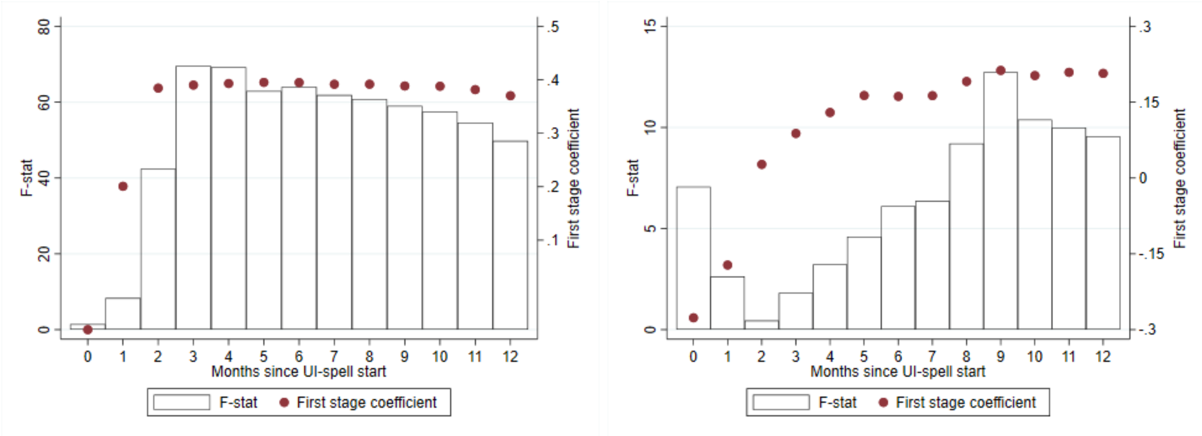
Notes: This table reports the first stage coefficients from estimations of Equation (2). [a] Caseworker refers to whether the instruments are based on the *realized* or the *day-of-birth-predicted* caseworker. [b] Leave-out refers to whether the instrument is defined as the leave-out spell mean (as opposed to the mean). [c] FE's refers to whether the regression includes job-center-unit FE's (j×u), job-center-unit and year FE's (j×u, y), or fully interacted job-center-unit-year FE's (j×u×y). [d] Clustering refers to whether standard errors are two-way clustered at the predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table D.3: First-Stage Estimates of On-The-Job Training Assignment

	Assignment to On-The-Job Training						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Classroom Training Instrument	-0.00	0.00	-0.01	-0.04*	-0.04**	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
On-The-Job Training Instrument	1.00***	1.00***	0.88***	0.49***	0.45***	0.21***	0.21***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)	(0.05)
Obs	167,222	167,222	167,222	167,222	167,222	167,222	167,222
F-stat	4053.81	2443.49	1916.51	131.04	107.03	18.97	9.99
Caseworker[a]	Realized	Predicted	Predicted	Predicted	Predicted	Predicted	Predicted
Leave-out[b]	-	-	spell	spell	spell	spell	spell
FE's[c]	-	-	-	j × u	j × u, y	j × u × y	j × u × y
Clustering[d]	-	-	-	-	-	-	yes

Notes: This table reports the first stage coefficients from estimations of Equation (2). [a] Caseworker refers to whether the instruments are based on the *realized* or the *day-of-birth-predicted* caseworker. [b] Leave-out refers to whether the instrument is defined as the leave-out spell mean (as opposed to the mean). [c] FE's refers to whether the regression includes job-center-unit FE's (j×u), job-center-unit and year FE's (j×u, y), or fully interacted job-center-unit-year FE's (j×u×y). [d] Clustering refers to whether standard errors are two-way clustered at the predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Figure D.2: First-Stage Robustness to Time Horizon



(a) Assignment to Classroom Training

(b) Assignment to On-The-Job Training

Notes: The figure shows coefficients from first-stage regressions of assignment to classroom training or on-the-job training within a given time horizon, on the two caseworker-tendency instruments (defined with the same time horizon). The time horizon refers to the month in which the assigned training program is set to start. Red dots represent the first-stage coefficient on the own instrument, i.e. in panel (a), the red dots represent coefficients on the classroom training instrument; in panel (b) red dots represent coefficients on the on-the-job training instrument. Bars represent the F-stat on both caseworker-tendency instruments. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels.

D.2 Monotonicity

Table D.4: Testing for Monotonicity with the Baseline Instrument

	Assignment to Classroom Training			
	q1	q2	q3	q4
Classroom Training Instrument	0.20***	0.39***	0.45***	0.47***
	(0.06)	(0.06)	(0.07)	(0.06)
On-The-Job Training Instrument	0.05	0.11	-0.00	0.01
	(0.05)	(0.09)	(0.08)	(0.09)
Obs	41,799	41,803	41,803	41,800
Dep var Mean	0.21	0.33	0.43	0.58
Dep var sd	0.41	0.47	0.50	0.49
F-stat (instruments)	6.35	19.30	20.57	35.21
P-value (F-stat)	0.002	0.000	0.000	0.000

Notes: This table implements a monotonicity test for assignment to classroom training based on the baseline instruments. The sample is partitioned into quartiles based on predicted assignment to classroom training, resulting from an OLS regression of assignment to classroom training on predetermined job-seeker characteristics. These characteristics include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), the typical occupation over the career (3-digit ISCO08), and an indicator for being in high risk of offshoring. Each column represents the coefficients from a quartile-specific first-stage regression based on the baseline instruments and including job-center-unit-year fixed effects. Standard errors (in parenthesis) are two-way clustered at the predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table D.5: Testing for Monotonicity with the Baseline Instrument

	Assignment to On-The-Job Training			
	q1	q2	q3	q4
Classroom Training Instrument	-0.01	0.01	0.05	0.01
	(0.04)	(0.04)	(0.05)	(0.05)
On-The-Job Training Instrument	0.10	0.13*	0.14*	0.35***
	(0.08)	(0.07)	(0.08)	(0.07)
Obs	41,802	41,803	41,806	41,805
Dep var Mean	0.12	0.20	0.27	0.39
Dep var sd	0.32	0.40	0.44	0.49
F-stat (instruments)	0.78	1.67	1.77	12.10
P-value (F-stat)	0.458	0.189	0.171	0.000

Notes: This table implements a monotonicity test for assignment to on-the-job training based on the baseline instruments. The sample is partitioned into quartiles based on predicted assignment to on-the-job training, resulting from an OLS regression of assignment to on-the-job training on predetermined job seeker characteristics. These characteristics include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08), and an indicator for being in high risk of offshoring. Each column represents the coefficients from a quartile-specific first-stage regression based on the baseline instruments and including job-center-unit-year fixed effects. Standard errors (in parenthesis) are two-way clustered at the predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table D.6: Testing for Monotonicity with the Reverse Instrument

	Assignment to Classroom Training			
	q1	q2	q3	q4
(Rev.) Classroom Training Instrument	0.12***	0.29***	0.39***	0.09***
	(0.04)	(0.06)	(0.07)	(0.03)
(Rev.) On-The-Job Training Instrument	-0.00	-0.10	0.04	-0.06**
	(0.03)	(0.08)	(0.08)	(0.03)
Obs	41,523	41,803	41,803	38,897
Dep var Mean	0.21	0.33	0.43	0.57
Dep var sd	0.41	0.47	0.50	0.50
F-stat (instruments)	4.81	10.98	18.15	3.67
P-value (F-stat)	0.009	0.000	0.000	0.026

Notes: This table implements a monotonicity test for assignment to classroom training based on "reversed" instruments. The sample is partitioned into quartiles based on predicted assignment to classroom training, resulting from an OLS regression of assignment to classroom training on job-seeker predetermined characteristics. These characteristics include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08), and an indicator for being in high risk of offshoring. For each quartile, a "reversed" instrument is constructed: using the average training assignment probability for job seekers with the same predicted caseworker but belonging to one of the other three quartiles. Each column represents the coefficients from a quartile-specific first-stage regression based on these 'reversed' instruments and including job-center-unit-year fixed effects. Standard errors (in parentheses) are two-way clustered at the predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

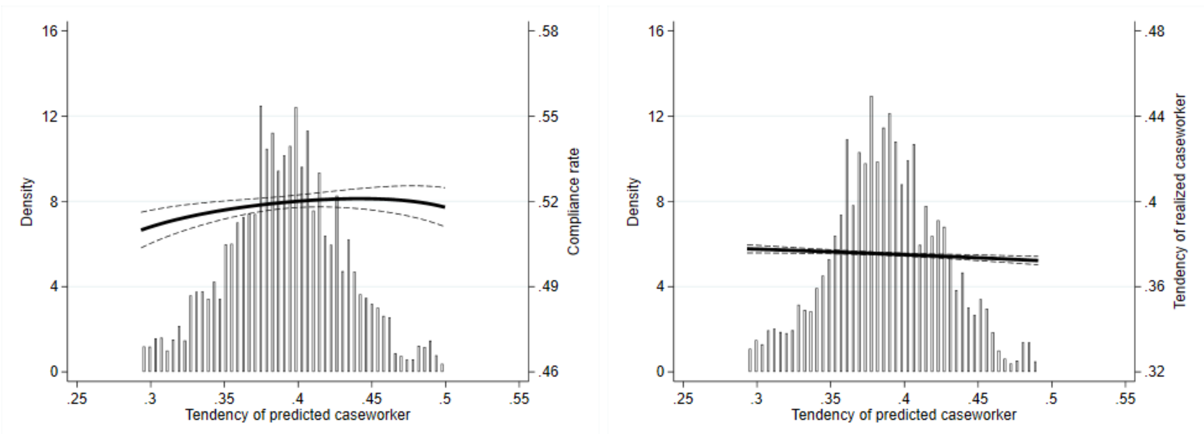
Table D.7: Testing for Monotonicity with the Reverse Instrument

	Assignment to On-The-Job Training			
	q1	q2	q3	q4
(Rev.) Classroom Training Instrument	-0.00	0.01	0.03	0.04
	(0.03)	(0.04)	(0.05)	(0.05)
(Rev.) On-The-Job Training Instrument	0.10*	0.19***	0.20***	0.11*
	(0.05)	(0.06)	(0.06)	(0.06)
Obs	41,802	41,803	41,806	41,805
Dep var Mean	0.12	0.20	0.27	0.39
Dep var sd	0.32	0.40	0.44	0.49
F-stat (instruments)	1.80	5.69	5.35	4.51
P-value (F-stat)	0.166	0.004	0.005	0.011

Notes: This table implements a monotonicity test for assignment to on-the-job training based on "reversed" instruments. The sample is partitioned into quartiles based on predicted assignment to on-the-job training, resulting from an OLS regression of assignment to on-the-job training on all predetermined job seeker characteristics. These characteristics include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08), and an indicator for being in high risk of offshoring. For each quartile, a "reversed" instrument is constructed: using the average training assignment probability for job seekers with the same predicted caseworker but belonging to one of the other three quartiles. Each column represents the coefficients from a quartile-specific first-stage regression based on these reversed instruments and including job-center-unit-year fixed effects. Standard errors (in parentheses) are two-way clustered at the predicted caseworker and job-seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

D.2.1 Monotonic Compliance to Caseworker Allocation Rule

Figure D.3: Monotonic Compliance to Classroom Instrument



(a) Compliance rate across caseworker tendency (b) Caseworker tendency if non-compliance

Notes: Panel (a) of this figure shows the correlation between the day-of-birth compliance rate and the classroom-training tendency of the predicted caseworker for all job seekers in the sample. Panel (b) shows the correlation between the classroom-training tendencies of the realized and predicted caseworker for jobseekers that do not comply with the day-of-birth allocation rule.. All plotted values in the figure have been demeaned by job-center-unit-year fixed effects and the on-the-job-training tendency for the predicted caseworker. In both panels, the bars represent the distribution of the classroom-training tendency for the *predicted* caseworker (excluding top and bottom 1%). The black line in panel (a) represents the coefficients from a local linear regression of the day-of-birth compliance rate on the classroom-training tendency of the *predicted* caseworker. The black line in panel (b) represents the coefficients from a local linear regression of the classroom-training tendency of the *realized* caseworker on the demeaned classroom-training tendency of the *predicted* caseworker. The local linear regressions are based on an Epanechnikov kernel with bandwidth 0.1. The day-of-birth compliance rate indicates whether the actual caseworker corresponds to the day-of-birth predicted caseworker. Dotted lines represent 95% confidence intervals.

D.3 Exclusion Test

Table D.8 shows caseworkers' on-the-job-training tendencies are uncorrelated with proxies for their experience and caseloads.

Table D.8: Experience, Caseload Size, and On-The-Job-Training Tendency

	Summary			On-The-Job Training Tendency	
	obs	mean	sd	coef×100	se×100
Caseload size					
- Meetings/year	576	430.6	235.6	0.0011	(0.0011)
- Jobseeker assignments/year	576	117.7	64.1	-0.0032	(0.0033)
Experience					
- Years	576	1.2	1.2	0.2864*	(0.1728)
- Meetings	576	868.9	783.8	0.0005*	(0.0003)
- Jobseeker assignments	576	255.0	237.2	0.0009	(0.0008)
Obs	576	1221		1221	

Notes: This table shows correlations between caseworkers' on-the-job-training tendencies and caseworkers' characteristics, e.g. their experience and caseload size, Columns (1)-(3) report the number of unique caseworkers along with the mean and standard deviation of experience and caseload size across caseworkers and years. Columns (4)-(5) report the coefficient and standard errors from a regression of caseworker experience or caseload size in a given year on the caseworker's on-the-job-training tendency in the same year, while controlling for job-center-unit-year fixed effects as well as the caseworker's classroom-training tendency. Standard errors (in parentheses) are two-way clustered on the predicted caseworker and job seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

Table D.9 shows a statically significant – yet *economically insignificant* – correlation between caseworkers' on-the-job-training tendencies and the timing and frequency of meetings with job seekers. Namely, the coefficients suggest that going from the least to the most on-the-job-training-inclined caseworker moves the job seeker's first meeting forward by 0.17 weeks (about one day; 3.4% relative to the mean) and increases the meeting frequency by 0.0085 meetings per week (about 6% relative to the mean).[44] Appendix Figure E.3b shows our conclusions are robust to controlling for the timing and frequency of meetings; we find no robust employment gains from assignment to on-the-job training.

[44]We approximate the least and most on-the-job-training-inclined caseworker by percentile 1 and 99 of the on-the-job-training instrument within a job-center-unit-year (see Appendix Table B.2)

Table D.9: Meetings, and On-The-Job-Training Tendency

		Within j-u-y		On-The-Job Training Tendency	
	obs	mean	sd	coef	se
Timing of first meeting	167,222	4.99	3.70	-1.00**	(0.50)
Frequency of meetings (per week)	167,222	0.15	0.10	0.05***	(0.02)
Observations	167,222				

Notes: This table shows correlations between caseworkers' on-the-job-training tendencies and meeting behaviors with individual job seekers. "Timing of first meeting" refers to the first week the job seeker meets with a caseworker (first week relative to UI-spell start). "Meetings/week" refers to the number of meetings per week during the first six months of unemployment. Columns (1)-(3) report the number of observations along with the mean and standard deviation of the meeting behavior across job seekers within job-center-unit-years. Columns (4)-(5) report the coefficient and standard errors from separate regressions of a given meeting behavior on the caseworker's on-the-job-training tendency. All regressions include job-center-unit-year fixed effects as well as a control for the caseworker's classroom-training tendency. Standard errors (in parentheses) are two-way clustered on the predicted caseworker and job seeker levels. *p<0.10 ** p<0.05 *** p<0.01.

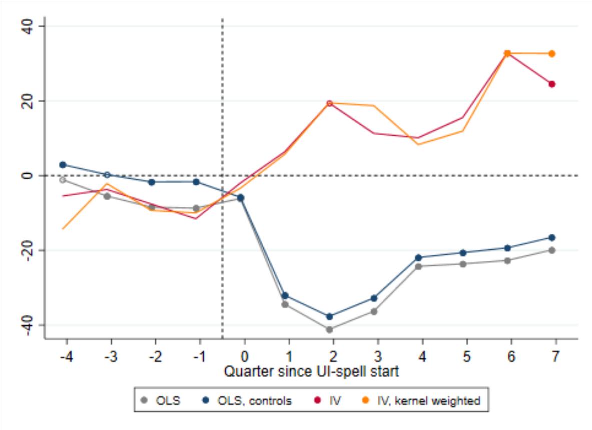
E Robustness Analysis

E.1 Local IV

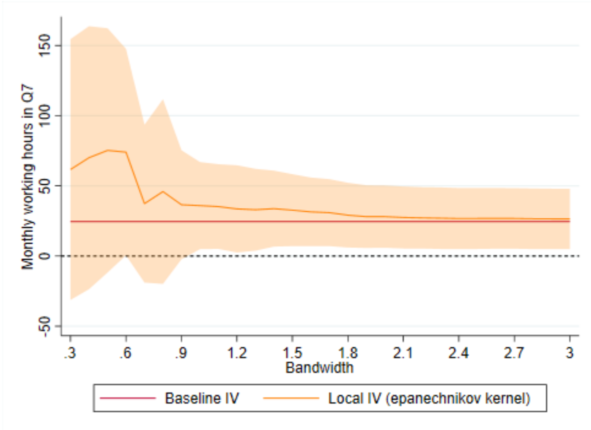
Our baseline TSLS specification controls linearly for the assignment to on-the-job training. Yet, our theory in Section C.3.1 and Blandhol et al. (2022) highlight the role of *flexible* controls for interpreting TSLS estimates as LATEs. In this section, we show our results are robust to estimating our TSLS specification around an evaluation point z_2' for the on-the-job-training instrument, following (Mountjoy, 2022).

In particular, we estimate our TSLS specification using an Epanechnikov kernel to weigh all observations around the mean of the on-the-job training instrument (within bandwidth 1.5 SDs). Figure E.1a plots this local IV estimate along with our baseline IV estimate. The figure shows our results are robust to a local estimation of Equations (2)-(3). Figure E.1b suggests this result holds across a wide range of kernel bandwidths.

Figure E.1: Local IV



(a) Local IV



(b) Effect in Q7 against bandwidth choice

Notes: This figure plots the effect of assignment to classroom training on average monthly working hours in a given quarter relative to job loss. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. In Panel (a), the orange line represents the local IV estimate obtained by using an Epanechnikov kernel (with bandwidth 1.5) to weigh all observations around the mean of the on-the-job-training instrument. The red line represents the baseline IV estimate and the gray (blue) line represents the OLS estimate (with controls). Full (hollow) dots indicate significance at the 5% (10%) level. In Panel (b), the orange line shows the local IV estimate obtained for a given choice of bandwidth on the Epanechnikov kernel. The light-orange area shows the 95% confidence intervals on the local IV estimate. The red line again represents the baseline IV estimate.

E.2 Separate Treatment Margins

Given extended monotonicity, our classroom tendency instrument shifts compliers into classroom training from two margins: from passive UI, $(0,0) \to (1,0)$, and from on-the-job training only, $(0,1) \to (1,1)$.

Following Section C.3.1, we can estimate the share of compliers (and their LATEs) along each of the treatment margins by regressing the treatment status indicators (and their interactions with outcomes) on the training tendency instruments:

$$T \in \{D_1(1-D_2), (1-D_1)(1-D_2), D_1D_2, (1-D_1)D_2\}$$

$$T_i = \beta^T_{q(i)} + \beta^T_1 Z_{1i} + \beta^T_2 Z_{2i} \tag{45}$$

$$Y_iT_i = \beta^{YT}_{q(i)} + \beta^{YT}_1 Z_{1i} + \beta^{YT}_2 Z_{2i}, \tag{46}$$

where $\beta_{q(i)}$ denote job center-unit-year fixed effects.

Along the passive UI margin, for example, the share of compliers and their LATE are

$$\mathbb{P}\left[(0,0) \to (1,0)\right] = \hat{\beta}_1^{D_1(1-D_2)} \tag{47}$$

$$\mathbb{E}\left[Y(1,0) - Y(0,0) | (0,0) \to (1,0)\right] = \frac{\hat{\beta}_1^{YD_1(1-D_2)}}{\hat{\beta}_1^{D_1(1-D_2)}} - \frac{\hat{\beta}_1^{Y(1-D_1)(1-D_2)}}{\hat{\beta}_1^{(1-D_1)(1-D_2)}}. \tag{48}$$

Table E.1 decomposes the baseline first stage (Column (1)) by the treatment margins (Columns (2) and (3)). The decomposition delivers two insights. First, the majority of compliers with the classroom-training instrument come from the passive margin (6%) rather than the on-the-job training margin (2%). Second, the instruments have twice the statistical power (F-stat) at the passive margin relative to the on-the-job training margin.

Table E.1: First-Stage Estimates by Margin

	Classroom Training by Margin		
	(1) Any Margin $(0,x) \rightarrow (1,x)$	(2) Passive Margin $(0,0) \rightarrow (1,0)$	(3) On-The-Job-Training Margin $(0,1) \rightarrow (1,1)$
Classroom Training Instrument	0.38***	0.27***	0.11***
	(0.04)	(0.03)	(0.02)
On-The-Job Training Instrument	0.05	-0.02	0.07***
	(0.03)	(0.04)	(0.03)
Obs	167,222	167,222	167,222
Obs by cell	65,082	45,338	19,744
F-stat	54.64	41.63	17.24
compliers	0.08	0.06	0.02

Notes: First-stage estimates for classroom training by on-the-job-training margin. Column (1) uses assignment to classroom training as the dependent variable; that is, it corresponds to our baseline first stage estimate (Column (1) of Table D.1). Columns (2) and (3) uses assignment to classroom training interacted with indicators for non-assignment and assignment to on-the-job training as the dependent variable. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. The complier share at the bottom of the table represents the share of compliers in the population who are shifted into classroom training by our classroom-training instrument. To calculate this share, we re-scale the coefficient on the classroom instrument by the difference in classroom-tendency for a classroom-restrained and a classroom-inclined caseworker (approximated by percentile 1 and 99 on the classroom-training instrument). *p<0.10 ** p<0.05 *** p<0.01.

Figure E.2 investigates whether our findings for the employment effects of classroom training are robust to focusing on the main treatment margin from passive UI, $(0,0) \rightarrow (1,0)$. The red line represents our baseline IV estimate, which captures the average employment effect across margins. The solid orange line shows the LATE along the passive UI margin, and the dashed orange line represents a local estimation around the mean value of the on-the-job training instrument. The figure shows that our conclusions about the positive employment effects of classroom training are robust to focusing on the compliers shifted from the passive margin into classroom training.

Figure E.2: Employment Effect by Treatment Margin

Notes: This figure shows the effect of assignment to classroom training on monthly working hours in a given quarter relative to job loss, by treatment margins. The red line represents the baseline IV estimate. The solid orange line represents the employment effect for compliers shifted from passive UI and into classroom training, $(0,0) \rightarrow (1,0)$. To facilitate comparison with the baseline estimate, these estimates are obtained while controlling linearly for the on-the-job-training instrument. The dashed orange line represents a *local* version of the solid orange line: the estimates are obtained using an Epanechnikov kernel (with bandwidth 1.5) to weigh all observations around the mean of the on-the-job-training instrument. All regressions include job-center-unit-year fixed effects. This figure shows no indications of statistical significance.

E.3 Control for Meeting Timing and Frequency

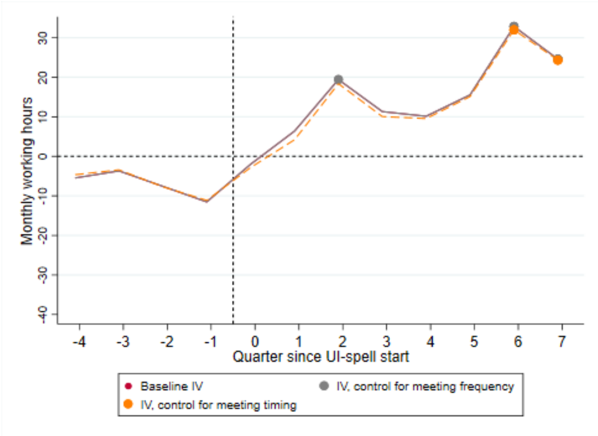
As discussed in section 5.4, one violation of the exclusion restriction is if caseworker-classroom-training tendencies affect employment rates through more frequent or earlier meetings. We test this possibility by re-estimating our main regressions while controlling for the frequency with which the job seeker meets with her caseworker and the timing of the first meeting.[45,46] Because the meeting frequency (and timing) is endogenous, we instrument the meeting frequency (timing) of job seeker i with the caseworker's general meeting frequency (timing), measured as a leave-out mean.

Figure E.3 presents the baseline IV estimate of the effect of assignment to training along with the IV estimate obtained while controlling for meeting frequency and timing. Evidently, our IV estimate for classroom training (Panel (a)) is robust to the inclusion of these controls. We take this finding as evidence in support of the exclusion restriction.

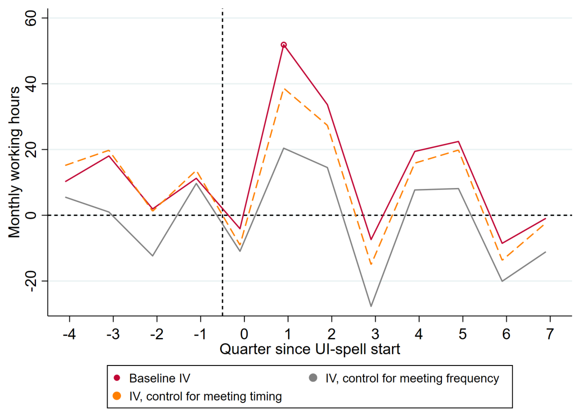
[45]We define meeting frequency for job seeker i as the number of caseworker meetings per week of unemployment. If the job seeker's UI spell is longer than 26 weeks, we only consider the first 26 weeks because meeting-frequency requirements change after 26 weeks. This measure includes all caseworker meetings held, regardless of the participating caseworker.

[46]Timing of the meeting is measured as the number of weeks between the UI spell start and the first meeting

Figure E.3: Controlling for Meeting Frequency & Timing



(a) Classroom Training



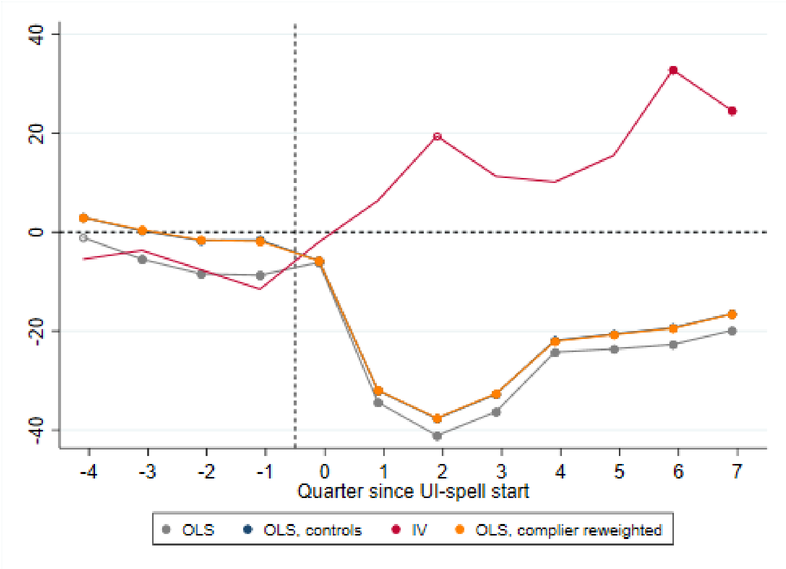
(b) On-The-Job Training

Notes: This figure shows the IV-estimate of the effect of assignment to classroom training (Panel (A)) and on-the-job training (Panel (b)) on average monthly working hours in a given quarter relative to job loss. The red line represents the baseline IV estimate. The orange and gray lines represent the IV estimate obtained by further controlling for caseworker meeting frequency and timing of the first meeting. Both of these controls are instrumented by corresponding leave-out means. All regressions include fully interacted job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

E.4 Complier-Characteristic Reweighted OLS

To test for effect heterogeneity across compliers and non-compliers, we assign all job seekers in the sample a weight according to their similarity with compliers in terms of observable characteristics. The weights are obtained by first partitioning the sample into eight subgroups based on the job seekers' (i) unemployment status in year t-1 and (ii) predicted probability of assignment to classroom training based on her predetermined characteristics. These characteristics include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08). Second, we compute the share of compliers in each subgroup and we then assign each job seeker a weight that reflects her similarity with compliers. We use these weights to estimate a complier re-weighted OLS. Figure E.4 shows the complier re-weighted OLS is highly similar to the standard OLS, suggesting effect heterogeneity based on observables *not* is driving the difference between IV and OLS.

Figure E.4: Complier-Characteristic Reweighted OLS



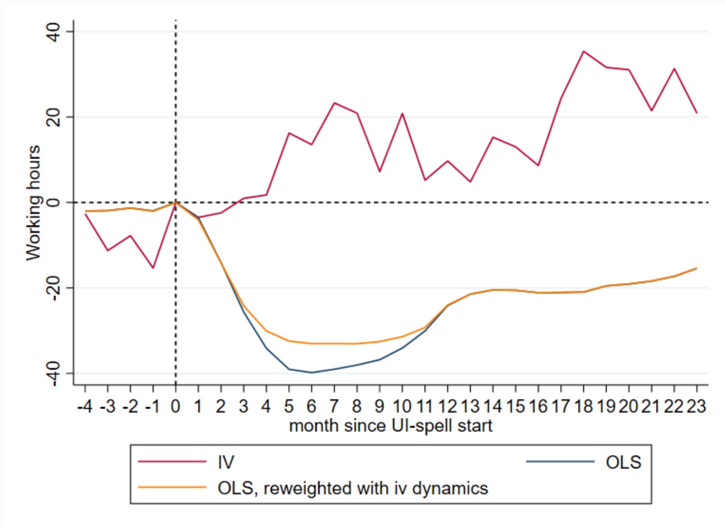
Notes: This figure shows the effect of assignment to classroom training on average monthly working hours in a given quarter relative to job loss. The red line represents the baseline IV estimate, and the blue line represents the OLS estimate (with controls). These controls include socio-demographics and labor market history from Table 3 as well as education levels (1-digit ISCED; 9 categories), previous industry (NACE level-1; 21 categories), and the typical occupation over the career (3-digit ISCO08). The orange line represents a complier re-weighted OLS. To obtain the complier weights, the population is partitioned into eight subgroups according to (i) the job seeker's unemployment status in year t-1 and (ii) her predicted probability of assignment to classroom training (based on the aforementioned set of controls). All job seekers are then assigned a weight according to the share of compliers in the subgroup. To calculate the share of compliers, the subgroup-specific first stage was re-scaled by p1 and p99 on the own-instrument for classroom training. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

E.5 OLS Reweighted by IV Training Dynamics

Figure E.5 plots three series based on our OLS and IV estimates for training-state probabilities (γ_{1t}^{s}), potential outcomes in a given state (β_{1t}^{1s}), and counterfactual outcomes (β_{0t}). We use these inputs to compute the average effect of assignment to classroom training.[47] Using the IV estimates for all three inputs, we obtain the baseline IV estimate. Correspondingly, using the OLS estimates for all three inputs, we obtain the OLS estimate. These are plotted in red and blue in Figure E.5. We also compute a hybrid estimate that uses the OLS estimate for potential outcomes (β_{1t}^{1s} and β_{0t}) and the IV estimate for training-state probabilities (γ_{1t}^{s}). That is, the hybrid corresponds to reweighting our baseline OLS estimate by the IV training dynamics. It is depicted in orange in the figure. Evidently, the baseline OLS and OLS reweighted by IV training dynamics are similar.

[47]We plug the estimates for γ_{1t}^{s}, β_{1t}^{1s}, and β_{0t} into Equation (4)

Figure E.5: OLS Reweighted by IV Dynamics

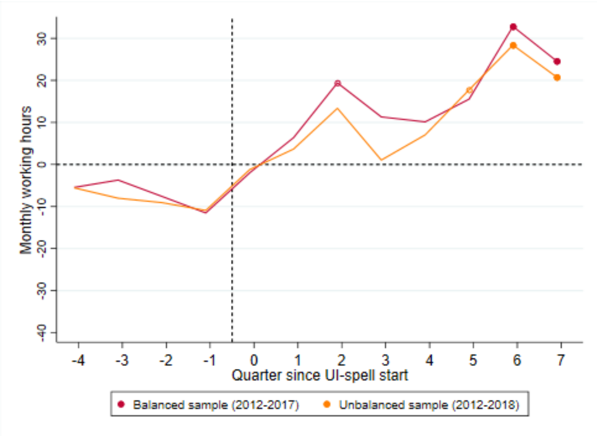


Notes: This figure plots the effect of assignment to classroom training on average monthly working hours in a given quarter relative to job loss. The effects are computed based on our OLS and IV estimates for γ_{1t}^{s}, β_{1t}^{1s}, and β_{0t}, according to equation (4). The red line represents the effects computed based on IV estimates only, the blue line represents the effects based on OLS estimates only. The orange line represents a hybrid: it is computed based on IV estimates for γ_{1t}^{s} and OLS estimates for β_{1t}^{1s} and β_{0t}. This figure shows no indication of statistical significance.

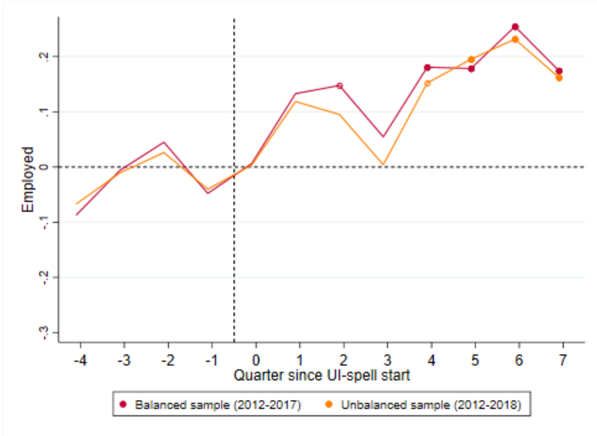
E.6 Balanced vs. Unbalanced Sample

Figure E.6 shows our results are robust to extending our current *balanced* sample of jobseekers, who became unemployed from 2012 to 2017, with an additional year. Adding job seekers who became unemployed in 2018 implies that the sample becomes unbalanced; we do not observe outcomes in the second year after job loss for job seekers from 2018. That said, the figure shows that our results are robust to extending the sample to job seekers who also became unemployed in 2018.

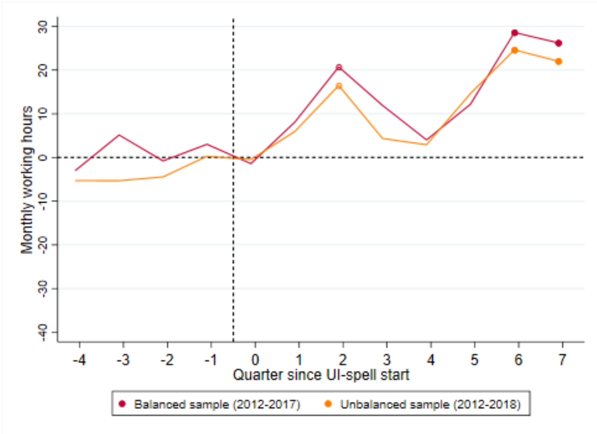
Figure E.6: Effect of Classroom Training Based on a Balanced vs. Unbalanced Sample



(a) Working Hours



(b) Extensive-Margin Employment



(c) Earnings

Notes: This figure plots the IV estimates of the effect of assignment to classroom-training programs on average monthly working hours, extensive margin employment, and earnings, in a given quarter relative to job loss. The red line represents our baseline estimates. These are based on a *balanced* sample of job seekers for whom we observe labor market outcomes over the full first two years after job loss, i.e., job seekers who became unemployed from 2012 to 2017. The orange line represents estimates based on the full, *unbalanced* sample of job seekers who became unemployed from 2012 to 2018. We do not observe the labor market outcomes in the full first two years after job loss for job seekers who became unemployed in 2018. All regressions include job-center-unit-year fixed effects. Standard errors are two-way clustered at the predicted caseworker and job-seeker levels. Full (hollow) dots indicate significance at the 5% (10%) level.

F Threat, Lock-in, and Post-program Effects

F.1 Training States

We split job seekers assigned to classroom training (assignees) into four mutually exclusive states. First, we identify the following dates for all assignees:

t_1: Assignment to training ($\sim$ First caseworker meeting)

t_2: Start of training

t_3: End of training

t_e: End of UI-spell

Second, define four mutually exclusive states, $s \in \{a, b, c, d\}$, for assignees in a given period t:

(a) Yet to be assigned:

$t < t_1$

(b) Yet to start training (or exit UI before training start):

$t_1 \leq t < t_2$

$t_e < t_2$ & $t \geq t_e$

(c) Undergoing training (or exit UI during training):

$t_2 \leq t \leq t_3$

$t_2 \leq t_e < t_3$ & $t \geq t_e$

(d) Done with training (or exit UI after/by end of training):

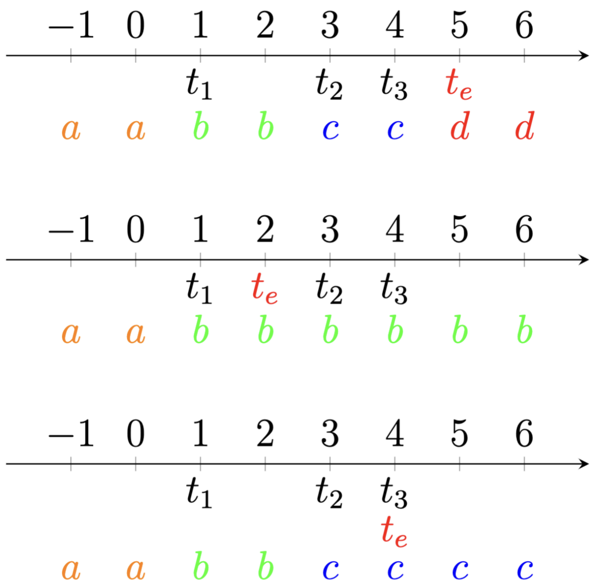
$t > t_3$

$t_e \geq t_3$ & $t > t_e$

All job seekers assigned to training at some point will belong to one of the states $s \in \{a, b, c, d\}$ in a given period t. If the job seeker exits unemployment after the end of the activity, she will transition through all four states. However, if she exits unemployment earlier, she will remain in her latest state in all later periods. Figure F.1 illustrates this point with three scenarios. In all three scenarios, a job seeker meets with her caseworker in month 1 relative to job loss (t_1), is assigned to training that starts in month 3 (t_2), and ends in month 4 (t_3). What differs across scenarios is when the job seeker exits unemployment, and hence how many states she progresses through.

In the first scenario, the job seeker exits unemployment in month 5. Therefore, she progresses through all four states: She is (a) yet to be assigned in months -1 to 0, (b) yet to start training in months 1-2, (c) undergoing training in months 3-4, and (d) done with training from month 5 and onward. In the second scenario, the job seeker exits unemployment even before she has started training. Therefore, she only progresses to state (b) and remains there in all periods onward. In the third scenario, the job seeker exits unemployment while undergoing training. Therefore, she only progresses to state (c) and remains there in all periods onward.

Figure F.1: A Job Seeker's Transition through Training States



Notes: This figure illustrates how job seekers may transition through training states.

F.2 Decomposition with Heterogeneous Counterfactuals

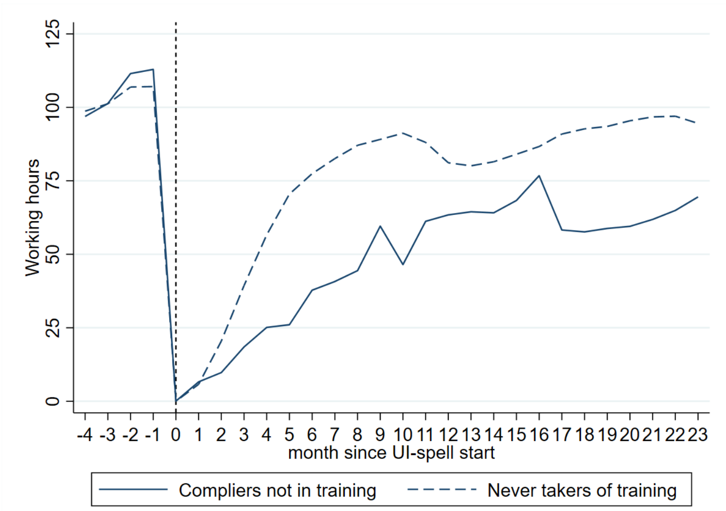
We relax the assumption of homogeneous counterfactuals by estimating state-specific counterfactuals for compliers assigned to training, β_{1t}^{0s}. To this end, we exploit that the average counterfactual for compliers can be written as a weighted average of the state-specific counterfactuals, using the state probabilities as weights:

$$\beta_{1t}^{0} = \sum_{s \in \{a,b,c,d\}} \gamma_{1t}^{s} \times \beta_{1t}^{0s} \tag{49}$$

We already estimated and plotted the *average* counterfactual outcome of compliers over time, β_{1t}^{0}, in figure A.4. Our challenge is that this average represents a mix of job seekers in the four training states, and splitting it into state-specific counterfactuals requires additional assumptions. Mathematically, Equation (49) constitutes T equations (one average outcome per period) in $4 \times T$ unknowns (four state-specific outcomes per period) and is not identified without restrictions on how outcomes vary across states or time. We will adopt three restrictions on how counterfactuals vary over time. We validate that the data support each restriction.

First, we allow for a general evolution in employment prospects since the time of job loss. For example, as time passes, more job seekers may find jobs for other reasons than changing training states. We measure this general evolution in job prospects based on the outcomes for never-takers, Y_t^{never}. These are job seekers that do not train even if assigned to the most training-inclined caseworker and hence are not shifted across training states by our caseworker instrument. We follow the approach of Bhuller et al. (2020) when identifying the outcome of never takers; see Online Appendix OA2.2 for further details. Figure F.2 shows the average employment rate for never-takers and compliers that do not train. Evidently, the two groups share a similar profile where job prospects improve rapidly in the first nine months following job loss and then flatten.

Figure F.2: Never-Takers and Compliers Not in Training



Notes: This figure shows the average monthly working hours in a given quarter relative to job loss, for never-takers and compliers not assigned to classroom training. All estimates rely on using percentiles 1 and 99 of the classroom-training instrument as proxies for training-restrained and training-inclined caseworkers. See further details in Online Appendix OA2.2. No indications of statistical significance in the figure.

Second, we allow the consequences of job loss to differ for compliers and never-takers. For example, if caseworkers target training to job seekers who struggle to find jobs otherwise, we would expect job loss to be more detrimental to the employment prospects of compliers. Indeed, Figure F.2 shows compliers have (increasingly) lower employment after job loss compared to never-takers. We model this gap by introducing a selection premium, δ_0, common to all compliers from the first month after job loss.[48]

[48]Mechanically, both never-takers and compliers have null (non-supported) working hours in the period of job

Third, we allow for state-specific selection premiums, δ_s. These premiums capture if job seekers switch into training based on their counterfactual in the period. To identify these state-specific premiums, we leverage that the training states kick in at different time horizons, as shown by the state probabilities in Figure 5. For example, if job seekers train when they face adverse job opportunities, we should observe that the mean counterfactual outcome of compliers is comparably low in months 2-6 when they train most intensively. Econometrically, we identify the state-specific selection premiums from periods with high state probabilities, γ_{it}^s.

Putting the pieces together, we model the state-specific counterfactual employment outcome of compliers as follows,

$$\beta_{1t}^{0s} = Y_t^{never} + \delta_0 \times \mathbf{1}(t \geq 1) + \delta_s \times \gamma_{1t}^s \tag{50}$$

The next section explains how we estimate the job loss penalty, δ_0, and state-specific selection premiums, δ_s. After that, we compute the state-specific counterfactuals and present a decomposition of the baseline IV estimate based on these counterfactuals.

Selection Premiums

In a first step towards estimation of the selection premiums, we insert (50) and estimated parameters in equation (49):

$$\hat{\beta}_{1t}^{0} = \sum_{s\in\{a,b,c,d\}} \hat{\gamma}_{1t}^s \times \underbrace{\left(\hat{Y}_t^{never} + \delta_0 \times \mathbf{1}(t \geq 1) + \delta_s \times \hat{\gamma}_{1t}^s\right)}_{\beta_{1t}^{0s}} \tag{51}$$

In this equation, blue font indicates parameters we already have estimated, and red font indicates unknown parameters. Hence, we have T equations in five unknowns. Note that Equation (51) point identifies the parameters if we have five data periods t with varying assignment rates. However, we have more than five periods and will estimate the unknown parameters by minimizing the sum of squared deviations as specified below.[49]

$$\min_{\delta_0,\delta_s} \sum_{t\in\mathcal{T}} \left[\hat{\beta}_{1t}^{0} - \sum_{s\in\{a,b,c,d\}} \hat{\gamma}_{1t}^s \times \left(\hat{Y}_t^{never} + \delta_0 \times \mathbf{1}(t \geq 1) + \delta_s \times \hat{\gamma}_{1t}^s\right)\right]$$

$$\text{s.t.} \quad \hat{Y}_t^{never} + \delta_0 \times \mathbf{1}(t \geq 1) + \delta_s \times \hat{\gamma}_{1t}^s \geq 0 \quad \forall (t, s)$$

loss (month 0). For this reason, we identify the job loss penalty for compliers from month 1 onward.

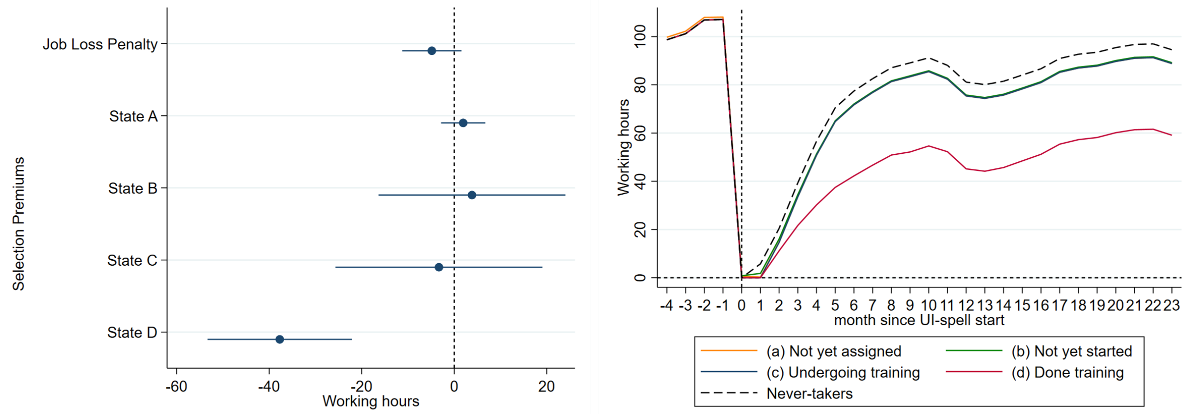
[49]This is inspired by Mogstad et al. (2018) who formulate the conversion of LATE estimates to target parameters as a minimization problem.

We solve the minimization problem subject to the constraint that counterfactual employment should always be non-negative. Note that we use all event months (-4: +23), except the month of and prior to job loss (0 and -1), when solving the constrained minimization problem.[50]

Panel (a) of figure F.3 plots the estimated selection premiums. The fact that δ_a, δ_b, δ_c, and δ_0 are statistically insignificant suggests that i) never-takers and compliers have similar employment rates prior to job loss, and ii) after a job loss, compliers have similar employment outcomes as never-takers as long as they have not completed their training. Hence, we do not find evidence suggesting that job seekers are assigned to or participate in training when they have worse job prospects. The fact that δ_d is statistically significant negative suggests that counterfactual employment rates are lower when compliers finish their training. The estimate suggests compliers would have been employed 35 hours *less* than never-takers had they not finished their training. The substantial negative selection premium for job seekers in state (d) is consistent with trainees being structurally challenged in the labor market.

Panel (b) of figure F.3 plots the counterfactual employment rate in a given state, $\hat{\beta}_t^{0s}$, that results from plugging the estimated never-taker outcome and selection premiums into (50). The figure shows that, following job loss, compliers in state (d) have much worse job prospects than never-takers.

Figure F.3: Heterogenous counterfactuals



(a) Selection premiums

(b) Heterogeneous counterfactuals

Notes: This figure shows the counterfactual employment rates for compliers in a given training state. Panel (a) plots selection premiums into the training states and 95% confidence intervals. Standard errors in this panel are constructed based on 100 bootstrap repetitions (see description of the bootstrap procedure in appendix OA2.1). Panel (b) plots the resulting state-specific counterfactual employment rates. There is no indication of statistical significance in Panel (b).

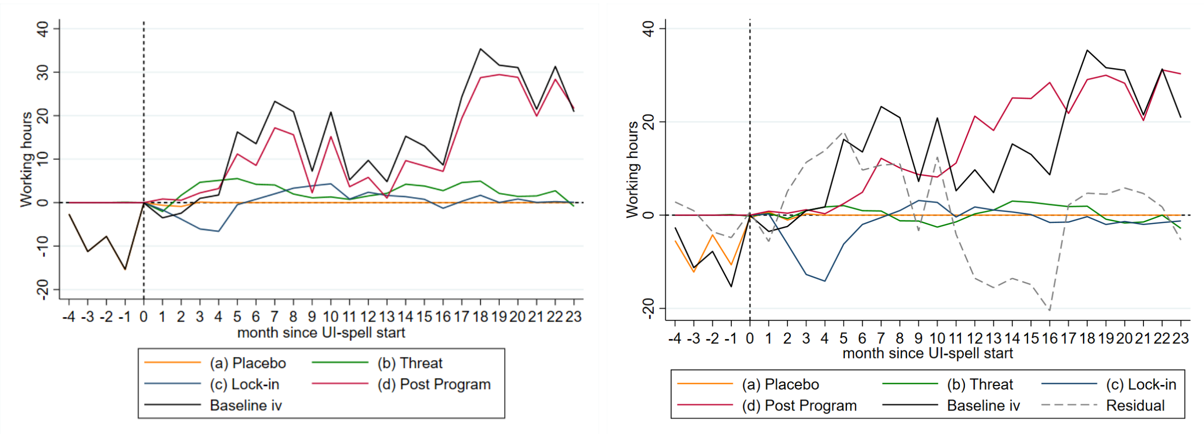
[50]We set the starting values to zero and test the identified minimum is invariant to this choice of starting values.

Decomposition with Heterogeneous Counterfactuals

We now have all inputs to do a decomposition of our IV estimate that allows for heterogeneous counterfactuals. Because our model of the counterfactuals has fewer parameters than we have data points, the decomposition will not match the data exactly. For this reason, we introduce a residual term to the decomposition.

Figure F.4 compares our decomposition of the IV estimate with homogeneous and heterogeneous counterfactual employment rates. This figure clearly shows our conclusions are robust to the assumption of homogeneous counterfactuals across states: by large, the threat, lock-in, and post-program effects do not depend on whether we assume homogeneous counterfactuals.

Figure F.4: Decomposition of Effect on Working Hours



(a) Homogeneous counterfactuals (b) Heterogeneous counterfactuals

Notes: This figure shows decompositions of the baseline IV estimate of the effect of assignment to classroom training on working hours in a given month relative to job loss. Panel (a) represents a decomposition with homogeneous counterfactuals, according to Equation (4), and (7). Panel (b) represents a decomposition with heterogeneous counterfactuals, according to Equation (4), and (50). No indications of statistical significance in this figure.

Online Appendix

What Works for the Unemployed? Evidence from Quasi-Random Caseworker Assignments

Anders Humlum, *U of Chicago*
Jakob R. Munch, *U of Copenhagen*
Mette Rasmussen, *U of Copenhagen*

OA1 Data

OA1.1 UI spells for 2012-2018

We identify all UI spells initiated in 2012-2018 based on the DREAM register. This register covers the entire Danish population and contains information about *weekly* receipts of any type of public transfers including UI benefits. We define a UI spell as the weeks in which the job seeker receives UI benefits or participates in activation programs (UI benefits hereafter), using the following DREAM codes: 111, 115, 121, 213, 214, 215, 216, 217, 218, 219, 231, 299. We identify the *start* of the UI spell as the first week with UI benefits, followed by at least eight consecutive weeks with no UI benefits. We define the *end* of the UI spell as the last week with UI benefits, followed by at least eight consecutive weeks with no UI benefits. Note this means that we allow for shorter breaks from receipt of UI benefits.

OA1.2 Linked Job-Seeker–Caseworker Data

A new data set from the Danish Agency for Labor Market and Recruitment (STAR) records the meetings between caseworkers and job seekers in Danish job centers from 2012-2019. This dataset contains information on the date, time, and type of all meetings. The data also record identifiers for the job seeker and caseworker attending each meeting, allowing us to link job seekers to their caseworkers. The job seeker identifiers can furthermore be linked to the registers in Statistics Denmark. The caseworker identifiers are administrative IDs and cannot be linked to the other registers.

This section describes our steps to clean and prepare the raw caseworker data. The goal is to end up with a data set that links unemployed job seekers (i) to the caseworker who participated in the job seeker's first face-to-face meeting, denoted her *realized caseworker*, and (ii) to her day-of-birth *predicted caseworker*.

Match UI Spells to Caseworker Meetings

First, we expand the sample of UI spells initiated from 2012-2018 with information about all caseworker meetings held during the UI spell.[51] Table OA1 shows we are able to match around 50% of all UI spells (75% of all job seekers) to at least one caseworker meeting. This reflects that many job seekers exit unemployment before having a caseworker meeting: the table shows the average UI spell length increases substantially, from 19 to 34 weeks, when we restrict to job seekers who had at least one caseworker meeting.

Timing, Type, and Contact of First Meeting

To ensure we do observe the *first* meeting for all job seekers, we impose three restrictions on the sample data. In particular, we only keep UI spells if the first registered meeting (i) takes place within the first 16 weeks of the UI spell (the 95th percentile), (ii) is a "regular meeting" or an "information meeting",[52], and if it (iii) takes place "in person" (i.e., not held via telephone or e-mail). Table OA1 shows these restrictions do not affect the average UI spell length.

Crosswalk of Caseworker Identifiers

To account for potential discontinuations of the caseworker identifiers, we also implement a crosswalk around months in which the vast majority (>95%) of caseworkers in a job center change identifiers. We provide more details below in Online Appendix OA1.3. Table OA1 shows the crosswalk reduces the number of unique caseworkers by 446 (3%).

Identifying the Realized Caseworker

Finally, we collapse the data to the first face-to-face meeting in the UI spell and denote the caseworker participating in this meeting as the *realized caseworker*. For the vast majority (99%) of job seekers, the first face-to-face meeting simply corresponds to the first meeting in the UI spell.

[51] Because UI spells are constructed based on weekly data, we do not know the exact start date of the UI spell. Therefore, we also include meetings held in the week prior to the start of the UI spell and meetings held in the week after the UI spell ends.

[52] To be precise, we require the meeting type to be coded as "Jobsamtale", "Jobsamtale med deltagelse af A-kassen", "Informationsmøde", "Informationsmøde uden mødepligt"

For a minority of the job seekers, the first face-to-face meeting corresponds to the second meeting in the UI spell. The reason is that some job centers first invite job seekers to an information meeting, in which one caseworker meets with a group of job seekers, and only thereafter invites the job seeker to a face-to-face meeting with the day-of-birth allocated caseworker. Hence, if the first meeting is an information meeting, we use the second meeting in the UI spell (provided it is a face-to-face meeting). If neither the first nor the second meeting was a face-to-face meeting, we drop the UI spell from the sample. Table OA1 shows when we collapse the data to the first face-to-face meeting, we keep almost all job seekers (99%) but lose about 45% of the caseworkers. This finding may reflect that some job centers have caseworkers who only meet job seekers if they become long-term unemployed.

Identifying the Day-Of-Birth-Predicted Caseworker

We now have a dataset that links job seekers to their *realized* caseworker. We link this dataset to the birth records of all job seekers and document the use of day-of-birth allocation rules in all job-center-unit-years. In other words, we now identify the *day-of-birth-predicted* caseworker for all job seekers. See details below in Online Appendix OA1.4. Table OA1 shows about 25% of the realized caseworkers end up as predicted caseworkers, and across all job-center-unit and years, the compliance rate between predicted and actual caseworkers is 42%.

Table OA1: Linked Job-seeker–Caseworker Data Restrictions

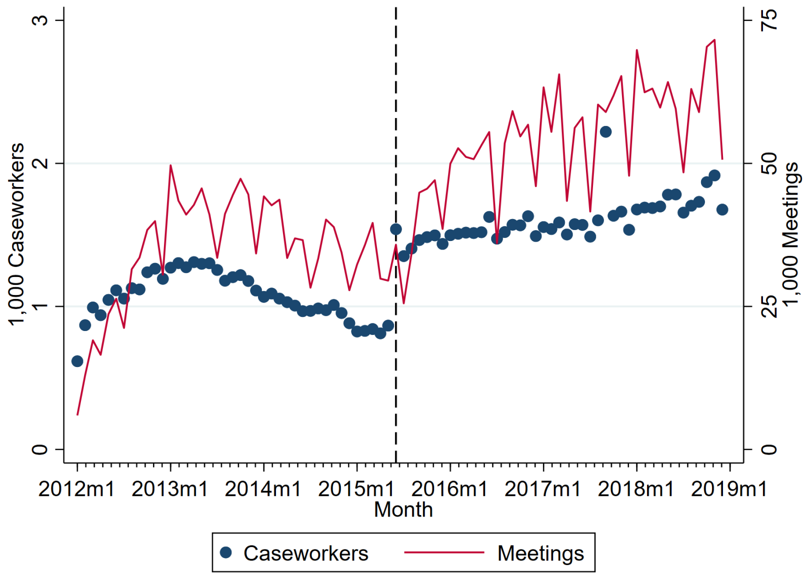
				Caseworkers		Average	
	UI-spells	Jobseekers	Jobcenters	Realized	Predicted	Weeks of UI	Compliance
UI-spells from 2012-2018	**2,200,785**	**940,971**	**94**	.	.	**19**	.
- who had at least one caseworker meeting	1,109,370	700,058	94	18,681	.	34	.
- timing, type and contact	944,462	633,445	94	14,372	.	33	.
- crosswalk caseworker id's	944,462	633,445	94	13,926	.	33	.
- identify realized caseworker	942,141	632,176	94	7,910	.	33	.
- identify day-of-birth protocol	934,922	628,352	94	7,910	1,949	33	0.42
Linked jobseeker-caseworker data	**934,922**	**628,352**	**94**	**7,910**	**1,949**	**33**	**0.42**

Notes: This table shows the number of units (UI-spells, job seekers, job centers, realized and predicted caseworkers) retained after each data restriction along with descriptives (average UI-spell length and the compliance between realized and predicted caseworker). Row (1) reports the statistics for all UI spells initiated from 2012-2018 in Denmark. Row (2) restricts the data to UI spells that had at least one caseworker meeting. Row (3) restricts the data to UI spells for whom the first meeting (i) took place within the first 16 weeks of job loss, (ii) was labeled as a "regular meeting" or an "information meeting", and (iii) was held "in person". Row (4) implements a crosswalk of caseworker identifiers. Row (5) restricts the data to UI spells for whom we can identify their realized caseworker. Row (6) restricts the data to UI spells for whom we can identify their day-of-birth predicted caseworker. Row (7) is identical to row (6) and summarizes the final linked job-seeker–caseworker data.

OA1.3 Crosswalk of Caseworker IDs

Figure OA1 shows the number of registered meetings (red line) as well as the number of active caseworker IDs (blue dots) in a given month. The black dotted line indicates the implementation of the 2015-employment reform in Denmark, which increased the frequency of caseworker meetings for all job seekers. Besides a sharp increase in the number of meetings around the reform, the figure shows the number of caseworkers generally follows the trend in meetings (unemployment). However, the figure also shows two significant jumps in the number of active caseworker IDs in June 2015 and September 2017. The sudden jumps in the number of active caseworker IDs (relative to the number of meetings) suggest a data break that likely could be caused by the adoption of new IT systems. Namely, the caseworker IDs are generated in the local job centers and may be discontinued if the job center adopts a new IT system.

Figure OA1: Caseworker IDs and meetings



Notes: This figure shows the registered number of caseworker IDs and meetings held in the Danish job centers from 2012 to 2019. The blue dots represent the number of unique caseworker IDs that were active in a given month (i.e., registered at least one meeting). The red line represents the total number of meetings held in the job centers in a given month. The black dotted line indicates a reform that increased the number of meetings during the first six months of unemployment.

We now define a job-center-specific churning rate that informs about the turnover of caseworker IDs in a given month. Let $caseworkers_{qt}$ represent all caseworker IDs that were active in job center q in month t. Further, let $exits_{qt}$ represent all caseworker IDs that were active in month t but not in month t+1, and $entries_{qt}$ represent all caseworker IDs that were active in month t but not in t-1. The rate of caseworker churning in a given job center q and month t can then be measured as

$$churning_{qt} = \frac{exits_{qt} + entries_{qt}}{caseworkers_{qt}} \tag{52}$$

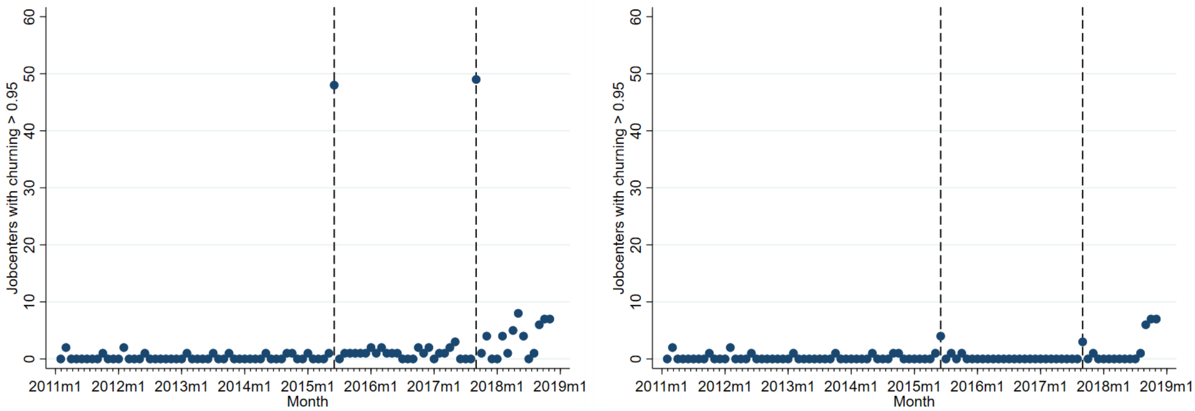
We bound this churning rate to be between 0 and 1.[53] When calculating the churning rate for a given job center q in a given month t, we will require

ii. the job center had at least two active caseworkers in period $t-1$, t and $t+1$

iii. the job center did not have churning $= 1$ in the previous period.[54]

Figure OA2a uses the raw data to plot the number of job centers in a given month with a churning rate ≥ 0.95 (denoted "high churning" hereafter). Note that close to half of all job centers experience a churning rate ≥ 0.95 in June 2015 and September 2017. This could reflect that many job centers changed IT system in these two months, such that caseworkers are assigned new IDs. To account for these breaks, we develop the crosswalk algorithm described below. Figure OA2b plots the number of job centers with a high churning rate after this crosswalk has been implemented. Clearly, the crosswalk reduces the number of job centers with a high churning rate significantly.

Figure OA2: job centers with churning ≥ 0.95



(a) Raw data

(b) Crosswalk

Notes: This figure is based on job centers with at least 4 active caseworker IDs in a given month. Each dot represents the number of these job centers that had a churning rate of at least 0.95. Panel (a) is based on the raw data; Panel (b) is based on the crosswalked data.

[53] In principle, the churning rate could be > 1, if some caseworker IDs enter and exit in the same month.

[54] If churning = 1 in two consecutive months, the first churning event is kept.

Crosswalk algorithm for a given job center and break month t:

1. Restrict to *meetings* registered within +/- 6 months from break month t (drop t).

   i. Require that job center have registered meetings in all months within the window.

   ii. Keep job seekers with at least one meeting before *and* after break month t.

2. Set up a transition matrix showing how job seekers that met with a given caseworker *before* the data break are distributed across caseworkers *after* the data break.

   i. Rows represent caseworkers that were active in the month before the break, $t-1$. Columns represent caseworkers that were active in the month after the break, $t+1$.

   ii. Find all job seekers that met with caseworker j in the *6 months* before the data break.

   ii. Count the number of meetings these job seekers have with any of the caseworkers in the *6 months* after the data break.

3. Do mapping from caseworker *before* the break to caseworkers *after* the break:

   i. Find the row-max in the transition matrix. I.e., for a given caseworker before the data break, what caseworker did most of her job seekers meet with after the data break.

   ii. If multiple before-caseworkers map into the same after-caseworker, use the mapping based on the largest crosswalk.

4. Accept mapping if certain criteria are met:

   i. The mapping must be based on a crosswalk of at least x job seeker meetings. The threshold x is data-driven and job center-specific and obtained as follows. For a given job center: Do Step 1-2 for a *placebo* break month (we use July 2016 since no job centers had churning=1 in this month). The off-diagonal elements in the placebo transition matrix inform us about the magnitude of crosswalks in a month *without* a break. Use the 95^{th} percentile of these crosswalks as threshold x if at least 3 active caseworkers. If less than three active caseworkers, use the average threshold for other job centers.

ii. The mapping must satisfy within-period consistency: Around a given data break month, the before-caseworker in one row cannot equal the after-caseworker in a different row. If this is the case, we disregard both mappings.

iii Implement crosswalk in a given job center and month if the churning rate in the job center in that month is > 0.95. If the churning rate is > 0.95 in consecutive months, we only implement the first churning event.

OA1.4 Day-Of-Birth Allocation Rules

Based on a survey of job centers and interviews with caseworkers, we know that job centers in Denmark often allocate job seekers to caseworkers based on the job seekers' day of birth. We impute the day-of-birth allocation rules that have potentially been applied in all 94 job centers in Denmark over time. The imputation involves two steps. Hereafter, we visually inspect the imputed day-of-birth rules to see what job-center-unit-years likely used this type of allocation rule.

Step 1: Identify Job-center Units

For each caseworker in a given job-center-year, we calculate the share of job seekers

- above age 30
- with an academic education[55]

We then check the distribution of the caseworker means in a given job-center-year: a bimodal distribution indicates the job center organizes the caseworker in units. Hereafter, we can then assign job seekers to a unit based on their individual characteristics (age or education). We assign caseworkers to a unit based on the characteristic of the majority of their assigned job seekers.

Step 2: Day-Of-Birth Distributions

For job centers without units

i. For each caseworker in a given job-center-year, count the total number of assigned job seekers whose birthdays fall a given day of the month (1-31).

[55] We define academic as having (at least) a bachelor's degree. Using Danish classifications, we include "60 Bacheloruddannelser", "70 Lange videregående uddannelser" and "PhD og forskeruddannelser".

ii. Across caseworkers in a given job-center-year, the caseworker with the highest number of job seekers born on a specific day of the month (1-31) becomes the *predicted caseworker* for that day.

For job centers with units

i. For each caseworker in a given job-center-unit-year, count the number of assigned job seekers who were born on a specific day of the month and belong to the same unit as the caseworker.

ii. Across caseworkers in a given job-center-unit-year, the caseworker with most job seekers born on a given day of the month (1-31) becomes the predicted caseworker for that day; that is, we use the mode.

We (visually) inspect the day-of-birth protocols and correct for differences in caseload size across caseworkers. For example. some job centers clearly use a day-of-birth rule but have one (or a few) caseworkers who deviate from the rule. If this caseworker also has a very large caseload, she will become the predicted caseworker for all job seekers (every day-of-birth). In this case, we place the "problematic" caseworker in her own unit and recompute the day-of-birth protocol for the entire job-center-unit-year.

Step 3: Visual Inspection

We now have an imputed day-of-birth allocation rule for all job-center-unit-years. Hereafter, we do a visual inspection of the imputed day-of-birth allocation rules. In particular, we check whether the imputed day-of-birth protocol appears to follow some *system*, which is clearly and easily readable by the unaided eye. For example, can we see one caseworker is responsible for days 1-7, another caseworker is responsible for days 8-15, and so on.[56] We also consider whether the job-center-unit-year as a whole tend to *comply* with the imputed rule. We drop job-center-unit-years where there is no clear system in the imputed day-of-birth allocation rule exists or where the compliance with the imputed rule is very low.

[56]More sophisticated, yet still clear, systems could involve caseworkers dividing even and odd days between them. However, "block" systems are easier to detect.

OA1.5 Outcomes

We base employment outcomes on the BFL register. This register covers the entire Danish population for 2008-2021 and is based on third-party reports to the Danish Tax Authorities. We use the variables ajo_loentimer and ajo_smalt_loenbeloeb to measure *any* working hours and earnings in a given month, and we define extensive-margin employment as months with non-zero working hours.

We use the DREAM register to separate non-supported employment from supported employment. In particular, we define non-supported employment as employment with no contemporaneous receipt of UI benefits and participation in activation programs (DREAM code cannot read 111, 115, 121,213, 214, 215, 216, 217, 218, 219, 231, 299) .

OA1.6 Offshorability Index

We use the O*NET database (version 26.1, November 2021) to construct an occupation-based measure of exposure to offshorability.[57] The O*NET database contains a wide range of occupational descriptors, for example, "Task ID 4.A.3.a.3: Controlling machines and processes", which can be used to characterize an occupation. Each descriptor is associated with an importance and level scale. Following O*NET guidelines, we construct one measure for each descriptor that incorporates both the level and the importance scale.[58] Because the occupations in O*NET are measured at the detailed onet-soc2019 level, and occupations in the Danish registers are measured at the coarser isco08 level, we apply a crosswalk and collapse the occupational characteristics from O*NET to the isco08 level. Finally, we do a standardization such that the occupational descriptors have a mean 0 and a standard deviation 1. Hereafter, we use the occupational descriptors to construct an index for the offshorability of a given isco08 code. Here, we follow the definition of offshorability from Autor and Dorn (2013). In particular, we use the following task items:

- Interpersonal interactions
  - 4.C.1.a.2.1 Face-to-face discussions

[57] O*NET data are downloaded from `https://www.onetcenter.org/db_releases.html`. We use four O*NET datasets: Work activities, Work context, Abilities, and Skills.

[58] Because importance is measured on a scale from 1-5 and level is measured on a scale from 1-7, we first do a normalization and hereafter aggregate the two.

- 4.A.4.a.4 Establishing and maintaining interpersonal relationships
- 4.A.4.a.5 Assisting and caring for others
- 4.A.4.a.8 Performing for or working directly with the public
- 4.A.4.b.5 Coaching and developing others

* Proximity to specific work place location
  - 4.A.1.b.2 Inspecting equipment, structures, or material
  - 4.A.3.a.2 Handling moving objects
  - 4.A.3.a.3 Controlling machines and processes
  - 4.A.3.a.4 Operating vehicles, mechanized devices, or equipment
  - 4.A.3.b.4 Repairing and maintaining mechanical equipment (*0.5)
  - 4.A.3.b.5 Repairing and maintaining electronic equipment (*0.5)

Table OA2: Five Largest Low- and High-Risk Occupations

Low-risk occupations	High-risk occupations
1. Shop sales assistant	1. Cleaners and helpers in offices, hotels and other establishments
2. Child-care workers	2. General office clerks
3. Home-based personal care workers	3. Clerical and customer support workers
4. Pedagogical work	4. Shelf fillers
5. Carpenters and joiners	5. Commercial sales representatives

Notes: This table shows the five largest occupations among job seekers at low and high risk of offshoring (i.e. the occupations most frequently held by job seekers in each group). Job seekers in the analysis sample are divided into low-risk and high-risk according to the offshorability of their previous occupation (using the 75th percentile in the job seeker distribution).

OA1.7 Imputation of Occupations

We use the AKM data to identify the previous occupation of all job seekers in our sample. AKM data allow us to see the yearly occupation of all employed individuals in Denmark from 1995-2017.[59] We use 4-digit isco codes to measure occupations. To account for the revision of isco codes in 2010, we apply an m:1 conversion key from isco88 to isco08 codes, such that we work with isco08 codes in all years. Following Hummels et al. (2014), we impute missing isco08-codes whenever possible. In particular, we merge workplace identifiers[60] onto the sample and

[59]To ensure we focus on employed individuals, we drop individuals in years when they have a missing or unknown industry code

[60]We use ARBNR ("arbejdsstednummer") from RAS, which is available from 1995 onwards.

use them to construct employment spells for the individual. If the individual has a missing isco08 code within an employment spell, we extrapolate from years in which the isco08 was not missing.[61] Finally, we identify the previous occupation of the job seekers in our sample. Here, we use the *typical* occupation over the job seeker's career, that is, the occupation in which she had most experience over the career (prior to the start of the UI spell).[62]

OA1.8 Validation of Job Plans

For information on training-program *assignments*, we rely on the individual job plans that caseworkers prepare for job seekers. The job-plan registrations are collected by the Danish Agency for Labor Market and Recruitment (STAR) and gathered in the so-called *Plan* data register (PLAN). Since 2015, this register has been the source of two well-known sources of information about ALMPs in the Danish labor market:

- Macro-level data: *jobindsats.dk*
- Micro-level data: *Danish Register for Evaluation of Marginalization (DREAM)*.

Importantly, whereas PLAN shows ALMP *assignments*, jobindsats.dk and DREAM show *participation* in ALMPs. Hence, PLAN data also include all the programs that were assigned but that the job seeker never participated in.

To the best of our knowledge, we are the first to use PLAN data for research on ALMPs, and this section serves as a validation of the dataset. First, we introduce the different types of ALMP assignments registered in PLAN data. Second, we benchmark PLAN data up against ALMP participation in the DREAM register. Third, to understand the content of a classroom training assignment in PLAN data, we benchmark PLAN data against two Danish registers, UDDA and VEUV, which contain information about enrollment in ordinary education (degree courses) and adult and continuing training (non-degree courses). The UDDA and VEUV registers mainly cover publicly provided education and thus provide a lower bound on the actual educational activity of job seekers.

[61] We require non-missing isco08-codes to be similar within an employer.
[62] In case of a tie, we use the lowest isco08 code.

OA1.8.1 PLAN Data

PLAN data contain information about all caseworker-*assigned* training programs for unemployed job seekers, including the start and end date of the programs and the status of the programs (e.g., canceled or completed).

We use PLAN data to distinguish between two broad classes of ALMPs: classroom training and on-the-job training.[63] Each class represents a composite of the activation categories used in Denmark, e.g., used on *jobindsats.dk* and in *DREAM*.[64] Table OA3 shows classroom training primarily includes "Ordinær uddannelse" (32%) and "Opkvalificering og vejledning" (66%), whereas on-the-job training primarily covers "Virksomhedspraktik" (72%) and "Løntilskud" (22%). In section OA1.8.2, we show how the activation categories in PLAN data match the activation categories in DREAM.

Table OA3: Training Classes and Activation Categories in PLAN Data

	(1) Total assignments	(2) Share (pct)
Classroom Training	**99,118**	**100**
- Opkvalificering og vejledning	65,410	66
- Ordinær uddannelse	31,357	32
- Andet	2,351	2
On-The-Job Training	**58,361**	**100**
- Virksomhedspraktik	41,903	72
- Løntilskud	12,683	22
- Andet	3,775	6

Notes: This table shows all assignments to classroom and on-the-job training, as well as the underlying activation categories, during the first 12 months of unemployment. Because a job seeker may be assigned to multiple training programs, the same job seeker can appear in multiple rows. Column (1) reports total assignments to a given activation category. Column (2) reports the share of assignments to a given activation category in percent of the total classroom (on-the-job) training assignments.

Table OA4 and OA5 report the different types of classroom training and on-the-job training programs registered in PLAN data for job seekers in our sample.[65] The table delivers two takeaways. First, on-the-job training programs include both public and private programs. For

[63]This distinction relies on two variables in PLAN data, *course_id* and *job_id*, each of which indicates the specific types of classroom and on-the-job training programs (see Tables OA4 and OA5).

[64]We identify the activation categories based the variables *aktivtyp*, *course_id*, and *job_id*. In particular, we ensure that all assignments to a given program type (within *course_id*, and *job_id*) belong to the same activation category (in *aktivtyp*).

[65]The two variables *course_id* and *job_id* show different types of classroom and on-the-job training programs, respectively. We enrich these variables with formats available at STARWIKI. *job_id*: https://starwiki.atlassian.net/wiki/spaces/FYS/pages/48890079/JobOrderTypeIdentifier *course_id*: https://starwiki.atlassian.net/wiki/spaces/FYS/pages/48890020/CourseTypeIdentifier

example, about one-third of all internships ("Virksomhedspraktik") is public ("offentlig"). This observation is important for the interpretation of our results because the existing literature has found different employment effects of public and private programs.

Second, the vast majority of classroom training programs are registered as "Øvrige forløb" (65%) or "AMU" (25%). The latter type of programs, "AMU," represent vocational training courses, which typically take place at dedicated facilities and target specific skills. The former type, "Øvrige forløb," is a rather generic category; the extent to which it represents actual education or job-search-assistance (JSA) courses is not clear from the title alone. In sections OA1.8.3 and OA1.8.4, we shed more light on the actual education content of classroom training in general and "Øvrige forløb" in particular.

Table OA4: On-The-Job Training Program Types

	(1) Total Assignments	(2) Share (pct)	(3) Duration (days)	(4) Timing (weeks)
Virksomhedspraktik	**41,903**	**72**	**31**	**23**
- Privat	27,407	47	32	23
- Offentlig	14,496	25	29	23
Løntilskud	**12,683**	**22**	**99**	**25**
- Offentlig	7,337	13	103	24
- Privat	5,346	9	94	27
Andet	**3,775**	**6**	**276**	**28**
- Rotationsvikar	2,595	4	155	28
- Voksenlærlig og elev	603	1	684	29
- Nytteindsats	511	<1	60	28
- Fleksjob, førtidspensionist etc.	66	<1	2,943	27

Notes: This table is based on all job seekers in the analysis sample who were assigned to on-the-job training within the first 12 months of unemployment. The table lists the specific types of on-the-job training programs that these job seekers are assigned to. Note that a job seeker may be assigned to multiple program types (i.e., she may appear in multiple rows) and even to multiple activities of the same type (i.e., she may appear multiple times in the same row). Column (1) reports the number of assignments to a given program type. Column (2) relates the number of program-specific assignments to total on-the-job training assignments. Column (3) reports the average duration of the program, measured in days. Column (4) reports the average timing of the program, measured in weeks relative to the UI spell start.

Table OA5: Classroom Training Program Types

	(1) Total Assignments	(2) Share (pct)	(3) Duration (days)	(4) Timing (weeks)
Opkvalificering og vejledning	**65,410**	**66**	**33**	**21**
- Øvrige forløb	63,975	65	34	21
- Privat udbudte kurser	727	<1	33	23
- Realkompetencevurdering	708	<1	6	20
Ordinær uddannelse	**31,357**	**32**	**34**	**22**
- AMU	24,895	25	20	21
- Erhvervsuddannelser	2,379	2	99	24
- AVU	1,802	2	66	22
- KVU, MVU, LVU	1,155	1	84	22
- Danskundervisning	555	<1	148	20
- Læse-, skrive- og regnekursus	329	<1	58	19
- SOSU uddannelser	242	<1	136	23
Andet	**2,351**	**2**	**81**	**23**
- Andet tilbud[a]	1,817	2	81	21
- Mentor	406	<1	93	29
- Mestring af sygdom, etc.	128	<1	45	31

Notes: This table is based on all job seekers in the analysis sample who were assigned to classroom training within the first 12 months of unemployment. The table lists the specific types of classroom training programs that these job seekers are assigned to. Note that a job seeker may be assigned to multiple program types (i.e., she may appear in multiple rows) and even to multiple activities of the same type (i.e., she may appear multiple times in the same row). Column (1) reports the number of assignments to a given program type. Column (2) relates the number of program-specific assignments to total classroom training assignments. Column (3) reports the average duration of the program, measured in days. Column (4) reports the average timing of the program, measured in weeks relative to the UI spell start.

OA1.8.2 PLAN vs. DREAM

Whereas PLAN data contain information on *assignments* to training programs, DREAM contains information about *participation* in training programs during unemployment. With one exception, the activation measures in DREAM have been based on PLAN data since 2015.[66] For this reason, we would expect a high correspondence between activation assignments in PLAN data and activation participation in DREAM.

Table OA6 shows the correspondence between training assignments in PLAN data and participation in DREAM. A given cell in this table reports the share of job seekers, who were assigned to some activation category in PLAN data (row), and end up participating in some activation category in DREAM (column). Both assignment and participation must be within the first 12 months after job loss. Because job seekers may be assigned to multiple types of

[66]Importantly, "6-ugers selvvalgt" in DREAM is based on a different data source (TASS), and when determining what an individual did in a given week, "6-ugers selvvalgt" is prioritized above other activation measures. Hence, the individual can be in "ordinær uddannelse" according to PLAN data but in "6-ugers selvvalgt" according to DREAM.

activation, the columns in a given row do not sum to 100.

The table shows a clear mass concentration in the diagonals of the first four columns. For the vast majority of job seekers, assignment to an activation category in PLAN data is accompanied by participation in a similar activation category in DREAM. For example, 87% of the job seekers, who were assigned to “(1) Opkval. og vejl.”, end up participating in the same category in DREAM. Importantly, the fact that the diagonals do not sum to 100% does not *only* reflect that some assignments never lead to participation. It also reflects that DREAM operates with the additional activation category “(5) 6-ugers selvvalgt,” which dominates the other categories. That is, some of the job seekers who are assigned to “(1) Opkval. og vejl.” may be classified in DREAM as participating in “(5) 6-ugers selvvalgt” (see footnote 66).

Table OA6: Activation in PLAN and DREAM

	Share participating in DREAM (pct)				
Assignment in Plan	(1) Opkval. og vejl.	(2) Ordinær udd.	(3) Virk. praktik	(4) Løntilskud	(5) 6-ugers selvvalgt
(1) Opkval. og vejl.	**87**	10	25	7	15
(2) Ordinær uddannelse	31	**75**	24	7	**30**
(3) Virk. praktik	35	10	**93**	18	14
(4) Løntilskud	34	9	60	**93**	13
(5) Andet	30	12	40	9	10

Notes: This table shows how assignments to training in PLAN data correspond with participation in training in the DREAM register. Each row represents all job seekers in the analysis sample who are assigned to a given activation type within the first 12 months after job loss according to PLAN. The columns report the share of these job seekers who end up participating in a given activation category in DREAM (participation within the first 12 months after job loss). Because job seekers may be assigned to multiple programs, the same job seeker can appear in multiple columns. That is, the columns in a row do not sum to 100. The activation categories in the columns are based on the following DREAM codes: “(1) Vejledning og opkvalificering” (213), “(2) Ordinær uddannelse” (214), “(3) Virksomhedspraktik” (217, 218), “(4) Løntilskud” (215, 216), “(5) 6-ugers selvvalgt uddannelse” (231).

OA1.8.3 PLAN vs. UDDA

The *Education register* (UDDA) contains information about enrollment in ordinary education, which we denote *Degree Courses*. Every year on September 30, Statistics Denmark takes *stock*: they register if an individual currently is *enrolled* in education and the day the individual enrolled in that education.[67]

To get an indication of the share of degree courses in classroom training, we compare classroom training assignees in PLAN data with *enrollments* in UDDA. However, note this comparison likely represents a *lower bound* of the share of degree courses in PLAN data: First of all, because some job seekers will be assigned to but never enroll in a degree course. For example,

[67] Statistics Denmark also takes stock of whether the individual has *completed* an education and register the day the individual obtained her degree. That is, they do not record the day the individual enrolled, only the completion date.

they may find a job upon assignment and, therefore, never enroll in education. Second of all, due to the nature of the UDDA register (the stock as of September 30), we cannot be sure to see all enrollments in degree courses. In particular, we lack information about some enrollments in courses with a short duration. For example, if the individual starts *after* September 30 of year t and the course ends *before* September 30 in the following year $t+1$, she will never appear as "enrolled" in the UDDA register (i.e., the enrollment date is unobservable to us).

Table OA7 reports the share of assignees to a specific classroom training program who also end up enrolling in an ordinary education in UDDA. Column (4) shows 4% of job seekers assigned to any classroom training program enroll in a degree course within the first 12 months from the UI spell start. Column (5) shows this share increases to 5% if we consider enrollments within the first 15 months from the UI spell start. In Column (6)-(7), the enrollment rates are weighted by the length of the assigned classroom training programs. Column (6) suggests degree courses can account for 9% of the days spent in classroom training.

Table OA7: Classroom Training Assignees Enrolled in UDDA

	(1) Total	(2) Share	(3) >1 program	(4) Enrollment (pct)	(5)	(6) Length-weighted enrollment (pct)	(7)
	assignees	(pct)	(pct)	month 0-11	month 0-14	month 0-11	month 0-14
Classroom training	**56,834**	**100**	**13**	**4**	**5**	**9**	**10**
- Øvrige forløb	44,381	78	15	3	5	4	5
- AMU	13,019	23	38	3	4	4	5
- Erhvervsuddannelser	1,652	3	48	29	32	52	56
- Andet tilbud	1,434	3	56	3	4	3	4
- AVU	1,391	2	51	6	8	11	13
- KVU, MVU, LVU	857	2	44	11	12	36	37
- Realkompetencevurdering	461	1	62	16	19	16	20
- Privat udbudte kurser	390	1	51	1	1	1	1
- Danskundervisning	370	1	46	2	3	2	3
- Mentor	320	1	65	2	2	1	1
- Læse-, skrive- og regnekursus	240	<1	60	8	10	9	12
- SOSU uddannelser	200	<1	42	41	43	71	72
- "Mestring af sygdom", etc.	91	<1	59	1	2	2	3

Notes: The table reports the share of assignees to a specific classroom training program who enroll in an ordinary education in the UDDA register. The table is based on all job seekers in the analysis sample (restricted to UI spells initiated from 2012-2016), who were assigned to classroom training within the first 12 months of unemployment. The table lists the specific *types* of classroom training programs these job seekers are assigned to. Because a job seeker may be assigned to multiple types of classroom training programs, the same job seeker may appear in multiple rows. Column (1) reports the total job seekers assigned to a given type of classroom training program during her UI spell. Column (2) reports the share of job seekers who were assigned to multiple program types, for example, both assigned to "Øvrige forløb" and "AMU". Columns (3)-(4) report the share of all classroom assignees who were enrolled in UDDA. Columns (5)-(6) report the share of all classroom assignees who were enrolled in UDDA, weighted by the total length of the assigned training programs. For Columns (3) and (5), the UDDA course must start within month 12, whereas for column (4) and (6) the UDDA course must start within month 15 of the UI spell start.

OA1.8.4 PLAN vs. VEUV

The *Course participant register* (VEUV) contains information about enrollments in publicly co-funded adult and continuing training courses that lead to formal qualifications. We denote these courses *Non-degree courses*. We use this register to check whether classroom training assignees in PLAN data (job seekers assigned to classroom training within the first 12 months of unemployment) become enrolled in a non-degree course within the first 12 (15) months from the UI spell start.

Because we only have VEUV data until September 2018, we can only observe course enrollments within all 12 (15) months after job loss for job seekers who initiated their UI spell no later than September (April) 2017. For simplicity, we restrict our sample to UI spells initiated from 2012-2016.

Table OA8 summarizes the correspondence between classroom training assignments in PLAN data and non-degree course enrollments in VEUV. Column (4) shows 41% of all classroom training assignees enroll in a non-degree course during the first 12 months after the UI spell start. Column (5) shows this share increases to 43% if we consider enrollment in courses that start within the first 15 months after the UI spell start. This finding suggests that non-degree courses constitute a relatively large share of classroom training. In Column (6)-(7), the enrollment rates are weighted by the length of the assigned training programs. Column (6) suggests non-degree courses can account for 44% of the total days spent in classroom training.

Table OA8 also reports the course enrollment rate for assignees to specific classroom training programs. We highlight two insights. First, the course enrollment rate is as high as 86% for assignees to vocational training courses ("AMU"; row 3, column 4). This finding is interesting because any enrollment in vocational training courses is registered in VEUV. Hence, when (100-86=) 14% of assignees to vocational training courses are *not* registered in VEUV, we have an indication of the extent to which enrollments follow assignments. In other words, we may take 14% as a proxy for the share of classroom assignees who never enroll in a program.

Second, about one-third of assignees to the generic course category "Øvrige forløb" enroll in a VEUV course within the first 12 months of unemployment (row 2, column 4). This share can*not* (only) be explained by job seekers being assigned to multiple types of programs.[68]

[68] A concern would be that assignment to "Øvrige forløb" is always accompanied by assignment to "AMU," which has a high course-enrollment rate. However, the majority (85%) of assignees to "Øvrige forløb" are *not* assigned to other types of programs (row 2, column 3). Hence, the high course-enrollment rate for assignees to

Rather, this finding suggests a significant share of “Øvrige forløb” is non-degree courses as registered in VEUV.

Table OA8: Classroom Training Assignees Enrolled in VEUV

	(1) Total assignees	(2) Share (pct)	(3) >1 program (pct)	(4) Enrollment (pct)	(5)	(6) Length-weighted enrollment (pct)	(7)
				month 0-11	month 0-14	month 0-11	month 0-14
Classroom training	**56,834**	**100**	**13**	**41**	**43**	**44**	**46**
- Øvrige forløb	44,381	78	15	31	34	31	34
- AMU	13,019	23	38	86	87	88	89
- Erhvervsuddannelser	1,652	3	48	54	56	42	44
- Andet tilbud	1,434	3	56	28	31	29	31
- AVU	1,391	2	51	76	78	80	82
- KVU, MVU, LVU	857	2	44	65	66	59	59
- Realkompetencevurdering	461	1	62	60	62	78	79
- Privat udbudte kurser	390	1	51	33	35	36	38
- Danskundervisning	370	1	46	81	82	86	87
- Mentor	320	1	65	23	27	19	22
- Læse-, skrive- og regnekursus	240	<1	60	80	83	93	93
- SOSU uddannelser	200	<1	42	56	57	35	36
- “Mestring af sygdom”, etc.	91	<1	59	19	21	19	22

Notes: The table reports the share of assignees to a specific classroom training program who enroll in a training course in the VEUV register. The table is based on all job seekers in the analysis sample (restricted to UI spells initiated from 2012-2016) who were assigned to classroom training within the first 12 months of unemployment. The table lists the specific *types* of classroom training programs these job seekers are assigned to. Because a job seeker may be assigned to multiple types of classroom training programs, the same job seeker may appear in multiple rows. Column (1) reports total job seekers assigned to a given type of classroom training program during her UI spell. Column (2) reports the share of job seekers who were assigned to multiple program types, for example, both assigned to “Øvrige forløb” and “AMU”. Columns (3)-(4) report the share of all classroom assignees who were enrolled in VEUV. Columns (5)-(6) report the share of all classroom assignees who were enrolled in VEUV, weighted by the total length of the assigned training programs. For Columns (3) and (5), the VEUV course must start within month 12, whereas for Columns (4) and (6), the VEUV course must start within month 15 of the UI spell start.

Table OA9 lists the ten most popular adult vocational training courses that classroom training assignees enroll in. As is evident, these popular courses target specific skills, for example, accounting (6, 7, 10), operating a computer (1, 4, 5, 8, 9) or a machine (2). [69]

“Øvrige forløb” cannot (only) be driven by assignments to other programs.

[69] With one exception, the same courses are in the top 10 if we focus on assignees to “Øvrige forløb” or “AMU”.

Table OA9: Popular Courses

Titel	Enrollments
1. Anvendelse af regneark til enkle beregninger	2,141
2. Gaffeltruck certifikatkursus B, 7 dage	1,707
3. Vurdering af basale færdigheder	1,503
4. Anvendelse af præsentationsprogrammer	1,454
5. Indskrivning og formatering af mindre tekster	1,442
6. Placering af resultat og balancekonti	1,265
7. Daglig registrering i et økonomistyringsprogram	1,240
8. Brug af pc på arbejdspladsen	1,225
9. Informationssøgning på internettet til jobbrug	1,207
10. Bilagsbehandling med efterfølgende kasserapport	1,155

Notes: This table reports the 10 vocational training courses that classroom training assignees most frequently enroll in. Classroom training assignees are based on our analysis sample (restricted to UI spells initiated from 2012 to 2016); assignment to classroom training must be within the first 12 months of unemployment. Course enrollments are based on VEUV; the course must start within the first 12 months of unemployment.

OA1.8.5 Back-of-the-Envelope Calculations for the Content of Classroom Training

What share of classroom training is made up of job-search-assistance (JSA) courses? We now present some simple calculations that put an *upper bound* on this share.

The calculations build on the idea that (i) classroom training includes three types of courses (JSA courses, degree courses, and non-degree courses) and that (ii) some assignments never lead to enrollment. Building on the previous sections, we have crude proxies for the enrollment rate into degree courses (UDDA) and non-degree courses (VEUV). Table OA10 reports these numbers as well as the share of classroom assignees who enroll in minimum one of these two courses, denoted *Any Education Course*. We also have a crude proxy for the non-enrollment rate of classroom training assignments if we are willing to assume the non-enrollment rate for "AMU" can be extrapolated to all classroom training programs. This proxy is also reported in Table OA10.

Based on these proxies, we calculate the residual share of classroom training assignees: job seekers assigned to classroom training but who (i) do not have a registered education enrollment and (ii) may not have enrolled in any classroom training program. Because we likely do not capture all education enrollments, we may interpret the residual as an *upper bound* of the share of classroom assignees who enroll in JSA courses.[70] Table OA10 reports the proxies and the residual, which suggests *maximum* 43% of classroom assignees enroll in JSA courses. The table also suggests JSA courses account for *no more* than 38% of total classroom training days.

[70] For example, because UDDA reports the stock of individuals enrolled as of September 30 in a given year, we will not capture all enrollments in degree courses, for example, if they are initiated after September 30 of year t.

Table OA10: Content of Classroom Training

	(1)	(2)	(3)	(4)	(5)
		Registered enrolment			
	Non-enrollment	Degree Course	Non-degree Course	Any Education Course	Residual
Share of classroom assignees	14	4	41	44	43
Share of classroom training days	12	9	44	50	38

Notes: This table summarizes the course enrolment rates for classroom training assignees. Column (1) shows the non-enrollment rate: the share of "AMU" assignees who do not enroll in VEUV (Columns (3) and (5) of Table OA8). Column (2) shows the enrollment rate in "degree courses": the share of classroom assignees who enroll in UDDA (Columns (3) and (5) of Table OA7). Column (3) shows the enrollment rate in "non-degree courses': the share of classroom assignees who enroll in VEUV (Columns (3) and (5) of Table OA8). Column (4) shows the enrollment rate in "any registered education": the share of classroom assignees who enroll in at least one of the two categories in Columns (2) and (3). Column (5) shows the residual, which is calculated as 100 - (Column (1) + Column (4)).

OA2 Estimation

OA2.1 Bootstrapping standard errors

We bootstrap standard errors on a given statistic, b, using a "Bayesian Bootstrap procedure":

1. Assign each observation in the sample a random and strictly positive weight. Obtain these weights by (i) drawing numbers from a uniform distribution and (ii) normalizing by the sum, $w_i = \frac{u_i}{\sum_i u_i}$.

2. Recompute the statistic with these weights.

3. Do N iterations of 1-2, such that we obtain an entire distribution of the relevant statistic, b. Compute standard errors based on this distribution:

$$se(b) = \sqrt{\frac{1}{N} \times \sum_{n}^{N} (b_n - \bar{b})^2}. \tag{53}$$

OA2.2 Complier Calculus

Given monotonicity, the population of job seekers can be split into three types: compliers, always-, and never-takers. *Compliers* are job seekers who would have received a different classroom-training decision had they been assigned to a different caseworker, namely, job seekers who (i) start in training k if assigned to a k-inclined caseworker and (ii) do not start in training k if assigned to a k-restrained caseworker. *Never-takers* are job seekers who do not start in training k irrespective of the encouragement of the caseworker. That is, even if assigned to the most k-inclined caseworker, never-takers do not start in training k. *Always-takers* are job seekers who start in training k irrespective of the encouragement of their caseworker. That is, even if assigned to the most k-restrained caseworker, always-takers do start in training k.

Share of Compliers, Always-, and Never-Takers

In the judge-leniency literature, the share of compliers, always-, and never-takers are often calculated in a setting with one treatment and a continuous instrument (Bhuller et al., 2020; Dahl et al., 2014). The idea is to approximate a "strict" and a "lenient" judge with some percentile of the residualized instrument,[71] and then re-scale the first-stage coefficient by the proxies for lenient and strict caseworkers.

Given extended monotonicity - and non-correlation of the instruments conditional on job-center unit and year fixed effects[72] - the approach from the single-treatment setting can easily be extended to our *multiple-treatment* setting. In particular, we take the following approach to identify the share of compliers, always-, and never-takers for treatment k:

1. Residualize training intervention D_{ki} and the own-instrument Z_{ki} on job-center-unit-year fixed effects, $\delta_{q(i)}$, and the cross-instrument Z_{-ki}. Add the unconditional mean and denote the residualized intervention and own-instrument $D^r_{k,i}$ and $Z^r_{k,i}$, respectively.

2. Identify a k-restrained and k-inclined caseworker as the 1st and 99th percentile of $Z^r_{k,i}$. Denote these percentiles $Z^{r,min}_k$ and $Z^{r,max}_k$, respectively.

3. Estimate the first-stage regression

$$D^r_{k,i} = \pi_{10} + \pi_{11} Z^r_{k,i} + u_{k,i}. \tag{54}$$

4. Predict treatment k for job seekers assigned to a k-restrained and -inclined caseworker

$$\hat{D}^{r,min}_k = \hat{\pi}_{10} + \hat{\pi}_{11} Z^{r,min}_k \tag{55}$$

$$\hat{D}^{r,max}_k = \hat{\pi}_{10} + \hat{\pi}_{11} Z^{r,max}_k \tag{56}$$

[71] The instrument is residualized by fixed effects representing the level at which randomization occurs

[72] Because the two instruments are uncorrelated, the distribution of the own-instrument, Z_{ki}, conditional on job-center unit and year fixed effects, $\delta_{j(i)}$, is largely invariant to controlling for the cross-instrument, Z_{-ki}, that is, $Z_{ki}|\delta_{j(i)} \sim Z_{ki}|\delta_{j(i)}, Z_{-ki}$.

5. Calculate the share of compliers, always-, and never-takers as

$$P_c = \hat{D}_k^{r,max} - \hat{D}_k^{r,min} = \hat{\pi}_{11} \times (Z_k^{r,max} - Z_k^{r,min}) \tag{57}$$

$$P_a = \hat{D}_k^{r,min} = \hat{\pi}_{10} + \hat{\pi}_{11} Z_k^{r,min} \tag{58}$$

$$P_n = 1 - P_c - P_a. \tag{59}$$

Note the above method is equivalent to a method in which we only residualize based on job-center-unit-year fixed effects in step 1:

1. Residualize training intervention D_{ki} and the instruments, Z_{ki} and Z_{-ki}, on job-center-unit-year fixed effects, $\delta_{q(i)}$. Add the unconditional mean and denote the residualized variables $D^r_{k,i}$, $Z^r_{k,i}$, and $Z^r_{-k,i}$,

2. Identify a k-restrained and k-inclined caseworker as the 1st and 99th percentile of the own-instrument $Z^r_{k,i}$. Denote these $Z_k^{r,min}$ and $Z_k^{r,max}$, respectively. A restrained and inclined caseworker, $Z_k^{r,min}$ and $Z_k^{r,max}$, will have approximately the same value of the cross-instrument if the two instruments are uncorrelated. That is, $Z_{-k}^{r,max} - Z_{-k}^{r,max} \sim 0$.

3. Estimate the first-stage regression. Due to extended monotonicity, the coefficient on the cross-instrument, π_{12}, is close to zero:

$$D^r_{k,i} = \pi_{10} + \pi_{11} Z^r_{k,i} + \pi_{12} Z^r_{-k,i} + u_{k,i}. \tag{60}$$

4. Predict treatment k for job seekers assigned to a k-restrained and -inclined caseworker:

$$\hat{D}_k^{r,max} = \hat{\pi}_{10} + \hat{\pi}_{11} Z_k^{r,min} + \underbrace{\hat{\pi}_{12}}_{\sim 0} Z_{-k}^{r,min} \tag{61}$$

$$\hat{D}_k^{r,min} = \hat{\pi}_{10} + \hat{\pi}_{11} Z_k^{r,max} + \underbrace{\hat{\pi}_{12}}_{\sim 0} Z_{-k}^{r,max} \tag{62}$$

5. Calculate the share of compliers, always takers, and never takers

$$P_c = \hat{\pi}_{11} \times (Z_k^{r,max} - Z_k^{r,min}) + \underbrace{\hat{\pi}_{12}}_{\sim 0} \times \underbrace{(Z_{-k}^{r,max} - Z_{-k}^{r,min})}_{\sim 0} \tag{63}$$

$$P_a = \hat{\pi}_{10} + \hat{\pi}_{11} Z_k^{r,min} + \underbrace{\hat{\pi}_{12}}_{\sim 0} Z_{-k}^{r,min} \tag{64}$$

$$P_n = 1 - P_c - P_a. \tag{65}$$

Potential outcomes

Building on the population shares obtained above, we can estimate the potential outcomes for compliers, always takers, and never takers with respect to a specific treatment k:

1. Residualize the outcome, Y_{it}, and instruments, Z_{1i} and Z_{1i}, on job-center-unit-year fixed effects, $\delta_{q(i)}$. Add the unconditional mean and call the residualized variables, $Y^r_{i,t}$, $Z^r_{k,i}$, and $Z^r_{-k,i}$.

2. Identify a k-restrained and k-inclined caseworker as the 1st and 99th percentiles of the own-instrument $Z^r_{k,i}$. Denote these $Z^{r,min}_k$ and $Z^{r,max}_k$, respectively. A restrained and an inclined caseworker, $Z^{r,min}_k$ and $Z^{r,max}_k$, will have approximately the same value of the cross-instrument if the two instruments are uncorrelated. That is, $Z^{r,max}_{-k} - Z^{r,max}_{-k} \sim 0$

3. Estimate the reduced form *conditional on treatment* and predict the outcome for job seekers assigned to a k-restrained and k-inclined caseworker:

$$Y^r_{i,t} = \alpha_0 + \alpha_k Z^r_{k,i} + \alpha_{-k} Z^r_{-k,i} + u_i, \quad \text{if} \quad D_{k,i} = 1 \tag{66}$$

$$\hat{Y}_t(D_k = 1, Z^{r,min}_{k,i}, Z^{r,mean}_{-k,i}) = \hat{\alpha}_0 + \hat{\alpha}_k Z^{r,min}_{k,i} + \hat{\alpha}_{-k} Z^{r,min}_{-k,i} = \hat{Y}^{always}_t \tag{67}$$

$$\hat{Y}_t(D_k = 1, Z^{r,max}_{k,i}, Z^{r,max}_{-k,i}) = \hat{\alpha}_0 + \hat{\alpha}_k Z^{r,max}_{k,i} + \hat{\alpha}_{-k} Z^{r,max}_{-k,i}. \tag{68}$$

4. Estimate the reduced form *conditional on non-treatment* and predict the outcome for job seekers assigned to a k-restrained and k-inclined caseworker:

$$Y^r_{i,t} = \alpha_0 + \alpha_k Z^r_{k,i} + \alpha_{-k} Z^r_{-k,i} + \epsilon_i, \quad \text{if} \quad D_{k,i} = 0 \tag{69}$$

$$\hat{Y}_t(D_k = 0, Z^{r,min}_{k,i}, Z^{r,min}_{-k,i}) = \hat{\alpha}_0 + \hat{\alpha}_k Z^{r,min}_{k,i} + \hat{\alpha}_{-k} Z^{r,min}_{-k,i} \tag{70}$$

$$\hat{Y}_t(D_k = 0, Z^{r,max}_{k,i}, Z^{r,max}_{k,i}) = \hat{\alpha}_0 + \hat{\alpha}_k Z^{r,max}_{k,i} + \hat{\alpha}_{-k} Z^{r,max}_{-k,i} = \hat{Y}^{never}_t. \tag{71}$$

5. Because always-takers start in training even if they are assigned to the most k-restrained caseworker, (67) identifies the outcome for always-takers.

6. Because never-takers do *not* start in training even if they are assigned to the most k-inclined caseworker, (71) identifies the outcome for never-takers.

7. To get the outcome for compliers if they do not start in training, note that job seekers who

do not start in training are a mix of never-takers and non-encouraged compliers:

$$\hat{Y}_{0,t}^{com} = \hat{Y}_t(D_k = 0, Z_k^{max} > Z_k^{min}) \tag{72}$$
$$= \frac{P_c + P_n}{P_c} \times \hat{Y}_t(D_k = 0, Z_{k,i}^{r,min}, Z_{-k,i}^{r,min}) - \frac{P_n}{P_c} \times \hat{Y}_t(D = 0, Z_{k,i}^{r,max}, Z_{-k,i}^{r,max}). \tag{73}$$

8. To get the outcome for compliers if they do start in training, note that job seekers who start in training are a mix of always-takers and encouraged compliers:

$$\hat{Y}_{1,t}^{com} = \hat{Y}_t(D_k = 1, Z_k^{max} > Z_k^{min}) \tag{74}$$
$$= \frac{P_c + P_a}{P_c} \times \hat{Y}_t(D_k = 1, Z_{k,i}^{r,max}, Z_{k,i}^{r,max}) - \frac{P_a}{P_c} \times \hat{Y}_t(D_k = 1, Z_{k,i}^{r,min}, Z_{-k,i}^{r,min}).. \tag{75}$$