FACULTY OF SOCIAL SCIENCES
Department of Economics
University of Copenhagen

Empirical Essays on the Effectiveness of Active Labor Market Policies:

Caseworkers, Vocational Training & Statistical Profiling

PhD Dissertation

Mette Rasmussen

Supervisors: Alexander Sebald & Steffen Altmann

Submitted: March 1, 2021

PhD Dissertation

Empirical Essays on the Effectiveness of Active Labor Market Policies:

Caseworkers, Vocational Training & Statistical Profiling

Mette Rasmussen
Department of Economics,
University of Copenhagen

Submitted: March 1st 2021
Supervisor: Alexander Sebald & Steffen Altmann

Contents

Acknowledgements

Above all, I am grateful to the Danish Ministry of Employment and the Agency for Labor Market and Recruitment for financing four years of research on the effectiveness of active labor market policies. I especially appreciate the novel data that they have provided me with. Further, I thank the many caseworkers and jobcenter directors that kindly have provided me with invaluable knowledge about the assignment of caseworkers in their local jobcenter.

I am thankful for the guidance provided by my supervisors, Alexander Sebald and Steffen Altmann, throughout my PhD. Further, I am indebted to my co-authors, Nikolaj Harmon, Robert Mahlstedt and Anders Humlum. I have learned a lot from our projects and collaborations.

I also owe a large thanks to the 'San Diego team', Sally Sadoff, Julie Cullen and Gordon Dahl, who hosted me and discussed research ideas with me during my stay at UC San Diego.

A special thanks is owed to my fellow PhD students. The Thursday running club, Christmas lunches, coffee breaks and causal Friday beers surely made the PhD journey enjoyable.

Finally, thanks to Magnus for endless discussions of my chapters and invaluable support. I promise to participate in cooking and talk a little less about caseworkers in the future. Also thanks to my friends and family for listening to my frustrations, complaints and excitement throughout my PhD.

Mette Rasmussen

March 1st, 2021

Introduction

How do policy makers help jobseekers transition out of unemployment? As unemployment is costly for the individual and inefficient for society, this is a question of utmost importance. Today, many OECD countries rely on active labor market policies (ALMPs) in their efforts to easen jobseekers' transitions out of unemployment and improve their subsequent labor market performance. Public expenditures on these policies are substantial. In 2017, the average OECD country spent about 0.5% of GDP on ALMPs, and Denmark was in the top with expenditures amounting to about 2%. Policy makers have a wide range of active labor market policies at their disposal, e.g. job search assistance and training facilitated through regular meetings with a caseworker. In order to chose the policies optimally, it is important to know what policies work best and for whom.

This PhD dissertation consists of three self-contained chapters that all study how different types of active labor market policies can help unemployed jobseekers transition out of unemployment. All three chapters rely on novel data provided by the Danish Agency for Labor Market and Recruitment.

The first chapter investigates the importance of caseworker quality for jobseeker transitions out of unemployment. The project combines a novel data set on meetings between jobseekers and caseworkers with rich Danish administrative data. To identify the individual impacts of caseworkers, I exploit that many jobcenters assign caseworkers to jobseekers based on their birthday. I verify that this effectively corresponds to a quasi-random assignment mechanism. The chapter offers three sets of results. First, I find that variation in caseworker quality is substantial and can explain about 6% of the heterogeneity in unemployment spells within a jobcenter and year. Assignment to a caseworker, who is one standard deviation above the mean, reduces the unemployment spell by about one week. Second, I find no smoking gun suggesting that this is at the expense of subsequent labor market performance. Rather, the caseworkers that reduce unemployment spell lengths tend to place jobseekers in jobs that are of similar quality, in terms of wages, hours and stability. As a result, these jobseekers have on average accumulated an additional 7,500 DKK (1,200 USD) and 35 working hours after two years. Third, I show that the variation in caseworker quality not necessarily is driven by unobserved personality traits only. Namely, I find that the high quality caseworkers tend to be more 'pro-active': They meet earlier and more frequently with the jobseekers, assign them earlier to training and tend to increase the jobseekers' use of network and unsolicited job search. Overall, this paper suggests that it would be Pareto improving to teach all caseworkers these pro-active strategies.

The second chapter (co-authored with Anders Humlum) investigates whether vocational training of jobseekers can help them re-attach to the labor market. To investigate this, we combine the novel data on caseworkers from the first chapter with new data on assignments to training. Specifically, we estimate the effectiveness of vocational training using a caseworker leniency instrument. This instrument exploits that i) jobseekers are quasi-randomly assigned to caseworkers, and that ii) caseworkers differ in their propensities to assign jobseekers to vocational training. Using our caseworker leniency instrument, we cannot reject that training courses on average have zero impacts on labor market outcomes after one year. In contrast, OLS regressions show strong negative correlations between training and employment, indicating that it is jobseekers with adverse job prospects who select into training. To investigate whether vocational training is more beneficial for workers who are exposed to rapid structural change, we zoom in on jobseekers whose previous jobs were in manufacturing. Although the estimates become more noisy, we find economically significant long-run benefits to vocational training for former manufacturing workers. This suggests that there is large heterogeneity in the benefits of training, which potentially could be reaped by better targeting of courses to workers.

The third chapter (co-authored with Nikolaj Harmon and Robert Mahlstedt) evaluates the effect of using statistical profiling tools to inform jobseekers (and their caseworkers) about their individual risk of long-term unemployment. These are low cost interventions, yet they could potentially be effective in speeding up jobseeker transitions out of unemployment, e.g. by restoring biased beliefs or by helping caseworkers target resources towards jobseekers in high risk of long term unemployment. In Denmark, a Machine Learning algorithm informs both newly unemployed jobseekers and their caseworkers of whether they belong to a group with a high risk of remaining unemployed for more than six months. Leveraging age discontinuities in the algorithm, we use a regression discontinuity design (RDD) to estimate the effect of being informed to be in high risk. We estimate that jobseekers marginally flagged as high-risk are between 5-14% less likely to be unemployed after 6 months. After 12 months however this difference has disappeared. Unfortunately, standard validity checks suggest that the identifying assumptions of our RDD may not hold. While our results thus points to statistical prediction tools as a promising way to speed up unemployment exits, more evaluation is necessary to reach any firm conclusion.

Finally, it should be noted that the third chapter builds on and repeats text parts from my master's thesis, "Behavioral Responses to Information about Individual Employment Prospects: Evidence from Denmark" (University of Copenhagen, 2019). The master's thesis was mainly concerned with the reconstruction of the historical version of the profiling algorithm, and included only very introductory RDD estimations. The project has been greatly extended in multiple dimensions since I defended my master's thesis. For example, we have worked extensively on the actual RDD analysis. This includes refining the initial strategy and carefully constructing samples, which allow us to explore heterogeneity in treatment effects. Further, we have done a thorough examination of the identifying assumptions and investigated the dynamics of the treatment effects.

Introduktion

Hvordan hjælper man borgere ud af ledighed? Ledighed er både omkostningsfuldt for individet og inefficient for samfundet. I forsøget på at øge transitionen fra ledighed til beskæftigelse beror mange OECD lande på en aktiv arbejdsmarkedspolitik. Udgifterne til indsatser i den aktive arbejdsmarkedspolitik er substantielle. I 2017 have det gennemsnitlige OECD land udgifter på omkring 0.5% af BNP. Danmark ligger i toppen med udgifter på op imod 2% af BNP. Når beslutningstagere skal designe den aktive arbejdsmarkedspolitik, kan de vælge mellem mange typer af indsatser, fx opkvalificering eller jobsøgningsassistance faciliteret via jævnlige møder med en sagsbehandler. For at designe beskæftigelsesindsatsen optimalt er det derfor altafgørende at vide hvilke indsatser, der virker bedst og for hvilke borgere.

Denne PhD afhandling består af tre selvstændige kapitler, der alle undersøger hvordan forskellige typer af indsatser i den aktive arbejdsmarkedspolitik kan hjælpe ledige med at afgå fra ledighed. Alle tre kapitler bygger på data, som Styrelsen for Arbejdsmarked og Rekruttering har stillet til rådighed.

Det første kapitel undersøger, hvor meget sagsbehandler kvalitet betyder for borgerens afgang fra ledighed. Dette projekt kombinerer et nyt datasæt over møder afholdt mellem den ledige og sagsbehandleren over ledighedsforløbet med dansk registerdata. Til at identificere effekten af individuelle sagsbehandlere udnytter jeg, at mange jobcentre i Danmark tildeler ledige borgere en sagsbehandler på baggrund af den lediges fødselsdag. Siden fødselsdage er så godt som tilfældige, svarer dette til at borgerne kvasi-tilfældigt tildeles en sagsbehandler. Kapitlet fremsætter tre sæt af resultater. For det første finder jeg, at variationen i sagsbehandler kvalitet kan forklare omkring 6% a variationen i afgangen fra ledighed inden for et jobcenter og år. Ledige borgere, der tildeles en sagsbehandler, hvis kvalitet ligger en standard afvigelse over gennemsnittet, vil afgå en uge tidligere fra ledighed. For det andet finder jeg, at den hurtigere afgang fra ledighed ikke har ugunstige effekter på senere præstation på arbejdsmarkedet. Tværtimod. De ledige borgere bliver beskæftigede i jobs, som hverken betaler en lavere løn, tilbyder færre timer eller er mere ustabile. Fordi de er startet i beskæftigelse en uge tidligere, har de også efter 2 år akkumuleret omtrent hvad der svarer til en uges fuldtidsbeskæftigelse i det gennemsnitlige job, 7.500 kroner og 35 timer. For det tredje finder jeg evidens for, at variationen i sagsbehandler kvalitet ikke nødvendigvis alene er drevet af ikke-observérbare personlighedstræk. Sagsbehandlere af højere kvalitet tenderer til at være mere 'pro-aktive': De holder møder tidligere og oftere med den ledige, sender den ledige tidligere i opkvalificering og så øger de den lediges brug af sit netværk samt uopfordrede ansøgninger. Umiddelbart indikerer kapitlets resultater, at det ville være en Pareto forbedring, hvis man lærte alle sagsbehandlere den 'pro-aktive' strategi.

Det andet kapitel (i samarbejde med Anders Humlum) undersøger om efteruddannelse kan øge lediges arbejdsmarkedstilknytning. For at undersøge dette kombinerer vi sagsbehandler datasættet fra det første kapitel med et nyt datasæt over sagsbehandleres anvisning til aktivering, herunder efteruddannelseskurser og jobrettede indsatser. Specifikt estimerer vi effekten af efteruddannelseskurser med et sagsbehandler strenghedsinstrument. Instrumentet udnytter, i) at ledige borgere kvasi-tilfældigt tildeles en sagsbehandler, og ii) at sagsbehandlere varierer i deres generelle tilbøjelighed til at anvise borgere til efteruddannelse (strenghed). Med vores sagsbehandler strenghedsinstrument kan vi ikke afvise at efteruddannelse i gennemsnit har ingen effekt på beskæftigelse efter et år. Det står i kontrast til OLS regressioner, som finder store negative sammenhænge mellem efteruddannelse og beskæftigelse. Det indikerer, at det er ledige med ugunstige jobudsigter, som selekterer ind i efteruddannelse. For at udforske om efteruddannelse er mere fordelagtigt for borgere, som er udsat for strukturelle ændringer, fx robot teknologi eller kinesisk importkonkurrence, zoomer vi ind på ledige, der tidligere var ansat i industrien. Selvom vores estimater bliver mere støjfyldte, finder vi økonomisk signifikante positive effekter af efteruddannelse på langt sigt.

Det tredje kapitel (i samarbejde med Nikolaj Harmon og Robert Mahlstedt) evaluerer effekten af at benytte et statistisk profilafklaringsværktøj til at informere borgeren og dennes sagsbehandler om den individuelle risiko for at blive langtidsledig. Dette er en indsats med relativt små udgifter, men som potentielt kunne øge afgangen fra ledighed. I Danmark har Styrelsen for Arbejdsmarked og Rekruttering udviklet et profilafklaringsværktøj, der bygger på en Machine

Learning algoritme. Profilafklaringsværktøjet informerer nyledige borgere og deres sagsbehandlere om, hvorvidt de tilhører en gruppe, der statistisk set er i høj risiko for blive ledig i mere end seks måneder. Ved at udnytte at den bagvedliggende algoritme har diskontinuiteter på alder, kan vi bruge et regressions diskontinuitets design (RDD) til at estimere effekten af at blive informeret om at være i høj-risiko. Vi finder, at borgere, som marginalt defineres som høj-risiko, er mellem 5-14 pct mindre tilbøjelige til at være ledige efter seks måneder. Efter 12 måneder er effekten forsvundet, da også borgere, der marginalt blev defineret som lav-risiko, er afgået fra ledighed. Uheldigvis finder vi ved standard validitetstjek, at den identificerende antagelse måske ikke holder. Selvom vores resultater indikerer, at statistisk profilering kunne øge afgangen fra ledighed, kræver det derfor en nærmere evaluering for at kunne drage en endelig konklusion.

Endelig bør det bemærkes, at det tredje kapitel i denne afhandling bygger på og gentager tekst fra mit kandidat speciale "Behavioral Responses to Information about Individual Employment Prospects: Evidence from Denmark" (Københavns Universitet, 2019). Specialet drejede sig mest om at rekonstruere den historiske version af profilafklaringsværktøjet. Det inkludere kun nogle enkelte og meget indledende RDD estimater. Projekt har ændret sig væsentligt, i mange og vigtige dimensioner, siden jeg forsvarede mit speciale. Eksempelvis har vi arbejdet meget med selve RDD analysen. Det inkluderer en forbedring af den initiale strategi, samt at konstruere to samples, som gør, at vi kan udforske heterogene effekter. Derudover har vi grundigt undersøgt de identificerende antagelser og udforsket de dynamiske effekter af interventionen.

Chapter 1

Caseworker Quality and Transitions Out of Unemployment: Evidence and Policy Implications

Caseworker Quality and Transitions Out of Unemployment: Evidence and Policy Implications

Mette Rasmussen*

March 1, 2021

Abstract

Can heterogeneity in caseworker quality explain part of the yet unexplained heterogeneity in transitions out of unemployment? And what is the policy implication? This paper estimates caseworkers' causal effects on jobseekers' transitions out of unemployment. I combine a novel data set that links jobseekers to caseworkers with rich Danish administrative data and exploit that jobseekers are quasi-randomly assigned to caseworkers in a large subset of the jobcenters in Denmark. The paper offers three sets of results. First, I find that heterogeneity in caseworker quality can explain about 6% of the heterogeneity in unemployment spells within a jobcenter and year. Assignment to a caseworker, who is one standard deviation above the mean, reduces the unemployment spell by about one week. Second, I find no smoking gun suggesting that this is at the expense of subsequent labor market performance. After two years, jobseekers assigned to a caseworker, who is one standard deviation above the mean, have on average accumulated an additional 7,500 DKK (1,200 USD) and 35 hours. This is because the jobseekers start about one week earlier in jobs of similar quality. Third, I show that high quality caseworkers tend to be more 'pro-active': They meet earlier and more frequently with the jobseekers and assign them earlier to training. Further, they make jobseekers increase the use of their network and send more unsolicited applications. The results suggest that caseworker quality is one dimensional, and that it would be Pareto improving to teach low quality caseworkers the pro-active strategies of the high quality caseworkers.

*University of Copenhagen; mette.rasmussen@econ.ku.dk

1 Introduction

There is large heterogeneity in transitions out of unemployment. While part of this is explained by jobseeker characteristics, a large part remains unexplained. Motivated by the fact that regular meetings with a caseworker constitute a main pillar in the public employment services across many OECD countries[1], this paper investigates what role *caseworkers* play. Caseworkers help jobseekers navigate out of unemployment, e.g. through advice on job search and training assignments. They likely have different strategies and take different actions, which could entail differences in effectiveness, or quality, across caseworkers. Extensive evidence shows that variation in the quality of school teachers and counselors has important consequences for later life outcomes (Chetty et al., 2014*a*,*b*; Jackson, 2018; Jackson et al., 2014; Mulhern, 2019; Rockoff, 2004). Similarly, caseworker quality could be important for transitions out of unemployment. This paper makes three contributions to our understanding of caseworker quality.

In a first step, I investigate to what extent *heterogeneity in caseworker quality explains the yet unexplained heterogeneity in transitions out of unemployment?* I define caseworker *quality* as the caseworker's individual impact on the speed with which jobseekers transition out of unemployment, and it may therefore also be denoted as caseworker *value-added* on transitions. I find that jobseekers assigned to caseworkers, whose value-added is one standard deviation above the mean, transition one week earlier out of unemployment. But does this simply reflect that caseworkers are trading off job quality against exit rates? It is well-known that job search entails a trade-off between exit rates and job quality (McCall, 1970). Presumably, not only jobseekers but also caseworkers face this trade-off, and to draw policy conclusions, it is important to understand how they manage the trade-off. Caseworkers could be specializing in specific dimensions, e.g. some caseworkers could improve exit rates at the expense of job quality, while others could do the opposite. In that case, caseworker quality should be evaluated along *multiple* dimensions, and it is not obvious that we should teach all caseworkers the strategy of the caseworkers facilitating fast exits. In a second step, I investigate the dimensionality of caseworker quality. In particular, whether *assignment to caseworkers with high value-added on transitions adversely affect subsequent labor market performance?* If there are no such adverse effects, it suggests that it would be Pareto improving to teach all caseworkers the strategy

[1] In 2018, Danish expenditures on PES admin and salaries was 0.38 % of GDP (OECD). See Rosholm (2014) for figures on other OECD countries.

of the high value-added caseworkers. To make concrete policy recommendations, the third part of the paper investigates *what high value-added caseworkers do differently?*

In this paper, I provide empirical evidence on the causal effects of individual caseworkers in the public employment services in Denmark. I thereby contribute to a sparse literature on the role of individual caseworkers in the labor market. The lack of studies is first of all a problem of data availability. It requires a data set that links caseworkers to jobseekers and contains information on the jobseekers' labor market outcomes. Second, identification requires random allocation of jobseekers to caseworkers. This paper overcomes both challenges.

First of all, I rely on a novel data set covering meetings between caseworkers and jobseekers in Denmark from 2011-18. By linking to rich administrative data, I obtain information on labor market outcomes of jobseekers over time. Based on novel data from the Danish employment authorities, I further enrich the panel with information on applied-for jobs, assignments to training and the frequency of caseworker meetings over the jobseekers' unemployment spell.

For identification, I exploit that a large subset of the jobcenters in Denmark assign jobseekers to caseworkers based on their birthday. Through a survey conducted among Danish jobcenters, I have information on local assignment rules over time. I verify that the jobcenters are using birthdays for assignment and in particular, that they use birth *day* of the month (1-31). I show that birth day of the month is *as-if random*, and I verify that this effectively means that jobseekers are *quasi-randomly* assigned to caseworkers.

Using state-of-the art methods from the teacher value-added literature (Chetty et al., 2014*a*; Jackson, 2018; Rockoff, 2004), I estimate causal effects of individual caseworkers on jobseeker transitions out of unemployment. In contrast to the majority of studies in the teacher value-added literature, however, I do *not* rely on a selection-on-observables approach. The paper most closely related to mine is Mulhern (2019), who exploits that school counselors are assigned to students based on the first letter of the last name. Instead of the first letter of the last name, I exploit that caseworkers are assigned to jobseekers based on birth day of the month. In either case, assignment is effectively quasi-random. Leveraging this quasi-random assignment, I can estimate causal effects of individual caseworkers on the length of the jobseekers' UI-spells. I use an empirical Bayes approach to reduce noise in the estimated caseworker effects, and in the spirit of the value-added literature, I denote the estimated caseworker effects as *value-added* or *quality*.

My paper offers three sets of results. First, I find substantial variation in caseworker quality. Within a jobcenter and year, heterogeneity in caseworker quality can explain about 6% of the heterogeneity in unemployment spells. In comparison, a rich set of jobseeker characteristics explains only about 3%. This suggests that caseworker quality is an important factor contributing to the heterogeneity in transitions. Assignment to a caseworker, whose value-added is one standard deviation above the mean, reduces the average UI-spell by about *one* week. Investigating the dynamics behind, I find large effects of value-added on exit probabilities early in the UI-spell. However, the effect diminishes over time as all jobseekers eventually exit.

Second, I investigate whether the positive effect on transitions is driven by caseworkers that trade off exit rates against labor market performance. I find no smoking gun. Two years after unemployment start, jobseekers assigned to caseworkers, whose value-added on transitions is one standard deviation above the mean, have on average accumulated an additional 7,500 DKK and 35 hours. This reflects that high value-added caseworkers push the majority of jobseekers from unemployment and into employment about one week earlier. Interestingly, the jobseekers are neither employed at lower wages nor in fewer hours, and they are not more likely to (re)enter unemployment or other public transfers. Hence, I find no evidence that sooner transitions out of unemployment come at the expense of the job quality. Finally, I find evidence suggesting that caseworkers not only improve transitions for the mean jobseeker but rather shifts the entire jobseeker distribution. Overall, the absence of negative effects on subsequent labor market performance suggests that it is Pareto improving to teach all caseworkers the strategies of the high value-added caseworkers.

Third, I find that high value-added caseworkers are more “pro-active”. They meet earlier and more frequently with the jobseekers and assign them earlier to training. Using registered job search as a proxy for actual job search, I find no evidence suggesting that higher value-added caseworkers increase the jobseeker’s search intensity. However, evidence suggests that they make jobseekers increase the use of their network and send more unsolicited applications. Although the pro-activeness may not explain the entire variation in value-added on transitions, it is an important insight for policy makers as it suggests that simple interventions potentially can reduce time spent in unemployment.

My paper contributes to a sparse literature on the individual impacts of caseworkers. Most of the existing studies rely on a conditional independence assumption. Using Swiss data, Behncke et al. (2010*a*) show that similarity between the caseworker and jobseeker increases the probability of employment, Behncke et al. (2010*b*) show that 'tougher' caseworkers increase employment, and Huber et al. (2017) note that this effect is not driven by assignments to active labor market programs (ALMPs) but rather by a particular 'counseling style'. Exploiting unplanned absences among caseworkers, Schiprowski (2020) shows that there is large heterogeneity in the effect of a face-to-face interaction across caseworkers in the productivity distribution. My research confirms that there are large differences in the individual impacts of caseworkers. I explicitly quantify how important this heterogeneity is for overall heterogeneity in jobseeker transitions out of unemployment.

In contemporaneous work, Cederlöf et al. (2020) examine caseworkers in Sweden by exploiting that Swedish - like Danish - jobseekers are assigned to caseworkers based on birthdays. Among other things, Cederlöf et al. (2020) also estimate the value-added of caseworkers. Reassuringly, the caseworker effects estimated in Cederlöf et al. (2020) are about the same magnitude as the ones estimated in my paper. In addition, my paper provides evidence on the *dimensionality* of caseworker quality, which is essential for drawing policy implications.[2] In particular, I show that caseworkers with higher value-added on transitions i) do not simply trade-off exit rates against job quality, and ii) improve transitions for the entire distribution of jobseekers rather than just for the mean jobseeker.

I further contribute to the discussion about the drivers of quality differences across caseworkers. Both Huber et al. (2017) and Schiprowski (2020) find that assignments to ALMP cannot explain the individual impacts of caseworkers, and Schiprowski (2020) notes that "The success of a caseworker thus appears to be mostly driven by unobserved personal qualities and counseling styles, which are difficult to replace". My paper suggests that differences in caseworker value-added are not just driven by unobservables: Caseworkers with high value-added assign jobseekers earlier to training programmes, and they meet earlier and more frequently with the jobseeker. The latter is in accordance with the investigation of successful counseling styles in Cederlöf et al. (2020). They find that 'active' caseworkers (those that have more frequent meetings) increase employment. However, Cederlöf et al. (2020) do not show that the 'active' caseworkers are the ones with high value-added on transitions. I directly link value-added to

[2]Cederlöf et al. (2020) also estimate caseworker value-added on earnings, but do not relate the effects on earnings to those on transitions. In principle, it could be different caseworkers that affect transitions and earnings (multi dimensionality), in which case the policy implication is unclear.

meeting frequencies, thereby suggesting that some caseworkers might have higher value-added *because* they are 'active'. Finally, it should be noted that I also find evidence suggesting that caseworkers with higher value-added induce jobseekers to send more unsolicited applications and make more use of their network. This channel comes closer to the 'unobserved counseling style' referred to by Schiprowski (2020).

Given the policy recommendation, this paper also relates to studies on caseworker discretion and a seemingly reluctance to change strategies. Conducting a randomized control trial, Behncke et al. (2009) show that caseworkers to a large extent ignore advice about optimal ALMP assignments, and Bolhaar et al. (2020) further show that not all caseworkers optimize their ALMP assignments even after learning about the programmes' effectiveness. This suggests that teaching low-value-added caseworkers the strategies of the high value-added caseworkers may not be simple but requires thought-through implementation.[3]

More broadly, this paper relates to the large literature on the effects of ALMPs (see Card et al., 2010, 2018). I show that the individual impacts of caseworkers offering the programs can explain about 6% of the variation in transitions out of unemployment. Further, my findings of potential mechanisms is in accordance with extensive evidence showing that earlier and more frequent meetings (Maibom et al., 2017; Van den Berg et al., 2012; Hägglund, 2011; Black et al., 2003) as well as early activation (Geerdsen, 2006; Geerdsen and Holm, 2007; Rosholm and Svarer, 2008) positively affect transitions out of unemployment. It is also in line with evidence showing that increased use of network in job search increases the employment probability (Glitz, 2017; Saygin et al., 2021).

Finally, my paper relates to the literature on individual impacts of e.g. school teachers (Chetty et al., 2014*a*,*b*; Jackson et al., 2014; Jackson, 2018; Rockoff, 2004; Kane and Staiger, 2008; Kane et al., 2013), school counselors (Mulhern, 2019), bosses and CEO's (Lazear et al., 2015; Bennedsen et al., 2020).

The rest of the paper proceeds as follows. Section 2 presents a simple model of caseworker value-added. Section 3 introduces the Danish unemployment insurance system and the assignment of caseworkers. Section 4 presents the data. Section 5 verifies that birthday assignment effectively result in a quasi-random assignment of jobseekers to caseworkers. Section 6 describes the empirical strategy. Section 7 presents the results and Section 8 concludes.

[3]It is probably easier to change the timing of meeting, yet a little more difficult to make caseworkers advice jobseekers on using their network and send unsolicited applications.

2 Model and identification

In this section, I present a simple model of caseworker value-added. Following the teacher value-added literature (Chetty et al., 2014*a*; Jackson et al., 2014; Kane and Staiger, 2008), I model individual labor market outcomes as a function of multiple input factors and explicitly separate the contribution of *caseworkers* from that of other input factors. Hereafter, I introduce my measure of caseworker quality and discuss identification.

2.1 Caseworker value-added model

Jobseeker i becomes unemployed in period t.[4] Upon unemployment, the jobseeker registers in her local jobcenter where she is assigned to a caseworker, $j = j(i)$. The labor market outcome of jobseeker i in period t is given by

$$Y_{it} = \beta X_{it} + v_{it} \quad , \quad v_{it} = \mu_{jt} + \epsilon_{it} \tag{1}$$

Here X_{it} represents observable factors, including jobseeker and jobcenter characteristics, and v_{it} is a residual representing unobservable factors. Importantly, the residual is comprised of a caseworker component, μ_{jt}, and idiosyncratic jobseeker-level variation, ϵ_{it}. One may think of the caseworker component, μ_{jt}, as the composite of a constant caseworker effect, μ_j, and random year-to-year fluctuations in caseworker effectiveness, ϕ_{jt}.[5]

$$\mu_{jt} = \mu_j + \phi_{jt} \tag{2}$$

The caseworker *effect*, μ_j, may be interpreted as the *quality* or *value-added* of the caseworker, since it captures the caseworker's individual impact on the labor market outcome Y_{it}. In this paper, I will define caseworker quality in terms of the caseworker's impact on the speed with which jobseekers transition out of unemployment. To measure transitions, I use the length of the individual's UI-spell as the labor market outcome, Y_{it}. The caseworker effects, μ_j, thereby capture a caseworker's individual effect on the length of the UI-spell. Note that throughout the paper, I will use the terms caseworker effect, quality and value-added interchangeably. These terms always refer to the caseworkers' individual impact on the speed with which jobseekers transition out of unemployment. In section 7.3, I discuss other quality measures and investigate the potential for caseworker quality to be multi dimensional.

[4]A jobseeker can have multiple unemployment spells, each of which are characterized by an (i,t) pair.
[5]The latter could be due to unplanned caseworker absences (illness) that temporarily affects their effectiveness.

2.2 Identification

Although caseworker effects, μ_j, are unobserved, it is important to note that they persist *across* jobseekers assigned to the *same* caseworker. This suggests that caseworker effects can be estimated by averaging over jobseekers assigned to the same caseworker. However, to obtain unbiased estimates of the causal caseworker effects, it is important to take potential selection of jobseekers to caseworkers into account. As noted in Jackson (2018), it is possible to obtain an unbiased estimate of the caseworker effects if there exists a set of conditioning variables T_{it}, such that jobseekers conditional on T_{it} are quasi-randomly assigned to caseworkers, $E(X_{it}|\mu_{jt}, T_{it}) = E(X_{it}|T_{it})$ and $E(\epsilon_{it}|\mu_{jt}, T_{it}) = E(\epsilon_{it}|T_{it})$. Conditional on T_{it}, the difference in average outcomes of jobseekers assigned to caseworker j and some other caseworker j' will be an unbiased estimate of caseworker j's individual effect.

To identify the effect of individual caseworkers, this paper exploits that a large subset of the Danish jobcenters assign jobseekers to caseworkers based on their *birthday*. Through a survey, I have information about which jobcenters that used birthday assignment from 2011-18 (see section 3). Relying on a novel data set, I use this information to construct a sample of jobseekers that were assigned to caseworkers based on birthdays (see section 4). In section 5, I show that this effectively means that I have a sample of jobseekers that were *quasi-randomly* assigned to caseworkers conditional on jobcenter unit and year, country of origin, quarter and age. Relating to the discussion above, this set of variables corresponds to T_{it}, which allows me to identify individual caseworker effects or *quality*. The identifying assumption is that jobseekers assigned to different caseworkers, conditional on T_{it}, have similar potential labor market outcomes. Given this assumption, conditional on T_{it}, the average UI-spell of jobseekers assigned to the same caseworker is an unbiased estimate of the caseworker's quality.

Two points regarding measurement of caseworker quality merit note. First, to use the quality measure for out-of-sample predictions, I will measure the caseworker's quality as the average UI-spell for all other jobseekers assigned to the same caseworker (leave-out mean). Second, to ease interpretation, I will code the quality measure such that higher quality caseworkers make jobseekers exit unemployment faster.

3 Institutional setting

In Denmark, unemployed individuals may receive unemployment insurance (UI) benefits up to two years. To be eligible, the individual must have paid contributions to a UI-fund[6] and have accumulated minimum one year of full time work over the previous 3 years.[7] The level of UI-benefits constitute 90% of prior monthly wages up to a maximum of 18,866 DKK (3,075 USD). In Denmark, around 85% of the Danish wage earnes are members of a UI-fund, and UI-benefit recipients account for around 75% of the unemployment in Denmark.

To remain eligible for UI-benefits, the jobseeker must live up to a number of requirements. Most importantly for this paper, the jobseeker will be assigned a caseworker from the municipal jobcenter and must attend regular meetings with the caseworker.[8] The assignment of a caseworker is described in detail in the next section. By law, the caseworker meetings should focus on concrete jobs, job search strategies and activation programs.[9] The frequency of meetings has increased over the last decade. Before 2015, jobseekers had to meet with a caseworker at least every third month[10], while the requirement since 2015 has been monthly meetings.[11]

Besides attending regular caseworker meetings, the jobseeker must be able to take a job with one day's notice and must actively search for jobs. The former includes accepting potential jobs referred to by the caseworker, and the latter, job search, must be in compliance with job search strategies laid out by the caseworker.[12] Since 2015, jobseekers have been required to document their job search on an online platform called *joblog*.[13] Although it is the UI-fund that sets a requirement for the minimum number of applied-for jobs registered in the joblog,[14] the caseworker from the jobcenter also monitors and base discussions on joblog activities. Hence, the caseworker can check whether job search is sufficient and in accordance with

[6]There are 24 UI-funds in Denmark which administer the payout of UI-benefits. UI-benefits are partly financed out of membership contributions and topped up by government subsidies. General eligibility requirements for UI benefit recipients are dictated by public policy. Yet, the UI-funds can dictate specific requirements. The UI-funds typically target a specific profession, occupation or education group. The UI-funds are organized in two larger organizations, the Trade Union Association and the Academic Association (a few UI-funds are not organized).

[7]Full time work can be in terms of hours (min. 1,924 working hours) or earnings (min. 238,512 DKK in 2020). Graduates are excepted from the latter requirement and are paid a reduced rate.

[8]It should be noted that the jobseeker also meet with the UI-fund, however, to a much letter extent, and I will focus entirely on caseworkers from the jobcenter hereafter.

[9]Lov nr 548 af 07/05/2019, §28-29

[10]LBK nr 990 af 12/09/2014, §17-19

[11]Monthly during the first 6 months, thereafter every third month. Lov nr 1486 af 23/12/2014, §16a

[12]BEK nr 1172 af 25/11/2019, §7, §14

[13]Lov nr. 548 af 07/05/2019 §20.

[14]Many UI-funds use universal thresholds of 1.5 - 2 joblogs/week (Fluchtmann et al., 2020)

strategies discussed in meetings. Finally, the jobseeker may be assigned to activation programs by the caseworker, e.g. job training or vocational training courses (see Humlum and Rasmussen, 2021). To remain eligible for UI-benefits, the jobseeker must comply with the assignments.[15]

3.1 Caseworker assignments and meetings

In Denmark, 94 jobcenters are responsible for the public employment services for individuals residing in a given municipality.[16] When an individual becomes unemployed, she will be assigned a caseworker from the jobcenter. It is key for this paper to understand how jobcenters assign caseworkers to jobseekers. In spring 2020, I conducted a survey among the Danish jobcenters to learn about this. The survey was sent via e-mail to the official mail box of each of the 94 jobcenters and comprised three questions regarding the organization of the jobcenter and especially, how caseworkers over time (2011-2018) have been assigned to jobseekers.[17] Often, a caseworker or the director would complete the survey on behalf of the jobcenter. A good 75 % of jobcenters completed the survey.[18] I now present four insights from the survey.

1. Jobcenters typically organize caseworkers in units

Many jobcenters in Denmark organize their caseworkers in smaller units. Typically, the units are based on jobseeker age (above or below 30) and profession (academic or non-academic).[19] I.e. jobseekers are first assigned to a unit based on observables. Subsequently, they are assigned to a caseworker within the unit. The latter assignment mechanism is key to this project and is discussed below. Since the assignment of caseworkers occur within a jobcenter unit and year, this is the level of randomization and I therefore include jobcenter unit and year indicators in all regressions. Based on the survey, I divide all jobcenters into four units by fully interacting jobcenter, profession and age indicators. To capture *profession*, I define an indicator for whether the jobseeker is part of the Trade Union Association,[20] and to capture *age*, I define an indicator

[15] Lov nr. 548 af 07/05/2019, §100-103.

[16] There are 98 municipalities in total, 4 jobcenters cover two municipalities.

[17] I sent out (multiple) reminders and interviewed some jobcenters via telephone. The survey questions were accompanied by examples. However, all answers were given in free text

[18] 66 jobcenters completed the survey, 5 could not help and 23 did not respond

[19] Anecdotally, the jobcenters tend to interpret 'academic' broader than just a college degree. Namely, 'academic' could also reflect the individual's previous profession or occupation.

[20] The 24 UI-funds in Denmark typically target specific professions and are broadly organized in two larger organizations: The Trade Union Association (FH) and the Academic Association (AC). A few UI-funds are not organized, and I will include them with jobseekers from the Academic Association

for whether the jobseeker is above 30.

2. Many jobcenters base caseworker assignment on jobseeker birthdays

There is variation in how jobseekers are assigned to caseworkers across jobcenters. E.g. some jobcenters base assignment on jobseeker birthdays while others base it on the jobseeker's previous industry. Even within the same jobcenter, there can be variation in assignment rules over time. Appendix table A1 shows how jobcenters are distributed across assignment rules in a given year from 2011-2018. Most importantly, the table shows that a large subset of the jobcenters report that they base caseworker assignments on *birthdays*. Out of the 66 responding jobcenters, 24 jobcenters report to use birthday assignment in the average year. In the analysis, I will rely on jobcenters that in a given year report to use birthday assignment. It should be noted that many of these jobcenters simply report that they use 'birthday assignment', without specifying whether they rely on birth day of the month (1-31), birth month (1-12), birth year (1946-2000) or the last (random) digit in the social security number. In section 5, I provide evidence clearly suggesting that the jobcenters are relying on birth day of the month (1-31). This is key to identification in this paper.

3. Assigned caseworker participates in the first individual meeting

The assigned caseworker invites and participates in the jobseeker's *first individual meeting* in the jobcenter. In the majority of jobcenters, this corresponds to the first meeting in the UI-spell. However, in a minority of the jobcenters, the meeting will be the second in the UI-spell, since the jobseeker first must participate in an information meeting with a group of other jobseekers. This knowledge allows me to identify the assigned caseworker from a data set containing information on the date and type of all meetings between jobseekers and caseworkers over the jobseeker's UI-spell.

4. Jobseekers tend to stay with the assigned caseworker

Although the jobseeker is supposed to have all meetings over the UI-spell with the assigned caseworker, this might not always be the case. Generally, there are two reasons for caseworker switches. First, it could be due to caseworker absences, e.g. illness or holidays, where a colleague replaces the assigned caseworker. Second, the jobseeker could actively request another caseworker. Here, it should be noted that jobseekers since 2016 have been required to book later meetings themselves (not the first meeting, which the assigned caseworker books).[21] The caseworker will tell the jobseeker when to book the next meeting, and the jobseeker must then book a meeting in due time in an online booking system. The booking system vary across jobcenters; in some jobcenters, the jobseeker is only allowed to book with the assigned caseworker, whereas in other jobcenters the jobseeker can choose between caseworkers. One might worry that the latter could lead to a surge in caseworker switching. However, many jobcenters report that they aim to minimize caseworker switching and strongly encourage the jobseeker to book with the assigned caseworker. In section 4.4, I show that jobseekers to a large extent stay with their assigned caseworker over the UI-spell. I show that is true both before and after self-booking of meetings was introduced. Hence, when I estimate the effect of assignment to a caseworker of a given quality, it will be the effect of multiple interactions with this caseworker.

4 Data

In this section, I introduce a novel data set on meetings between jobseekers and caseworkers, which I have been granted access to by the Danish Agency for Labor Market and Recruitment. Hereafter, I present the sample restrictions that I apply to the data in order to support my identification strategy. Based on Danish administrative data, I enrich the sample with relevant labor market outcomes and predetermined characteristics for the jobseekers. I further leverage new data on job search behavior and training assignments. Finally, I present descriptive statistics for the sample.

4.1 Caseworker Data

Since 2010, Danish jobcenters have been required to register all *meetings* held with jobseekers throughout their unemployment spell.[22] The meeting registrations are collected by the Agency for Labor Market and Recruitment and are used for monitoring of the meeting activity.[23] This has given rise to a novel data set covering meeting registrations from 2011-2018. It contains information on the date, time, type and contact form of each meeting as well as an authority, jobseeker and caseworker identifier. The jobseeker identifiers can be linked to Danish adminis-

[21] Lov nr. 548 af 07/05/2019, §34.
[22] BEK nr 418 af 23/04/2010, kap. 5
[23] Aggregated statistics on meeting activity in local jobcenters are made publicly available on jobnet.dk

trative data, whereas the caseworker identifiers cannot. I specifically focus on meetings between caseworkers and UI-benefit recipients[24] taking place in Danish jobcenters from 2011-2018.

Since I am the first to use the caseworker data, I have validated the quality. I summarize the three key findings here and refer to appendix B for a thorough examination of the validation exercise. First, I find that the data has a high coverage. Around 70% of new UI-spells have at least one meeting registration, and I find that the remaining UI-spells likely lack a registration because they exit unemployment before they even have a meeting.[25] Further, I find that the number of meetings registered over the UI-spell is in accordance with official rules. Second, I provide evidence strongly supporting the authenticity of the meetings. In particular, I find no bunching of meeting registrations on particular dates (e.g. December 31st), and I find that the vast majority of meetings are registered to take place Monday-Friday and during normal office hours for jobcenters. Third, I show that the caseworker identifiers are valid: The number of meetings registered per caseworker is well in accordance with anecdotal evidence on how many meetings a caseworker has in a typical work week. Overall, the validation exercise suggests that the caseworker data do in fact represent real meetings held between jobseekers and caseworkers in the Danish jobcenters.

4.2 Sample construction

I now explain the construction of the sample and refer to appendix table A2 for an overview of observations lost at each step. The starting point is the caseworker data, in particular the population of UI-benefit recipients who have at least one meeting registration with a caseworker in a jobcenter from 2011-2018. By linking to a register containing information on UI-benefit payments (DREAM), I construct UI-spells for all jobseekers in the sample. Here I follow the convention; I define a UI-spell as consecutive weeks with UI-benefit payments, while allowing for interruptions in payments of up to 3 consecutive weeks (Fluchtmann et al., 2020).[26] Thereby, I can order all meetings relative to the start date of the UI-spell.

In accordance with point 3 from section 3.1, I identify the *caseworker assigned* to a given

[24]Besides restricting to jobseekers that are classified as UI-benefit recipients in the meeting registration data, I cross check that these individuals do in fact receive UI-benefit within 4 weeks from a meeting registration

[25]These figures are for jobcenters that do not outsource all meetings with UI-benefit recipients to private entities (see appendix B).

[26] I define an exit from UI-benefits as 4 consecutive weeks with no UI-benefit payments. Likewise, I define an entry as the first time in which the individual receives UI-benefits after at least 4 consecutive weeks with no UI-benefits. The number of weeks from entry to exit gives the length of the UI-spell.

jobseeker as the caseworker from the jobseeker's first individual meeting in the UI-spell.[27] Hence, it is key that I correctly identify the first individual meeting. There could be different reasons why this might not be the case, e.g. if the meeting is not registered or if the UI-spell I have identified is actually a continuation of a previous UI-spell. I make a couple of restrictions to ensure that I identify the meeting correctly. First, I require that the first observed meeting for a given jobseeker takes place within 3 months from spell start, since this was required by law. Second, I drop UI-spells if they were initiated within 10 weeks from a previous UI-spell.[28] Third, I require that the first individual meeting is coded as a 'regular meeting' and to take place 'in person'. I drop the few jobseekers for whom I am unable to identify the assigned caseworker.

For identification, I rely on the fact that many jobcenters use birthdays to assign jobseekers to caseworkers. Based on the survey described in section 3.1, I therefore restrict the sample to all jobcenters and years in which the jobcenters report to base caseworker assignment on birthdays. To reduce noise, I further restrict the sample to caseworkers with at least 50 assigned jobseekers and who worked at least for one quarter. Since assignment to caseworkers occur within a jobcenter unit and year, I will include fully interacted jobcenter unit and year fixed effects in all regressions. I therefore also require at least 2 caseworkers per fixed effect cell. These restrictions leave me with a sample of jobseekers that - to the best of my knowledge - were assigned to caseworkers based on their birthday. In section 5, I test how well jobseeker birthdays predict assignment to a given caseworker in this sample, and based on this exercise I make one final restriction. Namely, to support my identification strategy, I drop caseworkers for whom jobseeker birthdays do not predict assignment well (see details in section 5). After this restriction, I obtain my final sample. This consists of 103,027 UI-spells distributed across 75,811 jobseekers, 467 caseworkers and 24 jobcenters.[29]

[27] I use the first meeting in the UI-spell, if this meeting was not coded an 'information meeting'. If the first meeting was an 'information meeting', I use the second meeting in the UI-spell, provided this was not an information meeting as well.

[28] When I examine whether assignment to caseworkers with higher value-added adversely affect subsequent labor market performance, I will consider re-entries into unemployment. Here the second UI-spell will count as a re-entry to UI-benefits within the first 10 weeks (see section 7.3).

[29] It should be noted that any of the sample restrictions potentially could affect my results. Yet, I save it for future work to test the robustness to the sample restrictions.

4.3 Labor market outcomes, mechanisms & characteristics

Labor market outcomes: Relying on Danish administrative data, I enrich the sample with relevant labor market outcomes. First, the income register (EINDKOMST) contain information on wages and working hours in a given month for the full population of wage earners in Denmark. I use this to measure accumulated wages and hours in a given month relative to UI-spell start. Second, the transfer register (DREAM) contains weekly information on transfers to the full Danish population. I use this register to construct mutually exclusive labor market states, such that I in a given week can see whether the jobseeker (re)enters unemployment, education subsidies, other public transfers, employment or drops out of the labor force.

Mechanisms: I will investigate three types of mechanisms - meetings, training and job search - through which caseworker quality might affect labor market outcomes. Note that these mechanisms are not mutually exclusive since e.g. advice on job search and assignment to training takes place at meetings.

First, some caseworkers could be better, because they *meet* more frequently or earlier with the jobseeker. To investigate this, I rely on the caseworker data, where I can see the frequency and timing of caseworker meetings over the UI-spell.

Second, some caseworkers could be better because they apply a different strategy for *assignments to training*. To investigate this, I rely on *job plan* data, which is described in detail in Humlum and Rasmussen (2021). Most importantly, the job plan data arises from the fact that caseworkers since 2010 have been required to prepare a job plan for the jobseeker. In the job plan, the caseworker notes all activities, such as training courses and job training, that she assigns the jobseeker to. With this data, I can therefore see whether the jobseeker is assigned to training, and when this training is set to start.

Third, some caseworkers could be better because they change individual *job search behavior*. To investigate this, I rely on *joblog* data. This data stems from the fact that jobseekers since September 2015 have been required to log applied-for jobs in a centralized system called joblog. Fluchtmann et al. (2020) show that the extensive margin coverage of the joblog data is high: 96% of new UI-spells register at least one job in the joblog. That said, it is not clear to what extent registered jobs in the joblog informs about the actual *number* of applied-for jobs. Fluchtmann et al. (2020) present evidence suggesting that many jobseekers log just the minimum requirement. I will therefore be cautious when interpreting on the *number* of joblogs per week. The real advantage of the joblog data, however, is that I can see the characteristics of the applied-for jobs. Fluchtmann et al. (2020) show that the applied-for jobs registered in the joblog indeed is representative of actual job applications.

For each applied-for job registered in the joblog, I can see what search method and channel the jobseeker made use of. The search methods refer to how the application was submitted (e.g. via e-mail, letter, in person) and search channels refer to how the job was found (e.g. unsolicited, advertised, through network). In the joblog data, I can also see the number of hours (part vs. full time) and commute distance (in kilometers) of the applied-for job. This allows me to investigate whether higher value-added caseworkers make jobseekers use more search methods and channels. Further, I can investigate whether they make jobseekers use their network, apply unsolicited or lower their job requirements (in terms of hours and commute distance).

Predetermined characteristics: I finally enrich the sample with information on jobseeker predetermined characteristics. Through the population register (BEF), I obtain information on demographics such as gender, age, origin, marriage status and number of children. I obtain information on highest completed education from the education register (UDDA).[30]. Based on the transfer register (DREAM), I can see past receipt of UI-benefits, education subsidies, parental leave subsidies or other types of public transfers. Further, I can see the previous industry as well as UI-fund. Finally, based on the income register (EINDKOMST), I obtain information on past wages, hours and number of employers.

4.4 Descriptive Statistics

I now present descriptive statistics for jobcenter units, caseworkers and jobseekers in the sample. This also includes a summary of UI-spells and labor market outcomes.

The jobcenter units are of particular interest, because they represent the level at which jobseekers are assigned a caseworker. Panel A of table 1 presents two key statistics describing the size of the jobcenter units. It shows that in the average jobcenter unit and year, there are around 7 caseworkers serving an inflow of 170 new jobseekers.

[30] I use the highest education completed up to one month after UI-spell start. This is to ensure that I see the relevant education for graduates who may not have received their diploma yet.

Table 1: Jobcenters and Caseworkers

	Obs	Mean	SD	Percentiles 10	25	50	75	90
A: Jobcenter $\times$ unit $\times$ years								
Caseworkers	606	6.95	3.49	3.00	4.00	6.00	9.00	12.00
Jobseekers	606	170.01	141.64	29.00	61.00	138.00	237.00	366.00
B: Caseworkers								
Caseload size	467	220.61	203.33	62.60	82.80	151.60	286.40	419.80
Meetings	467	985.46	814.36	226.00	367.40	765.60	1377.40	1990.40
Working weeks	467	84.53	53.12	27.80	42.00	74.00	106.00	174.80
Meetings / week	467	11.62	5.28	3.87	7.89	11.82	15.00	18.32

Note: To comply with data protection rules, all percentiles are based on at least 5 observations.
In panel A, the level of observation is a jobcenter $\times$ unit $\times$ year. It shows the number of caseworkers serving in a given jobcenter unit and year as well as the number of UI-spells initiated in a given jobcenter unit and year.
In panel B, the level of observation is a caseworker. Caseload size is the number of UI-spells assigned to the caseworker. This statistic involves only UI-spells from the final sample. The latter three statistics - meetings, working weeks and meetings/week - are based on all UI-spells in the caseworker data described in section 4.1. Meetings refer to the total number of meeting slots (unique time and date) registered for the caseworker. Working weeks refer to the total number of weeks in which the caseworker registered at least one meeting. Meetings/week is total meetings divided by working weeks for the caseworker.

Caseworkers: Panel B of table 1 presents a few descriptive statistics for caseworkers in the sample. The table shows that the mean (median) caseworker is assigned 220 (151) jobseekers over 84 (74) working weeks. Hence, the mean (median) caseworker works in about 1.6 (1.4) years in the jobcenter. Lastly, the table shows that the mean (and median) caseworker has approximately 12 meetings per week.

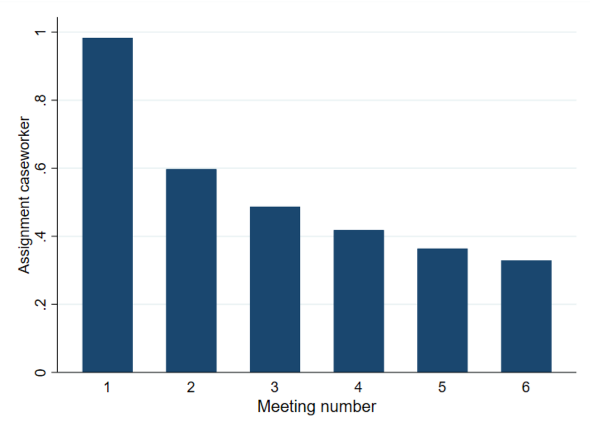
Figure 1a shows jobseeker survival with the assigned caseworker over the UI-spell. In particular, the figure plots the fraction of jobseekers that have a given meeting with the assigned caseworker conditional on surviving in unemployment up until that meeting. Since the vast majority of the jobseekers meet with their assigned caseworker in the first meeting of the UI-spell[31], the probability is close to 100% for the first meeting. The figure clearly shows that the probability of meeting the assigned caseworker decreases over time. E.g. from the first to the second meeting in the UI-spell, the probability of meeting the assigned caseworker drops to around 60%. In comparison, however, had the jobseeker randomly been assigned a *new* caseworker from the jobcenter unit and year, the probability that this would be the assigned caseworker would have been only 12.5%.[32] This suggests that jobseeker survival rates are fairly high. In fact, around 35 % of jobseekers are still meeting with their assigned caseworker when they get to the sixth meeting, which on average takes place in week 32 of the UI-spell.

[31]Recall that some jobseekers attend an information meeting before meeting with their assigned caseworker. Hence, these jobseekers meet the assigned caseworker in the second meeting.

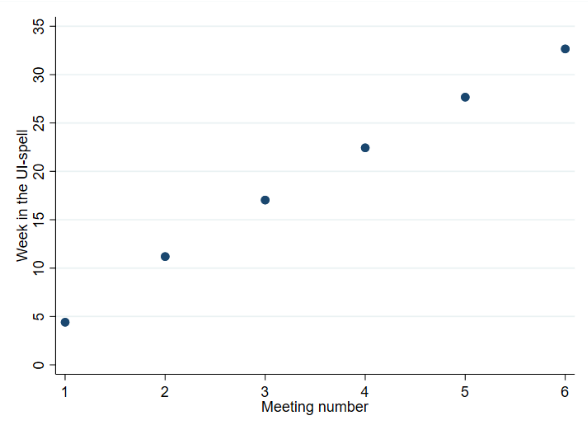
[32]The average jobseeker belongs to a jobcenter unit and year with 8 caseworkers.

One might worry that requiring jobseekers to book later caseworker meetings themselves would cause a significant decrease in caseworker survival. However, appendix figure A1 shows survival before and after September 2016, where self-booking of meetings became mandatory. The figure clearly shows that self-booking has not caused more caseworker switching. This indicates that, over the entire period of interest (2011-18), assignment to a given caseworker *is* predictive of the caseworker participating in later meetings. Hence, when I measure the impact of assignment to a high quality caseworker, I will be estimating the impact of multiple interactions with the caseworker.

Figure 1: Caseworker meetings & survival



(a) Caseworker survival



(b) Timing of meeting

Note: Panel 1a shows the fraction of jobseekers that have a given meeting in the UI-spell with their assigned caseworker (conditional on having the given meeting). The fraction is not one for the first meeting, since some jobseekers start with an information meeting and hence meet their assigned caseworker in the second meeting. Panel 1b shows the average timing of a meeting (conditional on having the meeting) relative to the UI-spell start.

Figure 1b shows the average timing of the different meetings in the UI-spell. Interestingly, the figure shows that jobseekers on average have their first meeting in week 4 of the UI-spell, which for the vast majority of jobseekers is the meeting in which they see their assigned caseworker for the first time. Note, however, that this need *not* imply that caseworker quality only can affect exit rates after week 4. Namely, the caseworker must first *invite* the jobseeker to the meeting, and there is much evidence showing that the mere invitation to a meeting have significant and positive effects on the exit rate (Hägglund, 2011; Black et al., 2003; Van den Berg et al., 2012; Maibom et al., 2017).

Jobseekers: Appendix table A3 presents summary statistics for jobseekers in the sample. It shows that there are equally many males and females in the sample, and that jobseekers on average are 40 years old and mainly non-immigrants. Around 50% of jobseekers have a vocational education, and a good 20 % have primary education only. Most of the jobseekers were employed in the previous year (90%) and earned about 210,000 DKK (34,000 USD). The jobseekers come from various industries, with the largest being public administration, health and education (22%) and trade and transportation (20%). Although the vast majority of jobseekers were employed in the past, a large chunk also experienced unemployment (40%) and received some kind of public transfers (60%). Finally, the table shows that the majority of jobseekers are members of a UI-fund organized under the Danish Trade Union Association (65%).

UI-spells: Panel A of table 2 summarizes the distribution of UI-spells for the jobseekers. This is a key outcome, since I will base the measure of caseworker quality on the length of the jobseekers' UI-spells. Therefore, it is important to note that some of the UI-spells in the sample are censored. The censoring occurs because i) my sample consists of jobseekers initiating a UI-spell from 2011-2018, and ii) I only have data on UI-benefit payments until ultimo 2019. This means that I can follow all individuals in the sample for at least 52 weeks after unemployment start.[33] Hence, no UI-spells below 52 weeks are censored. However, for jobseekers that initiate their UI-spell in 2018, the UI-spell may be censored if it exceeds 52 weeks.

There are two simple ways to deal with the censoring of UI-spells. One solution is to truncate all UI-spells at 52 weeks. Since the majority of jobseekers exit unemployment within 52 weeks, the truncation only affects 12% of all UI-spells. Another solution is to drop an entire year from the sample period, i.e. rely only on UI-spells initiated from 2011-2017. However, since I only have data on job search behavior from September 2015 to 2018, this would be very costly in terms of power. For this reason - and because it only affects 12% of UI-spells - I will use UI-spells truncated at 52 weeks for the remainder of the paper. Note that I test that my results are robust to truncation in section 7.1

Panel A of table 2 shows that truncated UI-spells have a mean (median) of 23.6 (19) weeks. Across all jobseekers, the standard deviation is 16 weeks, while it is 15.6 weeks across jobseekers within the same jobcenter unit and year. This shows that time and geography explains only little of the heterogeneity in UI-spells. Appendix table A4 further shows most of the heterogeneity remains if I simply control for jobseeker characteristics, e.g. education and past labor

[33]Recall that the maximum benefit duration is 104 weeks

market performance. With a rich set of jobseeker characteristics, I am only able to explain about 3% of the heterogeneity.[34]. In section 7.1, I investigate how much of the heterogeneity in UI-spells one can explain with heterogeneity in caseworker quality.

Table 2: UI-spells, hours and wages

			SD		Percentiles				
	Obs	Mean	Total	Within	10	25	50	75	90
A: UI-spells									
Raw	103,027	27.0	23.9	23.3	6.0	11.0	19.0	35.0	62.0
Truncated at 52 weeks	103,027	23.6	16.0	15.6	6.0	11.0	19.0	35.0	52.0
B: Accumulated hours and wages 24 months after unemployment start									
Accumulated hours	89,249	1,857.2	1,167.3	1,132.6	25.5	807.0	2,044.0	2,864.0	3,300.3
Accumulated wages	89,249	339,052.7	238,786.2	231,202.7	4,471.4	129,220.6	341,606.9	511,004.6	646,844.2

Note: Column 1 shows the number of observations, column 2 the unconditional mean, column 3 the total standard deviation and column 4 the standard deviation within jobcenter unit and year. Column 5-9 shows percentiles in the raw data. To comply with data protection rules, all percentiles are based on at least 5 observations. *Raw UI-spells* are defined as consecutive weeks with UI-benefit payments, while allowing for interruptions of up to 3 weeks. Some of the UI-spells will be censored. *Truncated UI-spells* corresponds to raw UI-spells truncated at 52 weeks. *Accumulated hours and wages* are based on the income register collected by the Danish tax authorities (Eindkomst). This register contains monthly information on hours and wages for all employees.

Labor market performance: Panel B of table 2 summarizes *accumulated wages and hours* for jobseekers two years after unemployment start. The table shows that the mean jobseeker has accumulated around 340,000 DKK and 1,900 hours two years after unemployment start. Interestingly, the standard deviation within a jobcenter unit and year is a solid 230,000 DKK and 1,100 hours. This suggests that there is a substantial heterogeneity not only in UI-spells but also in accumulated wages and hours across jobseekers *within* the same jobcenter unit and year. Although predetermined characteristics can account for part of this heterogeneity (10%), the largest part remains unexplained (see appendix table A4).

Aside from accumulated wages and hours, I will also investigate the labor market state that jobseekers enter over time. Namely, an exit from unemployment does not necessarily imply that the jobseeker enters employment. Appendix table A5 shows that although the majority of jobseekers (50%) become employed upon exit from unemployment, a non-negligible share eventually re-enter unemployment or other types of public transfers (around 40%), which could indicate that the job match quality was not the best. In 7.3, I will investigate whether this could reflect that some caseworkers trade of exit rates against job quality.

Finally, appendix table A6 present descriptive statistics for meeting frequency, job search and training assignments. These are the behaviors I will investigate in 7.4. What should be noted from the table is that the behaviors exhibit large variation, and hence, potentially could explain the heterogeneity in caseworker quality.

[34]The standard deviation reduces from 15.6 to 15.2 when controlling for all characteristics from A3

5 Validation of caseworker assignments

In the previous section, I constructed a sample of jobseekers that - to the best of my knowledge - were assigned to caseworkers based on their birthday. I now validate that this is indeed the case. In particular, I provide evidence suggesting that i) caseworker assignments are based on birth *day* of the month (1-31), and that ii) birth *days* are as-if random. This suggests that caseworkers effectively are quasi-randomly assigned to jobseekers. I provide a balance test supporting this conclusion.

5.1 Validation of birthday assignment

The validation of birthday assignments builds on a simple notion: If caseworker assignment is based on jobseeker birthdays, I should be able to *predict* assignment to a given caseworker within a jobcenter unit and year from information about jobseeker birthdays. Motivated by this, I run a regression for each caseworker j in the sample while only including jobseekers from the corresponding jobcenter, k. In particular, I define a dummy, $D(j)$, taking value one if jobseeker i was assigned to caseworker j. If caseworkers worked in all years and birthday rules were static, I would simply regress this dummy on birthday fixed effects along with fully interacted jobcenter, unit and year fixed effects, $\gamma_{k\times u\times y}$, the latter to take account of the fact that birthday assignment occurs within a jobcenter unit in a given year. However, caseworkers may not work in all periods and they might change birthday responsibility over time. To account for this, I include fully interacted birthday and quarter fixed effects, $\alpha_{b\times q}$, along with the jobcenter unit and year fixed effects.

$$D(j)_{ikuyqb} = \alpha_{b\times q} + \gamma_{k\times u\times y} + u_{ikuyqb} \quad , \quad \forall i \in k \tag{3}$$

As outlined in section 3.1, the jobcenters have reported that they use ’birthdays’ for assignment without specifying whether they rely on birth days (1-31), months (1-12) or years (1946-2000). I test all three interpretations by estimating equation (3) with the birthday effects representing days (1-31), months (1-12) and years (1946-2000).

To capture the predictive power of birthdays *within* a jobcenter unit and year, I save the within-jobcenter-unit-year-r^2 (within-r^2, hereafter) after each estimation of equation (3). In general, the higher the within-r^2, the better jobseeker birthdays (and quarters) predict caseworker assignment.[35] Besides a value of one, however, it is not obvious what a 'high' within-r^2 is. In fact, since I do not know the precise birthday rule, one wouldn't expect a within-r^2 of one.[36] To have a benchmark, I assign all jobseekers a placebo birthday[37] and save the within-r^2 from regressions similar to (3) including placebo instead of true birthdays.

Figure 2 presents the result from the birthday validation exercise. In all panels, the red line represents the distribution of the within-r^2 from regressions with *true* birthdays, whereas the black line represents that of regressions with *placebo* birthdays. Panel 2b and 2c show the result with birth months (1-12) and years (1946-2000). Clearly, the true and placebo distributions coincide in both figures. I.e. with the specification in (3), the placebo birth months and years are as good as the true birth months and years at predicting caseworker assignment. Hence, I find no evidence that jobcenters are using birth months nor years for assignment. Panel 2a shows the result with birth days (1-31). Here, the true distribution is clearly shifted to the right compared to the placebo distribution: True birthdays seem to be significantly better at predicting caseworker assignment compared to placebo birth days. This suggests that the jobcenters are using some type of birth day of the month (1-31) assignment rule.

5.2 Birthdays are as-if random

Having provided evidence that jobcenters in the sample use birth day of the month for caseworker assignment, I now test whether birth days are as-if random. To test this, I check whether jobseekers with specific characteristics are born on specific days of the month. I do this by regressing jobseeker predetermined characteristics on birth day of the month (1-31) fixed effects as well as jobcenter and year fixed effects.[38] In each regression, I test for joint significance of the birth day fixed effects.

Appendix table A7 reports the p-value on the F-test for joint significance of the birth day fixed effects. The first column shows that the birthday fixed effects are highly insignificant for the large majority of jobseeker characteristics. However, it also reveals that birth days are significant predictors of whether a jobseeker is of western or non-western origin (as well as

[35] Note that the within-r^2 also could be 'high' mainly due to over-fitting.

[36] As outlined, I do not know the exact birthday rule. Further, I allow the rule to switch every quarter, but the rule could be switching every week.

[37] I draw placebo birth days (months) by drawing integers from a uniform distribution defined from 1-31 (1-12). I draw placebo birth years from the sample distribution.

[38] The latter set of fixed effects is to allow for regional differences.

marriage and the number of children). This is likely driven by an institutional feature of the Danish immigration system. Anecdotally, if an immigrant arrives in Denmark without a birth certificate, she will be assigned January 1st as her birth day.[39] Since these immigrants mainly originates from non-western countries, there will be an over-representation of non-western jobseekers on January 1st. This is exactly what I see in the data: Non-western jobseekers constitute less than 5% of the entire sample, yet they represent more than 30% of jobseekers with birthday on January 1st. To test whether this is driving imbalances in other characteristics, I again regress jobseeker characteristics on birth day fixed effects while controlling for non-western origin. The result is shown in column 2, and it is evident that the imbalance in marriage and number of children were driven by non-western jobseekers. This exercise shows that birth days are as-if random conditional on country of origin.

5.3 Quasi-random assignment

The birthday validation exercise provides clear evidence suggesting that caseworker assignment is based on birth days (1-31) within a jobcenter unit and year. On top of this, I have shown that, conditional on country of origin, birth days are as-if random. Taken together, it suggests that caseworkers effectively are *quasi-randomly* assigned to jobseekers within a jobcenter unit and year. Since I will rely on this for identification of caseworker quality, I conduct a randomization test to verify that this is the case.

First, I make one final sample restriction to support the identification strategy. Based on the birthday validation in section 5.1, I want to keep the caseworkers for whom jobseeker birthdays were 'good' at predicting assignment. As discussed above, it is not obvious what the criterion for a 'good' prediction is. Further, choosing the criterion introduces a trade-off between precision (increasing the likelihood that birthdays were used for assignment) and power. I will set the criterion based on the placebo test. In particular, I restrict the sample to caseworkers for whom the within-r^2 on true birth days is above the median within-r^2 on placebo birth days.[40] Any criterion could of course affect the results, and I leave it for future work to test the robustness to this criterion. I rely on this sample throughout the rest of the paper.

I now conduct a randomization test inspired by Bhuller et al. (2020). The idea is to show

[39] This is also described in a Danish newspaper article: https://politiken.dk/debat/art8048808/Hvorfor-har-12-procent-af-mine-patienter-f%C3%B8dselsdag-1.-januar

[40] Note that I once again check that I have at least 2 caseworkers per randomization cell after this final restriction.

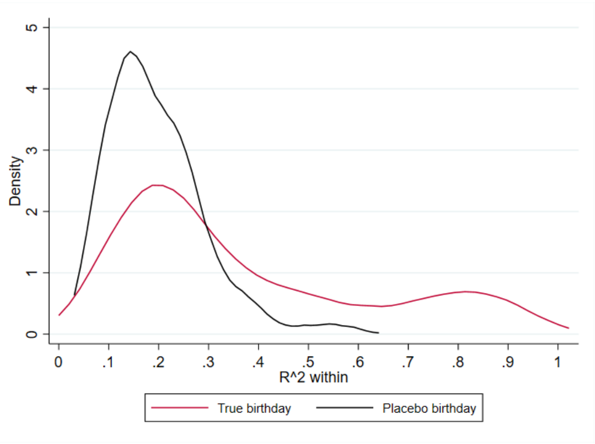
that higher quality caseworkers are no more likely to be assigned to a specific type of jobseeker. As described in section 2, I measure caseworker quality as the leave-out mean of (truncated) UI-spells, coded such that higher quality caseworkers make jobseekers exit faster from unemployment. To ease comparison, I standardize the quality measure as well as truncated UI-spells, such that they both have mean zero and standard deviation one.

I now regress first the truncated UI-spells and second the caseworker quality measure on predetermined jobseeker characteristics along with fully interacted jobcenter unit and year fixed effects, a dummy for non-western origin, quarter and age fixed effects. The jobcenter unit and year fixed effects are included to compare jobseekers assigned to different caseworkers at the level of randomization. I take out nation-wide differences common for jobseekers with a non-western origin, since I showed in the previous section that birthdays are as-if random conditional on country of origin. I include quarter fixed effects to take out potential seasonality and business cycles.[41] Although I find no indication that jobcenters use birth years for assignment (corresponding to age-based assignment), I include age fixed effects to be on the safe side as age is strongly correlated with labor market outcomes.

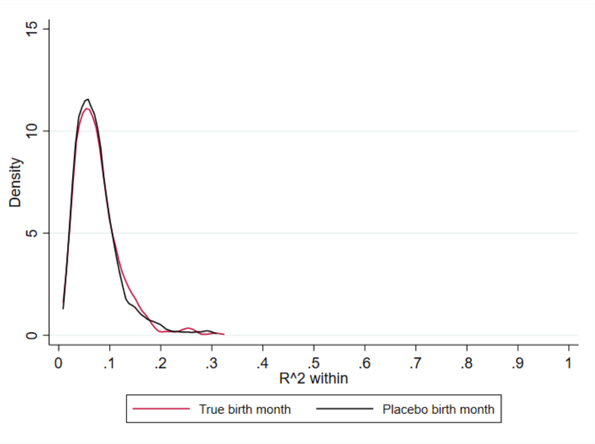
The randomization test is presented in table 3. The first column shows the coefficients on predetermined characteristics when using jobseeker UI-spells as the dependent variable. It shows that jobseeker characteristics strongly predict UI-spell length and hence are relevant to include in the randomization test. The second column is the actual randomization test, where the caseworker quality measure is used as the dependent variable. It shows that the majority of characteristics individually are unable to predict whether the individual is assigned a higher quality caseworker. Only a few characteristics are individually significant, e.g. 'master degree' and 'number of employers in t-1'. Although statistically significant (at the 10% and 5% level), their economic significance is limited (coefficients are 1/5 and 1/10 the size of those in the first column). The F-stat and corresponding p-value in the bottom of the table show that the characteristics jointly are insignificant. Hence they are jointly unable to predict caseworker assignment. Overall, I take this as evidence that jobseekers in the sample are quasi-randomly assigned to caseworkers.

[41] This may be important because caseworkers do not necessarily work in all months of the year: If there is seasonality in labor demand, one caseworker will look more effective than the other merely because she was working during high demand season and the other caseworker was working during low demand season.

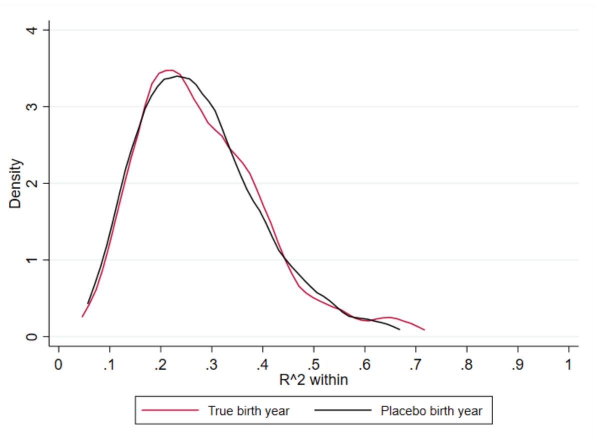
Figure 2: Birthday validation



(a) Birth day (1-31)



(b) Birth month (1-12)



(c) Birth year (1946-2000)

Note: The panels show within-r^2 distributions obtained by running a regression for each caseworker in the sample. For a given caseworker, a dummy for assignment to that caseworker is regressed on fully interacted birthday and quarter fixed effects along with jobcenter unit and year fixed effects. Note that only jobseekers from the caseworker's jobcenter are included in the estimation. The caseworker regression is run both with true and placebo birthdays, and the within-r^2 is saved after each regression. This captures how much of the variation within jobcenter unit and year that may be explained by birthday and quarter fixed effects. In panel a (b), birthdays represent days (months), and the placebo birthdays represent random draws for a uniform distribution defined from 1-31 (1-12). In panel c, birthdays represent years and the placebo birthdays represent random draws from the true birth year distribution (1946-2000).

Table 3: Randomization Test

	(1)		(2)		(3)	
	UI-spell		Leave-out mean		Covariates	
	b	se	b	se	mean	sd
Demographics[(1)]						
Male	-0.088***	(0.008)	0.001	(0.005)	0.484	0.500
Immigrant	0.201***	(0.016)	-0.013	(0.012)	0.050	0.218
Descendant	0.082	(0.074)	-0.004	(0.031)	0.002	0.043
Married	-0.039***	(0.008)	0.001	(0.005)	0.422	0.494
Number of children	0.002	(0.003)	0.000	(0.002)	0.806	1.039
Education[(2)]						
0. Missing	-0.027	(0.030)	-0.018	(0.015)	0.013	0.115
15. Preparatory course	0.291***	(0.048)	0.010	(0.030)	0.004	0.062
20. Upper secondary	0.048***	(0.016)	0.013	(0.009)	0.047	0.212
30. Vocationa educ.	-0.059***	(0.009)	0.004	(0.006)	0.476	0.499
35. Qualifying program	-0.019	(0.125)	-0.044	(0.060)	0.000	0.022
40. Short cycle tertiary	0.041**	(0.017)	0.005	(0.013)	0.047	0.211
50. Vocational bach.	-0.030**	(0.013)	-0.009	(0.009)	0.125	0.331
60. Bachelor	0.174***	(0.029)	-0.032*	(0.018)	0.014	0.118
70. Master	0.150***	(0.019)	-0.031*	(0.016)	0.049	0.215
80. PhD	0.148**	(0.062)	-0.055	(0.043)	0.003	0.053
Labor market history[(3)]						
UI-benefits in year t-1	-0.174***	(0.010)	0.010	(0.008)	0.408	0.491
UI-benefits in year t-2	-0.111***	(0.010)	0.010	(0.007)	0.424	0.494
Any employment in year t-1	-0.167***	(0.018)	0.007	(0.008)	0.900	0.300
Any employment in year t-2	0.069***	(0.015)	0.016*	(0.008)	0.922	0.268
Employment rate in year t-1	-0.068***	(0.025)	-0.021	(0.015)	0.642	0.352
Employment rate in year t-2	0.057***	(0.020)	-0.036*	(0.018)	0.659	0.348
Wage earnings in 1,000 DKK in year t-1	-0.000***	(0.000)	-0.000	(0.000)	212.235	147.811
Wage earnings in 1,000 DKK in year t-2	-0.000***	(0.000)	0.000	(0.000)	209.295	141.285
Number of employers in year t-1	-0.078***	(0.004)	0.006**	(0.003)	1.398	0.944
Number of employers in year t-2	-0.018***	(0.004)	-0.001	(0.002)	1.407	0.907
Public transfers in year t-1	0.051***	(0.009)	-0.003	(0.009)	0.627	0.484
Parental leave in year t-1	-0.029**	(0.013)	-0.000	(0.008)	0.083	0.276
Education subsidy in year t-1	-0.158***	(0.014)	-0.013	(0.009)	0.101	0.302
Previous industry[(4)]						
Real estate	0.231***	(0.030)	-0.010	(0.018)	0.011	0.106
Business services	0.194***	(0.011)	-0.010	(0.007)	0.115	0.320
Finance	0.406***	(0.034)	-0.011	(0.022)	0.010	0.098
Trade & transport	0.241***	(0.011)	-0.007	(0.007)	0.199	0.400
Manufacturing	0.187***	(0.015)	0.034	(0.023)	0.123	0.328
Communication & it	0.346***	(0.026)	-0.013	(0.015)	0.017	0.130
Culture	0.130***	(0.018)	-0.005	(0.010)	0.036	0.185
Agriculture, forestry & fishing	0.076***	(0.018)	0.012	(0.011)	0.026	0.159
Public administration, health, education	0.122***	(0.012)	0.002	(0.008)	0.224	0.417
Obs	103027		103027		103027	
Dep var Mean	0.000		-0.000			
Dep var sd	1.000		1.000			
Number of FE's	677		677			
F-stat (all cov)	127.171		1.213			
P-value (all cov)	0.000		0.187			

Note: *p<0.10 ** p<0.05 *** p<0.01. The first two columns report coefficients from a regression including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. The dependent variable is UI-spells (truncated at 52 weeks) in column 1, while it is caseworker quality (leave-out mean truncated UI-spells, multiplied by minus one) in column 2. In both columns, the dependent variable has been standardized to make comparison easier. Standard errors are two-way clustered at the caseworker and jobseeker level. Column three reports means and standard deviations of the covariates.

(1) *Demographics* rely on information from the population register (BEF and DREAM). Male, immigrant, descendant and married are dummies, while number of children is a count variable.

(2) *Education* rely on information from the education register (UDDA) and is based on the highest completed education (education completed up to 1 month after spell start is included). Omitted category is "10 Primary education".

(3) *Labor market history* variables rely on a register containing weekly information on UI-benefits and transfers (DREAM) and on the income register (Eindkomst). UI-benefits, any employment, public transfers, parental leave and education subsidy are all dummies. The employment rate, wages and number of employers are continuous and winsorized at the 99th percentile.

(4) *Previous industry* is based on the DREAM-register. It represents the dominating industry for the individual in the 12 months prior to the UI-spell start (the industry in which the individual had highest accumulated earnings). Omitted category is "Construction".

6 Empirical strategy

In the previous section, I showed that I have a sample of jobseekers that effectively were *quasi-randomly* assigned to caseworkers conditional on jobcenter unit and year, country of origin, quarter and age. In this section, I exploit this for identification of the causal effect of individual caseworkers. First, I describe how I obtain an unbiased estimate of the *magnitude* (variance) of caseworkers' individual effects on jobseeker transitions out of unemployment. This will allow me to answer the first question raised in the paper: *To what extent can heterogeneity in caseworker quality explain the heterogeneity in transitions out of unemployment?*

Second, I construct a measure of caseworker quality that allows me to do out-sample predictions. I.e. I will be able to identify the causal effect of assignment to a caseworker, who is one standard deviation above the average, on various labor market outcomes. Thereby, I can address the second and third question raised in the paper: *Does assignment to a caseworker with high value-added on transitions adversely affect subsequent labor market performance? And what are the high value-added caseworkers doing differently?*

6.1 Estimation of caseworker effects

I now estimate caseworkers' individual effects on jobseeker transitions out of unemployment. The ultimate goal is to obtain an unbiased estimate of the *variance* of the caseworker effects. In section 7.1, I will relate this variance to the variance in jobseeker transitions out of unemployment, thereby quantifying the importance of caseworker quality.

As described in section 2.2, I use the length of the (truncated) UI-spell to measure the jobseeker's rate of transition out of unemployment, Y_{it}. To identify the individual caseworker effects, I exploit that I have a sample of jobseekers that were *quasi-randomly* assigned to caseworkers conditional on jobcenter unit and year, country of origin, quarter and age. In all regressions throughout I will therefore include fully interacted jobcenter unit and year fixed effects, $\gamma_{k\times u\times t}$, a control for country of origin, δ_o, quarter and age fixed effects, κ_q and ψ_a (see discussion in section 5.3). The identifying assumption is that jobseekers assigned to different caseworkers, conditional on $T_{it} = \{\gamma_{k\times u\times t}, \delta_o, \kappa_q, \psi_a\}$, have similar potential labor market outcomes (UI-spell length). Since my identification strategy is very similar to that of Mulhern (2019), I follow her estimation approach closely.

1. FE-estimation

Given the identifying assumption, the average UI-spell for jobseekers assigned to the same caseworker will be an unbiased estimate of the caseworker's effectiveness conditional on T_{it}. Hence, I can estimate the following equation by OLS and obtain a fixed effect estimate of individual caseworker quality, $\hat{\mu}_j^{FE}$, for each caseworker in the sample:

$$Y_{it} = \alpha + \mu_j + \gamma_{k\times u\times t} + \delta_o + \kappa_q + \psi_a + \epsilon_{it} \tag{4}$$

Although the fixed effect estimates, $\hat{\mu}_j^{FE}$, are unbiased, they will contain estimation error. In particular, they will be *noisy* because caseworkers only are assigned a finite number of jobseekers (Jackson et al., 2014). Not taking account of this would lead me to overestimate the variance of caseworker effects.

2. Empirical Bayes approach

Following the teacher value-added literature, I use an Empirical Bayes approach to reduce estimation error in the caseworker estimates. Namely, by assuming that caseworker effects are normally distributed with mean zero, I can model the estimation error and shrink noisier estimates towards zero according to a signal-to-noise ratio.

Similar to Mulhern (2019), I use restricted maximum likelihood to estimate a mixed model that treats the caseworker effects, μ_j, as random. The model is similar to equation (4) except that it also includes a caseworker-year random effect, ϕ_{jt}.[42] This allows me to separate random year-to-year fluctuations in caseworker effectiveness, ϕ_{jt}, from the constant part, μ_j.

$$Y_{it} = \alpha + \mu_j + \phi_{jt} + \gamma_{k\times u\times t} + \delta_o + \kappa_q + \psi_a + \epsilon_{it} \tag{5}$$

Under the assumption of joint normality, I obtain consistent estimates for the variance of constant and time-varying caseworker effects, σ_μ^2, and σ_ϕ^2, and jobseeker errors, σ_ϵ^2. The former, $\hat{\sigma}_\mu^2$, is of particular interest and will be used in section 7.1 to quantify the importance of caseworker quality. It is instructive to note that I by estimation of (5), obtain an empirical Bayes estimate, μ_j^{EB}, for each caseworker. The empirical Bayes estimate corresponds to a simple caseworker mean, $\bar{\mu}_j$ multiplied by a caseworker-specific signal-to-noise ratio, λ_j:

$$\hat{\mu}_j^{EB} = \bar{\mu}_j \cdot \lambda_j \quad , \qquad \lambda_j = \frac{\hat{\sigma_\mu^2}}{\hat{\sigma_\mu^2} + (\sum_t(\hat{\sigma_\phi^2} + \frac{\hat{\sigma_\epsilon^2}}{n_{jt}})^{-1})^{-1}} \tag{6}$$

[42]In equation (4), ϕ_{jt} is part of the error term, ϵ_{it}.

It is evident that noisier estimates, e.g. due to fewer assigned jobseekers ($n_{jt} \downarrow$), higher jobseeker-level error ($\hat{\sigma_\epsilon^2} \uparrow$) or higher year-to-year fluctuations ($\hat{\sigma_\phi^2} \uparrow$), are shrunk towards zero.

6.2 Out-of-sample predictor for caseworker quality

I now construct an out-of-sample predictor for caseworker quality that can be used in a regression to predict the effect of assignment to a caseworker, who is one standard deviation above the mean, on various labor market outcomes.

First of all, to avoid mechanical endogeneity, the estimate for caseworker quality should not be based on the jobseeker, whose outcome I am trying to predict. Following the teacher value-added literature, I therefore form leave-one-out empirical Bayes estimates of caseworker quality. I obtain these by multiplying a caseworker-level leave-out mean[43], $\bar{\mu}_{j,-it}$, by the shrinkage factor, λ_j (see Jackson, 2018). To ease interpretation, I do two further twists. First, I multiply the leave-one-out empirical Bayes estimate by minus one, such that an *increase* in the predictor of caseworker quality is associated with *shorter* average UI-spells. Thereby a caseworker of higher quality makes jobseekers exit unemployment faster. Finally, I standardize by the standard deviation on caseworker effects estimated in 6.1.

$$va_{j,-it} = \frac{-\bar{\mu}_{j,-it} \cdot \lambda_j}{\hat{\sigma}_\mu} \tag{7}$$

Henceforth, I will refer to the leave-one-out empirical Bayes estimate as caseworker value-added, $va_{j,-it}$. I use this in a set of regressions to investigate multi dimensionality and mechanisms. The regressions will include the same set of controls, T_{it}, as specified above:

$$Y_{it} = \alpha + \tau va_{j,-it} + \gamma_{k \times u \times t} + \delta_o + \kappa_q + \psi_a + \epsilon_{it} \tag{8}$$

Given the identifying assumption, τ represents the effect on some labor market outcome, Y_{it}, of assignment to a caseworker whose value-added on transitions is one standard deviation above the mean. Since jobseekers can have multiple UI-spells and caseworkers serve multiple jobseekers, I will cluster standard errors on caseworker and jobseeker level when estimating (8).

[43] Following Jackson (2018), I obtain the leave-out-mean by taking an average over all jobseeker residuals except jobseeker i in period t. The jobseeker residuals, e_{it}, are obtained by OLS regression of $Y_{it} = \gamma_{k \times u \times t} + \delta_o + \kappa_q + \psi_a + e_{it}$

7 Results

In this section, I present the empirical results. First, I examine the magnitude of the estimated caseworker effects, thereby addressing the question of whether and by how much heterogeneity in caseworker quality can explain the heterogeneity in transitions out of unemployment. I also construct an out-of-sample predictor for caseworker quality, validate its relevance and use it to investigate the effect of caseworker quality on exit probabilities at different points in time. Second, I present evidence on the dimensionality of caseworker quality. As a first examination, I correlate caseworkers effects on transitions with those on subsequent labor market performance. Hereafter, I investigate the effect of assignment to a caseworker, whose value-added on transitions is one standard deviation above the mean, on the jobseeker's subsequent labor market performance. Third, I present evidence on mechanisms.

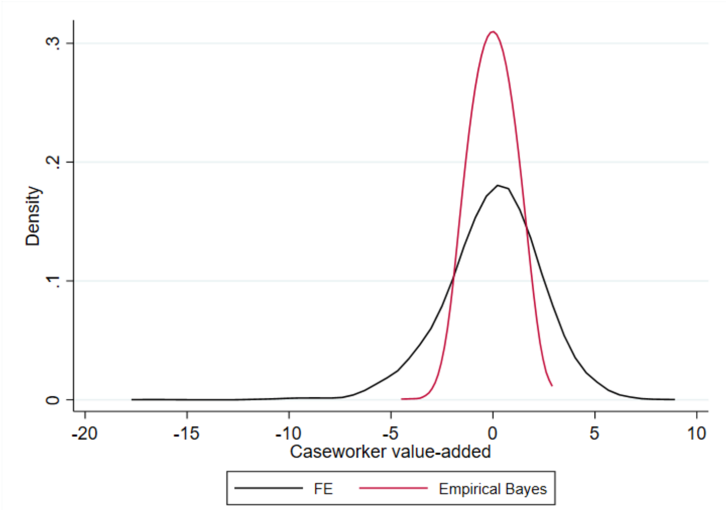
7.1 Magnitude of caseworker effects

Figure 3 shows the distribution of the estimated caseworker effects on unemployment spells. The black line represents the density of caseworker fixed effect (FE) estimates obtained by OLS regression of (4), while the red line represents the empirical Bayes (EB) estimates obtained by restricted maximum likelihood estimation of (6). The figure shows that the FE estimates have a much wider distribution compared to the EB estimates, especially due to very long tails. Table 4 shows that the standard deviation on the EB estimates (0.99 weeks) are about half the size of the standard deviation on the fixed effects (2.24 weeks). This corresponds to findings in the teacher value-added literature (see e.g. Rockoff, 2004), and it suggests that there is considerable noise in the FE's. Relying on the variance of the FE estimates would therefore cause one to over-estimate the magnitude of caseworker effects.

Although the standard deviation of the EB estimates are considerably smaller than that of the FE estimates, the former still suggests that caseworkers are important for jobseeker UI-spells. As already noted, table 4 shows that the standard deviation on the empirical Bayes estimates of caseworker value-added is 0.99 weeks. This means that jobseekers assigned to a caseworker, who is one standard deviation above the average, are expected to exit unemployment about one week earlier. Or equivalently; assignment to a caseworker at the 84^{th} vs. the 50^{th} percentile is expected to reduce the length of the UI-spell by about one week.

Column five of table 3 sets the standard deviation of the caseworker effects relative to the standard deviation of UI-spells within a jobseeker unit and year. It shows that variation in caseworker effects may explain 6% of the variation in UI-spells within a jobcenter unit and year. This is remarkable given that a rich set of jobseeker characteristics, including education, previous industry and labor market history, only is able to explain about 3% of the variation in UI-spells within jobcenter unit and year (see appendix table A4). It suggests that the quality of the assigned caseworker is an important contributing factor to the heterogeneity in transitions out of unemployment.

Figure 3: Magnitude of caseworker effects



Note: The figure plots the distribution of estimated caseworker effects. The black line represents the caseworker effects obtained by OLS estimation of (4). The red line represents the empirical Bayes estimates obtained by restricted maximum likelihood estimation of (6) where the caseworker effects are treated as random.

Table 4: Magnitude of caseworker effects

	Truncated UI-spells		Caseworker VA		Compare
	Mean	SD	SD	P-value	col 3 / col 2
FE	23.59	15.64	2.24	0.00	0.14
Empirical Bayes	23.59	15.64	0.99	0.00	0.06

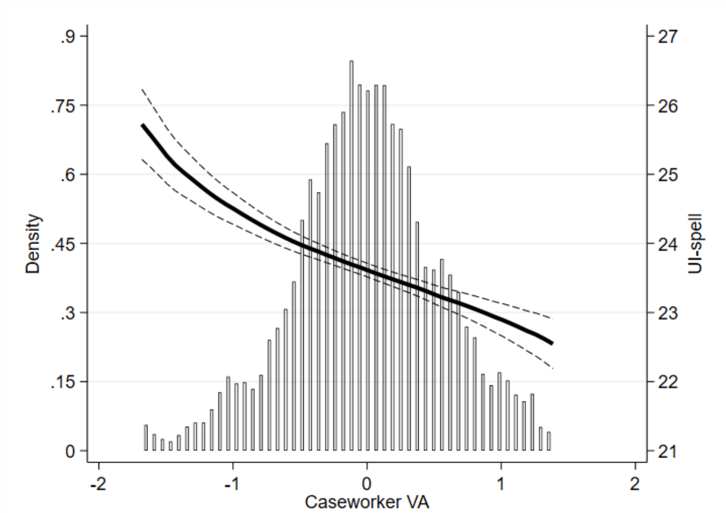
Note: Column 1 shows the unconditional mean for UI-spells truncated at 52 weeks, while column 2 shows the standard deviation within jobcenter unit and year and controlling for country of origin, quarter and age fixed effects. Column 3 shows the standard deviation on caseworker effects. The first row reports the standard deviation of the caseworker fixed effects opbtained by OLS estimation of (4). The second row reports the standard deviation directly obtained by restricted maximum likelihood estimation of (6) where the caseworker effects are treated as random. Column 4 shows the p-value on a test for joint significance of the caseworker effects (an F-test in row one, a likelihood ratio test in row two). Column 5 summarizes the magnitude of the caseworker effects by setting the SD of caseworker effects relative to the within-SD (col 3 / col 2).

Robustness: Using UI-spells truncated at 52 weeks to measure caseworker quality means that all caseworkers per definition are of similar quality after one year. One might be worried that I thereby lose an important nuance to caseworker quality. In the appendix, however, I show that I obtain similar results if I do not truncate UI-spells but instead reduce the sample period by one year. First, appendix figure A2 shows that caseworkers are ranked similarly when the empirical Bayes estimate of effectiveness are based on truncated UI-spells and when they are based on non-truncated UI-spells. Second, appendix table A8 shows that the variation in caseworker effectiveness can explain 5% of the variation in non-truncated UI-spells within a jobcenter unit and year. This is very close to the 6 % estimated for truncated UI-spells. It suggests that I do not lose important nuances by using truncated UI-spells.

7.2 Out-of-sample Predictions & Dynamics

I now investigate the effect of being assigned to a caseworker who (based on other jobseekers) is *predicted* to be one standard deviation above the mean. For this, I construct leave-one-out empirical Bayes estimates of caseworker effects on UI-spells (value-added, hereafter).

Figure 4: Caseworker value-added



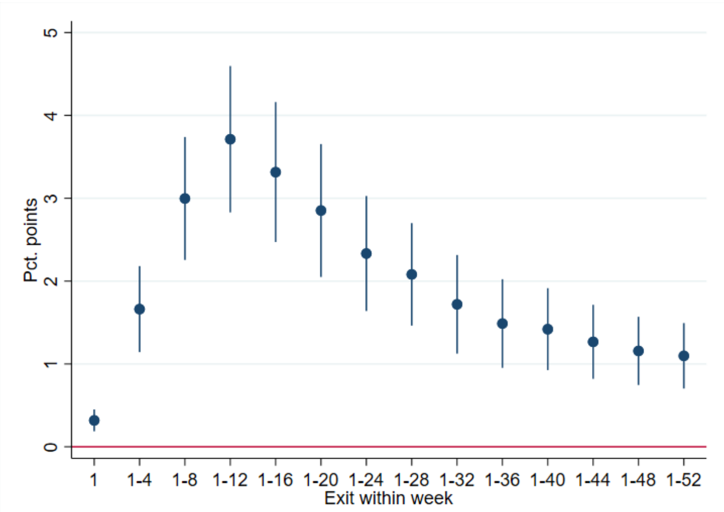
Note: The figure plots the length of the UI-spell against caseworker value-added (leave-out-one empirical Bayes estimates of caseworker effects). Plotted values are the residuals (with the unconditional mean added) from regressions on jobcenter × unit × year, non-western origin, age and quarter fixed effects. The line represents a local linear regression of UI-spell on caseworker value-added (degree 1, bandwidth 2). The dashed lines represent 95 percent confidence intervals. Bars represent the density of the caseworker value-added. Top and bottom 1 percent are excluded.

Figure 4 serves as a graphical representation of the impact of caseworker value-added on UI-spells. The black line (and 95% confidence intervals) represent a local linear regression of UI-spells on caseworker value-added. It clearly shows that UI-spells are monotonically decreasing in caseworker value-added, and further that the relationship is close to linear. Appendix table A9 reports the coefficient from a regression of UI-spells on caseworker value-added. It shows that assignment to a caseworker, who is *predicted* to be one SD above the mean, reduces the UI-spell by about one week. This estimate is highly significant (p<0.01), not statistically different from the effect presented in table 4, and shows that caseworker value-added is a relevant out-of sample predictor for the length of the average UI-spell.

Next, I investigate the impact of caseworker value-added on exit probabilities over time. Figure 5 shows the coefficients and 95% confidence intervals from regressions of a dummy for exiting unemployment within a given number of weeks relative to unemployment start on caseworker value-added and the usual set of controls, T_{it}. The figure shows that the impact of caseworker value-added kicks in early in the spell and that the accumulated effect increases over the first 12 weeks. Jobseekers assigned to a caseworker, who is predicted to be one standard deviation above the mean, are 3.7 pct. points more likely to exit unemployment within the first 12 weeks. Compared to the mean exit probability within the first 12 weeks (0.313), this corresponds to an effect of a about 12%. However, the figure also shows that the impact of caseworker value-added diminishes over the UI-spell, as the jobseekers assigned to caseworkers with lower value-added also eventually exit unemployment. I.e. the low value-added caseworkers eventually catch up with the high value-added caseworkers.

Another interesting point from figure 5 merits note. Namely, the figure shows that higher value-added caseworkers affect exit rates before the first meeting *on average* takes place. Namely, recall from figure 1b that the first meeting on average takes place in week 4 of the UI- spell. This might at first seem worrisome, but it actually hints at (one of) the mechanisms through which I find that caseworker quality works. Namely, in section 7.4, I show that higher value-added caseworkers tend to meet *earlier* with the jobseeker compared to low value-added caseworkers. On top of this, recall from the discussion in section 4.4, that much evidence shows that the mere invitation to a meeting increases the exit rate out of unemployment. I.e. the high value-added caseworkers might not only meet earlier but also *invite* the jobseeker earlier for meetings, which would induce more exits already before the meeting takes place.

Figure 5: Dynamics of caseworker value-added



Note: The figure plots the coefficient and 95 % confidence intervals from regressions of an exit dummy on caseworker value-added (leave-one-out empirical Bayes estimates). The exit dummy takes value one if the individual exits unemployment within a given number of weeks relative to unemployment start. All regressions include job-center × unit × year fixed effects, a control for country of origin, age and quarter fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. *p<0.10 ** p<0.05 *** p<0.01

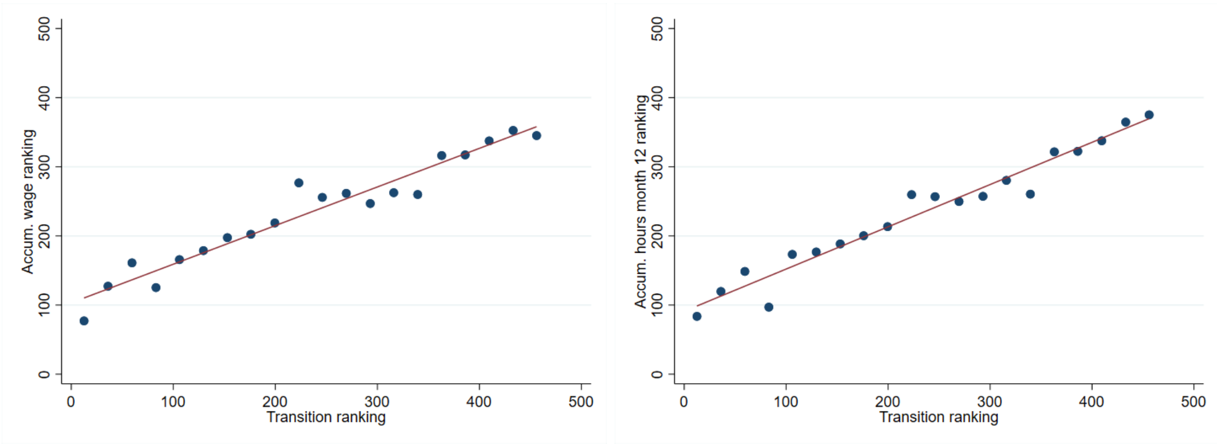
7.3 Dimensionality of caseworker quality

I now present evidence on the dimensionality of caseworker quality. In particular, I investigate whether assignment to a caseworker who ranks higher in terms of value-added on transitions out of unemployment adversely affect the jobseeker's subsequent *labor market performance*.

I will first consider two broad measures of labor market performance: Wage earnings and working hours accumulated over time relative to unemployment start. Hereafter I will investigate two channels through which accumulated wages and hours could be affected. The first channel is the individual's *labor market state* over time. Namely, the individual could leave unemployment to start employment, however, she could also start an education, receive other public transfers or drop out of the labor force. While the individual would accumulate wages and hours in the former state, she would not in the latter states. Further, since none of the states are absorbing, the individual could re-enter unemployment again over time, e.g. if the initial job match was poor. The second channel is the *job quality*. Here I will specifically consider monthly wages and hours conditional on employment. Finally, I end the section by investigating whether high value-added caseworkers only improve transitions for the mean jobseeker, potentially at the expense of other jobseekers.

Accumulated wages and hours: As a first exploration of whether caseworkers that improve transitions out of unemployment tend to do so at the expense of labor market performance, I correlate the ranking of caseworker effects on transitions as well as on accumulated hours and wages. I obtain the caseworker effects by estimation of (4), using UI-spells, accumulated hours and wages as the outcome. I then rank caseworkers according to their effect on each outcome. I code the ranking such that a caseworker ranks higher if she makes jobseekers transition faster out of unemployment, accumulate more wages and hours. Figure 6 plots the rank-rank correlations. The figure clearly shows a positive correlation between caseworker effects on transitions and wages as well as hours. Appendix table A10 reports a correlation coefficient of 0.56 between transitions and wages and 0.61 between transitions and hours (p<0.01). These positive associations suggest that caseworkers are not specializing in a single dimension. Rather, it suggests that good caseworkers improve both transitions *and* labor market performance.

Figure 6: Rank-rank correlations of caseworker effects



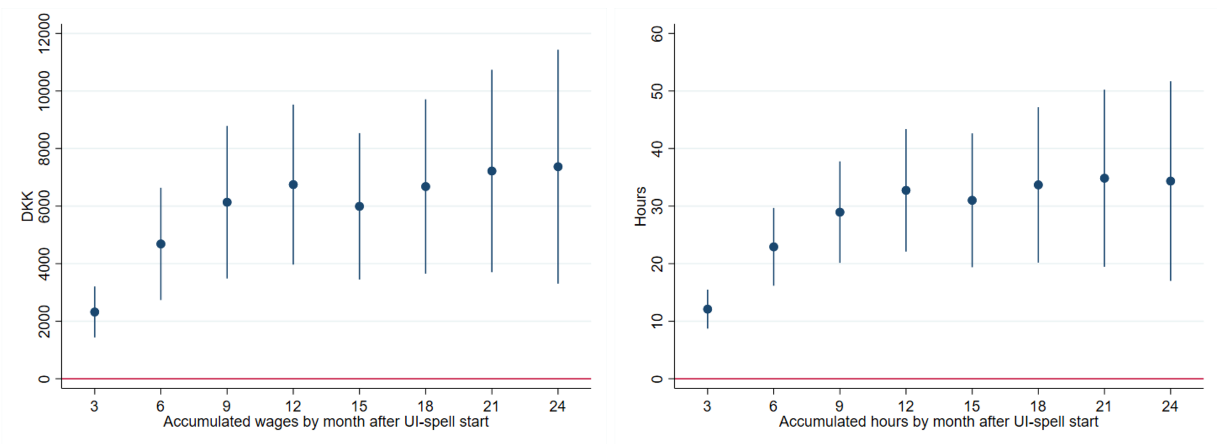
(a) Transitions vs. accumulated wages (b) Transitions vs. accumulated hours

Note: Each figure plots caseworker ranking in terms of transitions out of unemployment against caseworker ranking in terms of accumulated wages or hours by month 12. Note that caseworkers always are ranked within a jobcenter unit and year, controlling for country of origin, quarter and age fixed effects. *Transition ranking* is based on truncated UI-spells and coded such that a caseworker ranks higher if the mean jobseeker has a shorter UI-spell. *Wage (hours) ranking* is based on accumulated wage (hours) by month 12 relative to unemployment start. It is coded such that a caseworker ranks higher if the mean jobseeker has higher wages (hours).

Next, I use the leave-one-out empirical Bayes estimates to identify the causal effect of assignment to a caseworker, whose value-added on transitions is one standard deviation above the mean, on subsequent labor market performance. In particular, I regress accumulated wages and hours by a given month on caseworker value-added and the usual set of controls, T_{it} (see equation 8). Figure 7 shows the coefficients (and 95% confidence intervals) on caseworker value-added every third month after unemployment start. The figure shows that assignment to a caseworker with a higher value-added on transitions tend to *positively* affect the jobseeker's accumulated wages and hours. The effect is increasing over the first 12 months, after which it stabilizes. Two years after unemployment start, jobseekers assigned to a caseworker, whose value-added on transitions is one standard deviation above the mean, have accumulated about 35 additional working hours and 7,500 DKK more in wage earnings. Compared to the mean, this corresponds to 2% higher accumulated wages and hours after two years (see table A11).

Note that the effect on accumulated hours and wages corresponds to about *one* extra week of employment in the average full time job.[44] This is interesting in light of the finding that the high value-added caseworkers make jobseekers exit unemployment *one* week sooner. It suggests that high value-added caseworkers primarily make jobseekers exit unemployment for entries into employment, and to a lesser extent for entries into e.g. education, other transfers or to drop out of the labor force. Further, it suggests that the jobs in which the jobseeker becomes employed is not of lower quality. I investigate this further below.

Figure 7: Accumulated wages and hours since spell start



(a) Wages in DKK

(b) Hours

Note: Panel A (B) plots coefficients and 95% confidence intervals on caseworker value-added obtained by regressing accumulated wages (hours) by a given month relative to UI-spell start on caseworker value-added. Note that 100 DKK approximately corresponds to 16 USD. All regressions include jobcenter × unit × year fixed effects, a control for country of origin, age and quarter fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. *p<0.10 ** p<0.05 *** p<0.01.

[44]Considering individuals working full time (37.5 hours per week ∼ 158 hours per month) in a given month, they on average earned around 30,500 DKK per month.

Labor market states: I now construct five mutually exclusive states that an individual may belong to in a given week after unemployment start. In particular, the individual may receive UI-benefits[45], education subsidies, other transfers, be employed[46] or have dropped out of the labor force. I then evaluate the effect of caseworker value-added on the probability to belong to any of these states. Each panel in figure 8 shows the coefficients (and 95% confidence intervals) on caseworker value-added obtained from a regression of a dummy for belonging to a given state in a given week on caseworker value-added and the usual set of controls, T_{it}.

Comparing panel 8a and 8b it is evident that the effect of caseworker value-added on UI-benefits is close to a mirror image of the effect on employment. Assignment to a caseworker, whose value-added on transitions is one standard deviation above the average, *increases* the probability to be employed in a given week by almost as much as it *decreases* the probability to receive UI-benefits.[47] Further, the panels confirm the notion that jobseekers assigned to low value-added caseworkers eventually catches up with those assigned to high value-added caseworkers. Two years after UI-spell start, the positive effect on employment has disappeared as also the jobseekers assigned to low value-added caseworkers have found jobs.

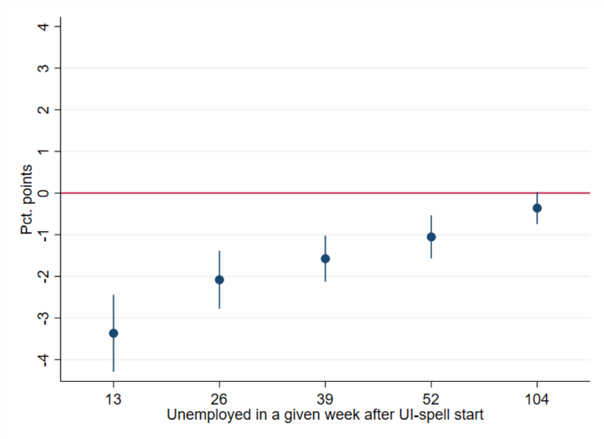
The latter three panels in figure 8 show that there is no or only tiny effects of caseworker value-added on the probability to enter other labor market states during the first two years. Panel 8c shows that caseworker value-added has *no* impact on the probability to receive education subsidies, which suggest that jobseekers are not leaving unemployment sooner to start an education. Likewise, panel 8e shows no real impact on the probability to drop out of the labor force. If anything, caseworker value-added has a tiny negative effect on the probability to drop out of the labor force. Panel 8d, shows that caseworker value-added in terms of transitions has a statistically significant, yet very small, effect on the probability to receive other transfers in the beginning. However, it should be noted that this effect is small, becomes insignificant from week 52 and has turned to a null in week 104. I.e. in the longer run, jobseekers assigned to high value-added caseworkers are no more likely to receive other types of transfers.

[45] This outcome is different from exit probabilities shown in figure 5. For an exit, I require that the individual does not receive UI-benefits in four consecutive weeks. I.e. the individual can have small breaks (up 3 weeks) from UI-benefits without 'exiting'.

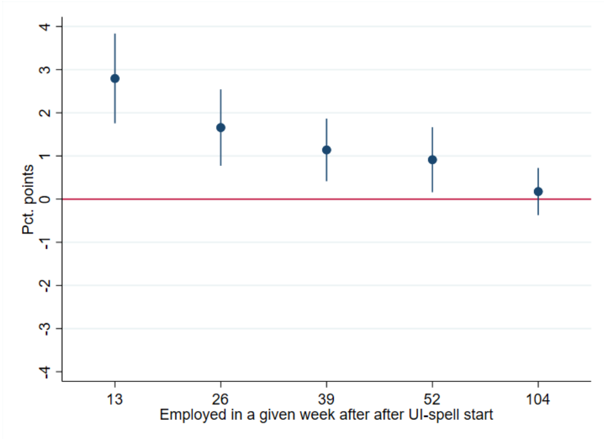
[46] Employment includes self-supported employment only (no receipt of transfers). I.e. if the individual is in some type of supported employment, she belongs to the 'other transfers' or 'UI-benefits' category.

[47] E.g. it decreases the probability to receive UI-benefits by about 3.4 pct. points in week 13 while increasing the probability to be employed by 2.8 pct. points in the same week.

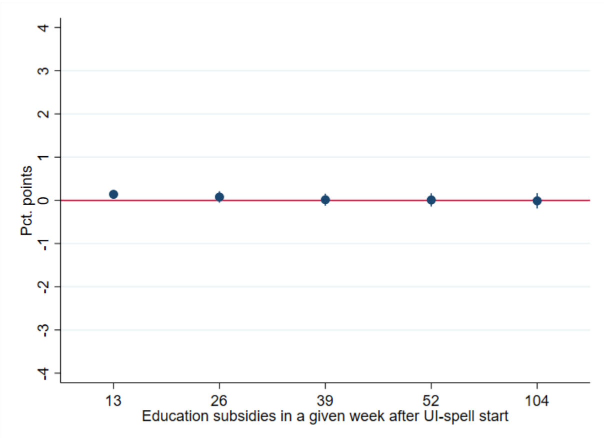
Figure 8: Labor market state in a given week after UI-spell start



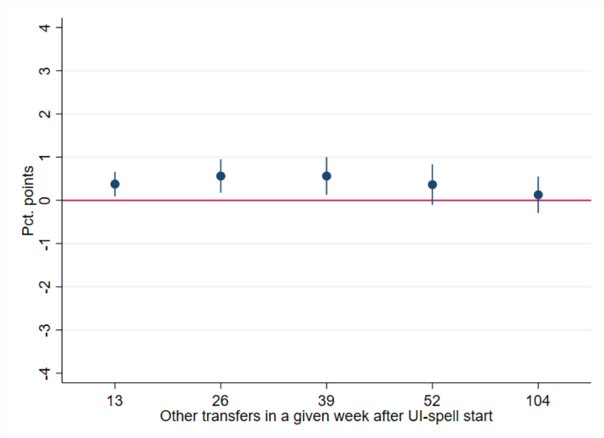
(a) UI-benefits



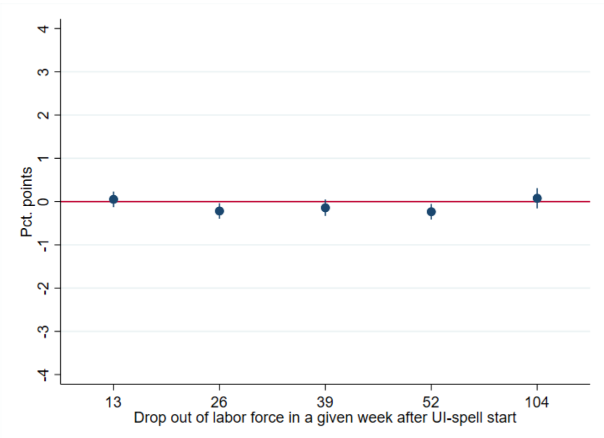
(b) Employed



(c) Education subsidies



(d) Other transfers



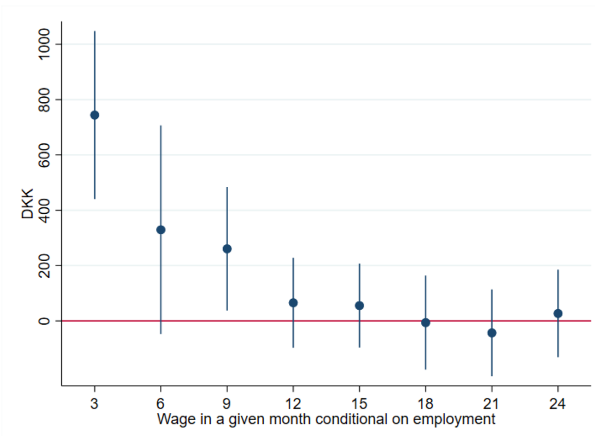
(e) Drop out of labor force

The individual is in a given week classified according to five mutually exclusive states: Recipient of UI-benefits, education subsidies, other public transfers, employed or having dropped out of the labor force (no transfers nor wage income). Each panel plots coefficients and 95% confidence intervals on caseworker value-added obtained by regressing a dummy for belonging to a given state in a given week after unemployment start on caseworker value-added. All regressions include jobcenter × unit × year fixed effects, a control for country of origin, age and quarter fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. *p<0.10 ** p<0.05 *** p<0.01.

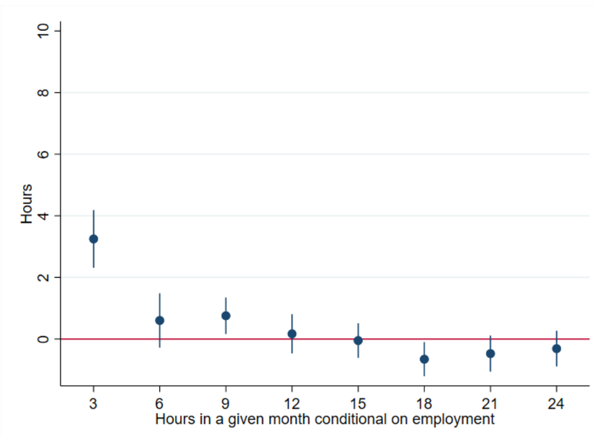
Overall, figure 8 shows that assignment to a caseworker, whose value-added in terms of transitions is one standard deviation above the average, causes the vast majority of jobseekers to exit unemployment for an entry to *employment*. Looking two years beyond unemployment start, the figure shows that sooner exits does not come at the expense of (re)entries to unemployment, other public transfers or entirely leaving the labor force. This could indicate that the high value-added caseworkers are not downgrading on job match quality, since one would expect the job to end sooner if the initial employer-employee match is very poor.[48]

Job quality: Next, I investigate whether caseworker value-added affects job quality. Namely, one could speculate that some caseworkers have higher value-added because they trade off fast exits against job quality, e.g. in terms of wages or hours. To investigate this, I regress monthly wages and hours *conditional* on employment, on caseworker value-added and the usual set of controls, T_{it}. Figure 9 shows the coefficients (and 95% confidence intervals) on caseworker value-added. By conditioning on employment, I introduce a bias and the estimates should therefore be interpreted with caution (see discussion below). Nevertheless, taking the estimates at face value, the figure shows *no* evidence that assignment to a caseworker with higher value-added on transitions negatively affects job quality. If anything, the figure shows a small positive effect on wages and hours in month 3, which quickly hereafter turns into null effects.

Figure 9: Wages and hours conditional on employment in a given month



(a) Wages in DKK conditional on employment



(b) Hours conditional on employment

Note: Panel A (B) plots coefficients and 95% confidence intervals on caseworker value-added obtained by regressing wages (hours) in a given month after UI-spell start, *conditional* on employment in that month, on caseworker value-added. Note that 100 DKK approximately corresponds to 16 USD. All regressions include jobcenter × unit × year fixed effects, a control for country of origin, age and quarter fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. *p<0.10 ** p<0.05 *** p<0.01.

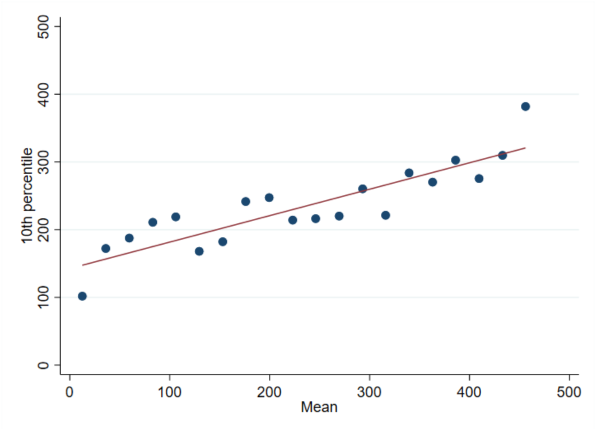
[48] Yet, I cannot rule out that the jobseekers are making job-to-job transitions due to an poor job match quality.

I will argue that there is reason to believe that figure 9 presents *lower* bounds on the true causal effect of caseworker value-added. For the line of reasoning, it is instructive to think about the marginal jobseeker. That is, the additional jobseeker, who the high value-added caseworker is able to move out of unemployment and into employment, but who the low value-added caseworker was not able to move. One could imagine that both high and low value-added caseworkers move high ability (high employment and earnings potential) jobseekers out of unemployment. In addition, the high value-added caseworker might move some jobseekers of lower ability into employment (the marginal jobseeker). Given that ability and wages are correlated, this would tend to generate a negative bias of the coefficient on caseworker value-added in regressions with conditional wages on the right hand side. Hence, if the marginal jobseekers are of lower ability, the estimates in figure 9 might represent a *lower* bound on the true causal effect on job quality. That suggests that assignment to high value-added caseworkers does *not* negatively affect job quality measured in terms of wages and hours.

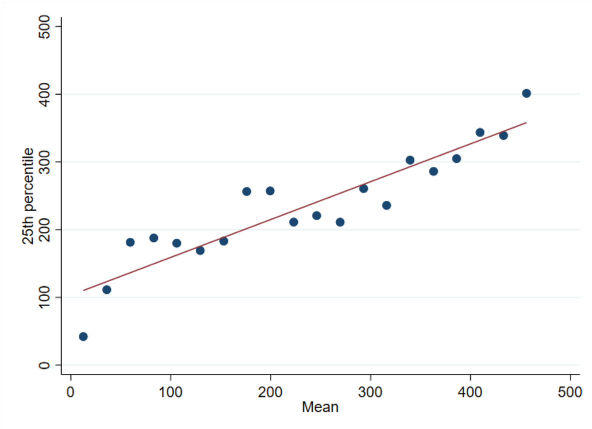
Welfare and policy implications: The analysis above suggests that assignment to a high value-added caseworker does *not* adversely affect later labor market performance. In fact, jobseekers assigned to a caseworker, whose value-added in terms of transitions is one standard deviation above the average, have on average accumulated an additional 7,500 DKK and 35 hours after two years. This reflects that higher value-added caseworkers push the vast majority of jobseekers from unemployment and into employment. After two years, there is no evidence that sooner exits from unemployment causes the average jobseeker to (re)enter unemployment, other transfers or to drop out of the labor force. Nor do I find evidence suggesting that they are employed in jobs of lower quality in terms of wages and hours.

The above analysis has focused entirely on caseworker impacts on the *mean* jobseeker. It suggests that caseworkers who improve the transitions out of unemployment for the mean jobseeker also improves subsequent labor market performance. Given a utilitarian welfare function, the policy implication seems obvious: We should teach the low value-added caseworkers the tricks of the high value-added caseworkers and thereby improve outcomes for the mean jobseeker. However, it seems relevant to ask whether the policy recommendation is the same in case of other welfare weights? E.g. Rawlsian welfare weights, where we care about the least off rather than the mean jobseeker. To investigate this, I correlate the caseworker's ranking in terms of the *mean* UI-spell with her ranking in terms of a given *percentile* of the UI-spell distribution. I code caseworker rankings such that a caseworker ranks higher if the UI-spell of the jobseeker is shorter. The rank-rank correlations are shown in figure 10. Clearly there is a positive correlation between the caseworker's ranking on the mean jobseeker and jobseekers at the 10th, 25th, 75th and 90th percentile of the distribution (see exact correlations in appendix table A12). This suggests that high value-added caseworkers improve transitions for the entire distribution of jobseekers. This again suggests that caseworker quality is one dimensional, and that it would be welfare improving to teach low value-added caseworkers what the high value-added caseworkers do. I investigate this in the next section.

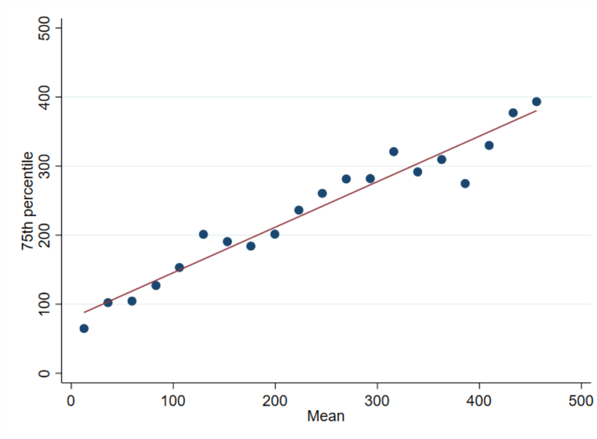
Figure 10: Rank-rank correlations of caseworker impacts
Mean vs. some jobseeker percentile



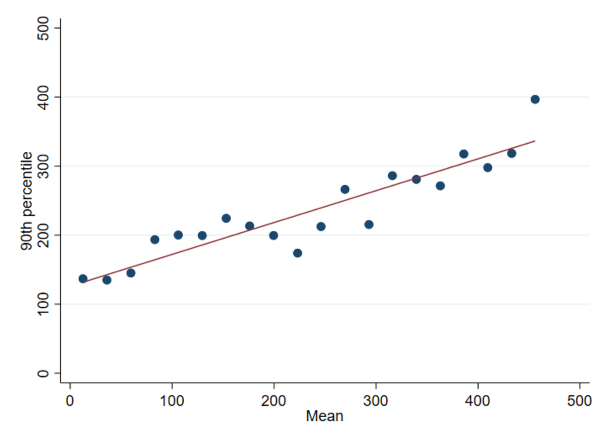
(a) 10th percentile



(b) 25th percentile



(c) 75th percentile



(d) 90th percentile

Note: Each figure plots caseworker ranking on the *mean* jobseeker's UI-spell against caseworker ranking on the UI-spell of some *percentile* of the jobseeker distribution. Three details of the caseworker rankings merit note. First, caseworkers always ranked within a jobcenter unit and year, controlling for country of origin, quarter and age fixed effects. Second, ranking is coded such that a caseworker ranks higher if the jobseeker of interest (the mean or a given percentile) has a shorter UI-spell. Third, to comply with data protection rules, all percentiles are based on at least 5 observations.

7.4 Mechanisms

In this section, I investigate what higher value-added caseworkers do differently. I will examine three overall mechanisms through which caseworker value-added could affect labor market outcomes: Meetings, training and job search.

For each mechanism, I will define a set of *behaviors* that could explain why some caseworkers make jobseekers transition faster out of unemployment. Specifically, I will regress a given behavior on caseworker value-added and the usual set of controls, T_{it}. One challenge when interpreting the coefficient on caseworker value-added, however, is that many behaviors change over the UI-spell. Since caseworker value-added is defined in terms of UI-spell length, this creates a mechanical correlation between value-added and the behavior studied. For example, one behavior is the number of job search methods used by the jobseeker in the average week of the UI-spell. As shown in appendix figure A6, jobseekers tend to decrease the number of job search methods over the UI-spell. This creates a positive correlation between caseworker value-added and the number of search methods.[49] This in turn suggests that the estimated coefficient on caseworker value-added will be the composite of the true causal effect on search methods and a positive bias term. Depending on the sign of the estimate, one might still be able to draw conclusions (rule out certain effects). However, it also means that one should be cautious when interpreting the estimates. For this reason, I will examine the dynamics of all behaviors used to study the three mechanisms.

Meetings: I first investigate whether higher value-added caseworkers have different meeting strategies, e.g. in terms of the frequency or timing of meetings. There is much evidence suggesting that earlier and more frequent caseworker meetings positively affect transitions out of unemployment (Graversen and Van Ours, 2008; Maibom et al., 2017; Hägglund, 2011; Van den Berg et al., 2012; Rosholm, 2014). Hence, the differences in value-added could potentially be explained by differences in frequency and timing of meetings across caseworkers.

To investigate this, I calculate the number of weeks between the jobseeker's UI-spell start and first meeting, and the number of meetings per week.[50] Appendix figure A3 shows that the number of meetings per week is fairly constant over the UI-spell. Hence, I regress both out-

[49] Recall that value-added is defined such that higher value-added caseworkers reduce UI-spell length.

[50] Since meeting requirements change after 26 weeks of unemployment I consider only the first 26 weeks.

comes on caseworker value-added and the usual set of controls, T_{it} (see equation 8), and expect no bias. Table 5 reports the coefficient on caseworker value-added from these regressions. The first column shows that jobseekers assigned to a caseworker, whose value-added is one standard deviation above the mean, have their first meeting about 0.4 weeks sooner. Relative to the mean, this is an effect of -8%. The second column shows that these jobseekers have 0.010 more meetings per week (+6% relative to the mean). The estimates are significant at the 1% level, and suggest that better caseworkers meet not only earlier but also more frequently with the jobseeker. The fact that they meet earlier - and likely also send out the meeting invitation earlier - could explain why caseworker value-added improve exit rates already before jobseekers *on average* has their first meeting. I.e. this reconciles the finding in figure 1b and 5.[51]

Table 5: Meetings

	Weeks until 1st meeting	Meetings/week
Caseworker VA	-0.370***	0.010***
	(0.038)	(0.002)
Expected bias	0	0
Obs	103027	103027
Dep. var mean	4.476	0.171
Dep. var SD	3.318	0.120

Note: Each column reports the coefficient on caseworker value-added (leave-one-out empirical Bayes estimates). All regressions include fully interacted jobcenter unit and year fixed effects, a control for country of origin, quarter and age fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. *Weeks until first meeting* refers to the number of weeks between the UI-spell start and the first meeting in the jobcenter. *Meetings per week* refers to the number of caseworker meetings pr. week of unemployment. If the UI-spell is longer than 26 weeks, only the first 26 weeks are considered, because meeting frequency requirements change after 26 weeks. *p<0.10 ** p<0.05 *** p<0.01

One might worry that the effect on the meeting frequency is driven by the fact that jobseekers since 2016 have been required to book later meetings themselves. I.e. could it reflect that jobseekers book the meetings more frequently when assigned to a 'better' caseworker? Here, it is important to note that the jobseeker may affect the timing but not the frequency of later meetings, which is dictated by the caseworker. As such, reversed causality is unlikely. Another concern is that a labor market reform in 2015 changed the timing and frequency of meetings. Appendix table A13 addresses these two concerns by estimating the effect on meeting timing and frequency before and after 2015. The table shows that the effects are similar. This suggests

[51]Figure 1b shows that the first meeting *on average* takes place in week 4. Figure 5 shows that caseworker value-added improve exit rates earlier than week 4. If higher value-added caseworkers have their first meeting earlier and send out the invitation for this meeting earlier, it would reconcile these two figures.

that the effects on meeting frequency and timing are neither driven by the reform nor by self-booking of meetings.

Training: I now investigate whether higher value-added caseworkers are different in terms of training assignment strategies. First, it is instructive to discuss how training assignments potentially could explain the results found in this paper.

Participation in training could improve the skills of the jobseeker and thereby positively affect the probability to exit unemployment. However, such positive effects are expected to materialize in the longer run. In their meta analysis, Card et al. (2010) show that *participation* in classroom training and on-the-job training generally does *not* increase the exit rate in the short run. This is mainly due to locking-in effects; participants in training have less time to search for jobs compared to non-participants, which tends to prolong time in unemployment (Jespersen et al., 2008; Humlum and Rasmussen, 2021). Therefore, it is not clear that *participation* in training can explain why some caseworkers make jobseekers exit *faster* from unemployment. However, it is important to note that assignment to training not necessarily implies participation. There is much evidence suggesting that training *assignments* are associated with so-called threat effects; jobseekers exit unemployment merely to avoid participation in training activities (Geerdsen, 2006; Geerdsen and Holm, 2007; Rosholm and Svarer, 2008). If higher value-added caseworkers are more likely to assign jobseekers to training or just assign them earlier, such threat effects could explain why their jobseekers exit unemployment faster. Finally, it should be noted that since training are associated with locking-in effects, it could also be the other way around: Some caseworkers could have lower value-added because they to a large extent make jobseekers participate in training and thereby 'lock' them into unemployment.

Based on individual job plans, I can see whether the caseworker ever assigns the individual to a given activity, broadly defined as courses or job training, and when the activity is set to start. Appendix figure A4 shows that the probability to be assigned to any activity is increasing in UI-spell length. This implies that caseworker value-added and the probability to be assigned to training is negatively correlated. Hence, when I regress a dummy for assignment to training on caseworker value-added and the usual set of controls, T_{it}, the estimated coefficient will be the composite of the true causal effect and a negative bias term. Panel A of table 6 reports the estimated coefficients on caseworker value-added. Since the estimates are negative, and the bias term is expected to be negative, the coefficients in panel A do not allow me to draw conclusions.

To address the trend in training assignment probabilities, I now condition on UI-spells of at least 8 weeks and define a dummy taking value one if the jobseeker is assigned to a given activity that starts during those 8 weeks.[52] I regress this dummy on caseworker value-added and the usual set of controls, T_{it}. Panel B of table 6 reports the coefficients on caseworker-value added. Interestingly, the estimated coefficients change from negative in panel A to positive in panel B. The table shows a positive effect of caseworker value-added on the probability to be assigned to job training (p<0.001), yet small and insignificant effects on the probability to be assigned to course training. Although the estimates should be interpreted with caution, they suggest that jobseekers assigned to caseworkers, whose value-added is one standard deviation above the mean, are more likely to be assigned to job training.

Table 6: Training assignment conditional on spell length

	Ever assigned		
	Any training	Course training	Job training
A: Unconditional			
caseworker VA	-0.017***	-0.018***	-0.001
	(0.006)	(0.006)	(0.003)
B: Spells $\geq$ 8 weeks			
Caseworker VA	0.008*	0.004	0.005***
	(0.004)	(0.004)	(0.001)
Expected bias	-	-	-
Obs (uncond.)	103027	103027	103027
Obs (spell>8)	87,678	87,678	87,678
Dep. var mean (uncond.)	0.435	0.340	0.217
Dep var mean (spell>8)	0.127	0.100	0.032

Note: Each column reports the coefficient on caseworker value-added (leave-one-out empirical Bayes estimates). All regressions include fully interacted jobcenter unit and year fixed effects, a control for country of origin, quarter and age fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. In panel A, all spells are included in the estimation. The dependent variable is a dummy taking value one is the individual is ever assigned. In panel B, only spells of at least 8 weeks are included. The dependent variable is a dummy taking value one if the individual was assigned to an activity that started during the first 8 weeks of the UI-spell. *Expected bias* refers to the sign of the expected bias of the unconditional estimate (panel A). This is based on the correlation between training and UI-spell length (see appendix figure A4). *p<0.10 ** p<0.05 *** p<0.01

To investigate whether higher value-added caseworkers might assign jobseekers *earlier* to training, I calculate the number of weeks between UI-spell start and the start of the assigned training activity (hence, this outcome is conditional on being assigned to training). I regress this on caseworker value-added and the usual set of controls, T_{it}. Table 7 reports the coefficients on

[52]It should be noted that this is not a perfect solution, as it may not be random which jobseekers that are left after a given number of weeks. For this reason, and for the sake of power, I choose a relatively narrow window.

caseworker value-added. The estimates suggest that jobseekers assigned to a caseworker, whose value-added is one standard deviation above the mean, are assigned to training activities that start close to one week *earlier* (0.88 weeks earlier for courses, 1.25 for job training). This is interesting, since the high value-added caseworkers also push jobseekers one week earlier out of unemployment.

Table 7: Training timing

	Weeks until training start		
	Any	Course	Job training
Caseworker VA	-0.887***	-0.881***	-1.248***
	(0.222)	(0.272)	(0.243)
Obs	44820	35025	22330
Dep. var mean	15.964	17.040	23.894
Dep. var SD	12.710	14.874	18.392

Note: The figure reports the coefficient on caseworker value-added (leave-one-out empirical Bayes estimates). All regressions including fully interacted jobcenter unit and year fixed effects, a control for country of origin, quarter and age fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. . Training assignments is based on individual job plans. *Weeks until a given activity* is the number of weeks between UI-spell start and the start of the activity (job training, course training or any of the two) registered in the job plan. *p<0.10 ** p<0.05 *** p<0.01

Job search: I finally investigate whether higher value-added caseworkers affect the individual's job search behavior. To investigate this, I rely on joblog data which contain information on applied-for jobs for individual jobseekers since 2015. In particular, I investigate the effect of caseworker value-added on registered job search in the *average* week of the UI-spell[53].

First, one might hypothesize that some caseworkers have higher value-added on transitions, because they make jobseekers increase the job search *intensity*. As a crude proxy for job search intensity, I consider the *number of applied-for jobs* registered in the average week (joblogs, hereafter). Appendix table B13 shows that the number of weekly joblogs is fairly constant over the UI-spell.[54] Therefore, I regress the number of joblogs in the average week on caseworker value-added and the usual set of controls, T_{it}, and expect no bias. The first column of table 8 reports the coefficient on value-added. Surprisingly, it suggests that jobseekers assigned to caseworkers, whose value-added is one standard deviation above the mean, register 0.04 *fewer* joblogs in the average week (2.5% relative to the mean). Hence, high value added caseworkers

[53] I consider the first 26 weeks of the UI-spell

[54] Disregarding the first and last two weeks of the UI-spell. Hence, I will also disregard these weeks when evaluating the number of joblogs in the average week

do not make jobseekers *register* more applied-for jobs. As discussed in section 4.4, this does not rule out that they increase the jobseeker's *actual* number of applied-for jobs. Further, the fact that jobseekers assigned to higher value-added caseworkers tend to register *fewer* jobs could suggest that these caseworkers put less weight on monitoring.

Second, the caseworker could affect the *number of search methods and channels* that the jobseeker makes use of. I interpret this as another proxy for search intensity. To investigate this, I rely on the fact that jobseekers in the joblog must register the search method (e.g. 'applied via e-mail') and the search channel (e.g. 'using my network'). Appendix figure A6 and A7 show that the number of search methods and channels per joblog is decreasing over the UI-spell. This mechanically creates a positive correlation between caseworker value-added and search methods (channels). Being aware of a potential positive bias, I now count the number distinct search methods (channels) per joblog in the average week and regress this on caseworker value-added and the usual set of controls, T_{it}. The coefficients on caseworker value-added is shown in column 2 (3) of panel A in table 8. Interestingly, the table shows that both coefficients are estimated to be zero. Given that the bias is expected to be positive, this suggests that I can rule out that higher value-added caseworkers increase the number of search methods and channels. This conclusion is further supported by panel B, where I condition on UI-spells of at least 8 weeks. Hence, relying on joblogs, I find no evidence that higher value-added caseworkers increase job search intensity.

Third, some caseworkers could have higher value-added on transitions because they make jobseekers use *specific search channels*. E.g. they could make jobseekers send more unsolicited applications or make more use of their network, which has been shown to be important for employment probabilities (Glitz, 2017; Saygin et al., 2021). To investigate this, I consider the fraction of applied-for jobs found through the jobseeker's network and the fraction of applied-for jobs that were sent unsolicited. Appendix figure A8 and A9 show that the probability to make use of these search channels is decreasing over the UI-spell. This mechanically creates a positive correlation between caseworker value-added and the search channels. Being aware of a potential positive bias, I now regress the two search channels on caseworker value-added along with the usual set of controls, T_{it}. The coefficients on caseworker value-added are reported in column 4 and 5 of panel A in table 8. The estimated coefficients on caseworker value-added are positive. However, as these coefficients presumably are affected by a positive bias term, I cannot conclude that higher value-added caseworkers make jobseekers increase the use of their network and unsolicited search solely based on panel A. In panel B of the table, I therefore condition on UI-spells of at least 8 weeks and consider the use of network and unsolicited search during these first 8 weeks. Interestingly, the effects on network and unsolicited search remain positive and do not change much compared to panel A (yet, the effect on network becomes insignificant). Compared to the mean, the estimates suggest that assignment to a caseworker, who is one standard deviation above the average, increases the use of network and unsolicited search by about 5% in the average week.

Finally, some caseworkers could have higher value-added on transitions because they make jobseekers apply for jobs that are less attractive but potentially easier to get (a reservation wage story, see McCall, 1970). To investigate this, I consider the fraction of applied-for jobs that were part time and the average commute distance of the jobs. Appendix figure A10 and A11 show that the probability of applying for part time jobs and jobs with a longer commute is increasing over the UI-spell. I.e. jobseekers tend to become more willing to apply for less attractive jobs. This implies that caseworker value-added and both job type measures will be negatively correlated. Being aware of a negative bias, I now regress both job type measures on caseworker value-added and the usual set of controls, T_{it}. The coefficients on caseworker value-added are reported in column 6 and 7 in panel A of table 8. Since the expected bias is negative, and because the estimated coefficients in panel A are zero and negative, it means that I cannot conclude whether higher value-added caseworkers do in fact increase the probability to apply for part time jobs and jobs with a longer commute. In panel B of the table, I again condition on UI-spells of at least 8 weeks and obtain similar coefficients as in panel A. This suggests that higher value-added caseworkers do *not* induce the jobseeker to apply more for part time jobs or jobs with a longer commute. It suggests that the higher value-added caseworkers do not make jobseekers less 'picky' in the jobs they apply to.

Table 8: Joblog behavior

	Intensity			Specific channels		Job type	
	Joblogs/week	Methods/log	Channels/log	Network	Unsolicited	Part time	Commute (km)
A: Unconditional							
Caseworker VA	-0.043***	-0.002	0.000	0.004**	0.011**	-0.001	-0.889***
	(0.014)	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)	(0.317)
B: Spells $\geq$ 8 weeks							
Caseworker VA	-0.043***	-0.000	0.000	0.003	0.011**	-0.002	-0.962***
	(0.015)	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)	(0.326)
Expected bias sign	0	+	+	+	+	-	-
Obs (uncon.)	39,621	38,694	38,694	38,694	38,694	38,694	21,154
Obs (spell>8)	35,875	34,648	34,648	34,648	34,648	34,648	15,435
Dep. var mean (uncon.)	1.707	0.644	0.592	0.061	0.219	0.139	30.762
Dep var mean (spell>8)	1.733	.651	.601	.066	.222	.135	30.382

Note: Each column reports the coefficient on caseworker value-added (leave-one-out empirical Bayes estimates). All regressions include fully interacted jobcenter unit and year fixed effects, a control for country of origin, quarter and age fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. While panel A does *not* condition on spell length, panel B conditions on spells of at least 8 weeks and show only behavior during these 8 weeks. All dependent variables in the table is based on Joblog data and represent joblog-behavior in the average week (not considering behavior in the first and last two weeks of the UI-spell, see Fluchtmann et al. (2020)). *Joblogs per week* refers to the number of joblogs per week *Search methods* refers to the number of search methods per joblog (examples of search methods are: 'via e-mail', 'letter', 'phone', 'in person'). *Search channels* refers to the number of search channels per joblog (examples of search channels are: 'unsolicited', advertised, 'through network'). *Network* refers to the fraction of applied-for jobs that was found through the network. *Unsolicited* refers to the fraction of applied-for jobs that were unsolicited. *Part time job* refers to the fraction of applied-for jobs that were part time. *Commute distance* refers to the average commute distance (in kilometers) of applied-for jobs. This latter variable is the only information that was not mandatory to register in the joblog, and hence this information is not available for all joblogs (hence, the fewer observations in this column). *p<0.10 ** p<0.05 *** p<0.01

Overall, the analysis suggests that variation in caseworker quality is not merely driven by differences in personality traits and unobserved ability. Rather, the evidence put forward suggests that better caseworkers tend to be more pro-active; they meet earlier, more often and tend to assign jobseekers earlier to training. Using registered job search as a proxy for actual job search, I find no evidence suggesting that better caseworkers make jobseekers search more intensively or apply for less attractive jobs. However, I find evidence suggesting that better caseworkers make jobseekers increase the use of their network and unsolicited applications.

8 Conclusion

In this paper, I investigate the importance of caseworker quality for jobseeker transitions out of unemployment. To draw policy recommendations, I further provide evidence on the dimensionality of caseworker quality and the mechanisms through which caseworker quality could be affecting labor market outcomes.

I find that variation in caseworker quality can explain around 6% of the heterogeneity in transitions out of unemployment for jobseekers initiating a UI-spell in the same jobcenter and year. This is a remarkable impact, since various jobseeker characteristics only can explain around 3% of the heterogeneity in transitions. Assignment to a caseworker, who is one standard deviation above the mean, reduces the average UI-spell by about *one week*. Investigating the dynamics behind this effect, I find that caseworker quality have large effects on the exit probability in the beginning of the UI-spell. However, the effect diminishes over time as jobseekers assigned to lower quality caseworkers also eventually exit unemployment.

I find *no* evidence suggesting that caseworkers, who improve transitions out of unemployment, adversely affect the subsequent labor market performance of jobseekers. In fact, two years after unemployment start, jobseekers assigned to a caseworker, whose value-added on transitions is one standard deviation above the mean, have on average accumulated about 7,500 DKK and 35 hours more than those assigned to the mean caseworker. This effect reflects that caseworkers with higher value-added tend to push jobseekers from unemployment and into employment. I find no evidence suggesting that sooner exits from unemployment come at the expense of subsequent job quality. Jobseekers are not employed at lower wages and hours, nor are they more likely to re-enter unemployment, e.g. due to poor job matches. Finally, I provide evidence suggesting that better caseworkers not only improve transitions for the mean jobseeker but for the entire distribution of jobseekers. All together, this suggests that caseworker quality is one dimensional, and importantly, it suggests that there is no reason why all caseworkers shouldn't adopt the strategies of the high value-added caseworkers.

In a final step, I investigate what the high value-added caseworkers are doing differently. I find that the high value-added caseworkers tend to be more *pro-active*: They meet earlier and more frequently with the jobseeker, and they assign the jobseeker earlier to training. Using registered job search as a proxy for job search behavior, I find no evidence suggesting that higher value-added caseworkers induce jobseekers to search more intensively. However, I find evidence suggesting that they make jobseekers increase the use of their network and unsolicited search. Although this may not explain the entire variation in caseworker value-added, it is an important insight for policy makers. It suggests that simple interventions potentially could reduce time spent in unemployment.

References

Behncke, S., Frölich, M. and Lechner, M. (2009), 'Targeting labour market programmes—results from a randomized experiment', *Swiss Journal of Economics and Statistics* **145**(3), 221–268.

Behncke, S., Frölich, M. and Lechner, M. (2010*a*), 'A caseworker like me–does the similarity between the unemployed and their caseworkers increase job placements?', *The Economic Journal* **120**(549), 1430–1459.

Behncke, S., Frölich, M. and Lechner, M. (2010*b*), 'Unemployed and their caseworkers: should they be friends or foes?', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **173**(1), 67–92.

Bennedsen, M., Pérez-González, F. and Wolfenzon, D. (2020), 'Do ceos matter? evidence from hospitalization events', *The Journal of Finance* **75**(4), 1877–1911.

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

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.

Bolhaar, J., Ketel, N. and van der Klaauw, B. (2020), 'Caseworker's discretion and the effectiveness of welfare-to-work programs', *Journal of Public Economics* **183**, 104080.

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. (2020), 'What makes a good caseworker?', *mimeo* .

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014*a*), 'Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates', *American Economic Review* **104**(9), 2593–2632.

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014*b*), 'Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood', *American Economic Review* **104**(9), 2633–79.

Fluchtmann, J., Glenny, A., Harmon, N. and Maibom, J. (2020), 'The dynamics of job search in unemployment: Beyond search effort and reservation wages', *mimeo* .

Geerdsen, L. P. (2006), 'Is there a threat effect of labour market programmes? a study of almp in the danish ui system', *The Economic Journal* **116**(513), 738–750.

Geerdsen, L. P. and Holm, A. (2007), 'Duration of ui periods and the perceived threat effect from labour market programmes', *Labour Economics* **14**(3), 639–652.

Glitz, A. (2017), 'Coworker networks in the labour market', *Labour Economics* **44**, 218–230.

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**(10-11), 2020–2035.

Hägglund, P. (2011), ‘Are there pre-programme effects of active placement efforts? evidence from a social experiment’, *Economics Letters* **112**(1), 91–93.

Huber, M., Lechner, M. and Mellace, G. (2017), ‘Why do tougher caseworkers increase employment? the role of program assignment as a causal mechanism’, *The Review of Economics and Statistics* **99**(1), 180–183.

Humlum, A. and Rasmussen, M. (2021), ‘Is vocational training of workers a bad investment?’, *mimeo* .

Jackson, C. (2018), ‘What do test scores miss? the importance of teacher effects on non–test score outcomes’, *Journal of Political Economy* **126**(5), 2072–2107.

Jackson, C. K., Rockoff, J. E. and Staiger, D. O. (2014), ‘Teacher effects and teacher-related policies’, *Annual Review of Economics* **6**(1), 801–825.

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.

Kane, T. J., McCaffrey, D. F., Miller, T. and Staiger, D. O. (2013), Have we identified effective teachers? validating measures of effective teaching using random assignment, *in* ‘Research Paper. MET Project. Bill & Melinda Gates Foundation’, Citeseer.

Kane, T. J. and Staiger, D. O. (2008), Estimating teacher impacts on student achievement: An experimental evaluation, Technical report, National Bureau of Economic Research.

Lazear, E. P., Shaw, K. L. and Stanton, C. T. (2015), ‘The value of bosses’, *Journal of Labor Economics* **33**(4), 823–861.

Maibom, J., Rosholm, M. and Svarer, M. (2017), ‘Experimental evidence on the effects of early meetings and activation’, *The Scandinavian Journal of Economics* **119**(3), 541–570.

McCall, J. J. (1970), ‘Economics of information and job search’, *The Quarterly Journal of Economics* **84**(2), 113 – 126.

Mulhern, C. (2019), ‘Beyond teachers: Estimating individual guidance counselors’ effects on educational attainment’.

Rockoff, J. E. (2004), ‘The impact of individual teachers on student achievement: Evidence from panel data’, *American economic review* **94**(2), 247–252.

Rosholm, M. (2014), ‘Do case workers help the unemployed?’, *IZA World of Labor* .

Rosholm, M. and Svarer, M. (2008), ‘The threat effect of active labour market programmes’, *scandinavian Journal of Economics* **110**(2), 385–401.

Saygin, P. O., Weber, A. and Weynandt, M. A. (2021), ‘Coworkers, networks, and job-search outcomes among displaced workers’, *ILR Review* **74**(1), 95–130.

Schiprowski, A. (2020), ‘The role of caseworkers in unemployment insurance: Evidence from unplanned absences’, *Journal of Labor Economics* .

Van den Berg, G., Kjærsgaard, K. and Rosholm, M. (2012), ‘To meet or not to meet (your case worker) - that is the question’, *IZA Discussion Paper 6476* .

A Additional Tables and Figures

Table A1: Caseworker assignment rules in Danish jobcenters

	Jobcenters using a given assignment rule							
	2011	2012	2013	2014	2015	2016	2017	2018
UI-fund	5	5	5	5	12	16	20	22
Industry	2	2	2	2	1	1	1	1
Caseload size	13	13	13	13	15	13	13	13
Birthday	28	28	28	28	21	20	18	17
Random[(1)]	10	10	10	10	9	9	8	8
Residual[(2)]	8	8	8	8	8	7	6	5
Obs	66	66	66	66	66	66	66	66

Note: The table is based on a survey conducted among Danish jobcenters in Spring 2020. The survey asked how the jobcenters in a given year from 2011-18 have been assigning jobseekers to caseworkers (specifically, how it was decided which caseworker that would participate in the jobseeker's first meeting). Note that 66 refers to the number of jobcenters that completed the survey (5 jobcenters could not help and 23 did not respond).

[(1)] Refers to jobcenters reporting that it is 'random'. Although intriguing, it should be noted that it was unclear how the randomness in assignments were ensured.

[(2)] Residual category refers to jobcenters i) that have only 1 or 2 caseworkers, ii) where they report that the 'jobseeker is choosing her caseworker' or iii) where the respondent does not know about teh particular year.

Table A2: Sampling

	Meetings	Spells	Jobseekers	Caseworkers	Jobcenters
Caseworker data	**4,401,504**	**1,110,742**	**651,089**	**14,043**	**94**
No missing covariates[(1)]	4,395,767	1,109,387	650,373	14,036	94
Timing of first meeting[(2)]	4,002,681	1,000,816	606,610	13,628	94
Distance to previous spell[(3)]	3,509,206	865,551	584,912	13,292	94
Identify assignment caseworker and meeting [(4)]	3,506,162	864,036	584,148	13,290	94
Only the assignment meeting[(5)]	864,036	864,036	584,148	10,088	94
Meeting type[(6)]	803,372	803,372	557,007	7,632	94
Meeting contact[(7)]	787,073	787,073	548,383	7,468	94
Assignment mechanism[(8)]	151,953	151,953	106,782	1,486	29
Caseload size[(9)]	141,998	141,998	101,479	612	25
Caseworker work experience [(10)]	141,088	141,088	100,930	601	25
Min. 2 caseworkers per randomization cell[(11)]	140,781	140,781	100,698	601	25
Sample	**140,781**	**140,781**	**100,698**	**601**	**25**
Sample with r^2 restriction [(12)]	**103,027**	**103,027**	**75,811**	**467**	**24**

Note: The caseworker data consists of all jobseekers who entered unemployment from 2011-2018 and has at least one meeting registration in a jobcenter.

(1) I drop observations if predetermined characteristics are missing. This includes demographics (age, gender, immigrant, decendent, western, married, children), labor market history (past unemployment, employment, receipt of public transfers) and UI-fund association. Note that I allow for missing education and/or industry, and code missings as a category.

(2) I require that the first meeting takes place in the same week or up to 13 weeks after spell start. This is based on the rule sin PES specifying that the first meeting had to take place within 3 months (1 month) until 2015 (since 2015).

(3) I require min 10 weeks between the current and any potential previous UI-spell. This is to be sure that we are observing new UI-spells and not a continuation of old spells.

(4) I identify the *assigned* caseworker based on the jobseeker's *first individual meeting*. In most jobcenters, jobseekers have the first meeting with their assigned caseworker. In a few jobcenters, jobseekers first go to a group meeting (an information meeting with a group of other jobseekers) and only in the second meeting meet their assigned caseworker. I therefore define the *assigned caseworker* as the caseworker from the first meeting if this is not an information meeting (coded as "Informationsmøde"), and otherwise as the caseworker from the second meeting. Note that some jobseekers leave unemployment before making it to the second meeting, and I cannot define an assignment caseworker for these jobseekers. Hence, they are dropped from the sample.

(5) I restrict the sample to the first individual meeting only. Hence the level of observation goes from spell$\times$meeting to spells.

(6) I require that the first individual meeting is coded as a "job interview" ("jobsamtale" or "jobsamtale med deltagelse af a-kasse")

(7) I require that the first individual meeting takes place in person ("personlig kontakt")

(8) Based on the jobcenter survey, I know the yearly assignment mechanism for all responding jobcenters over the sample period. I keep all jobcenter $\times$ years in which they used report to use "birthday" assignment.

(9) I restrict the sample to caseworkers that have at least 50 cases

(10) I restrict the sample to caseworkers who work in at least 10 weeks

(11) Randomization occurs at the jobcenter $\times$ unit $\times$ year level, and hence I require at least two caseworkers per randomization cell.

(12) For each caseworker, I test how well jobseeker birth day of the month (1-31) predict assignment to a particular caseworker. In particular, for each caseworker, I run a regression with a dummy for assignment to casewrker j on birthday $\times$ quarter an jobcenter $\times$ unit $\times$ year FE's. I then save the within-r^2 from each of these regressions. To have a benchmark, I run a similar set of regressions including placebo birthdays instead of true birthday fixed effects (again interacted with quarters). Finally, I drop caseworkers if the within-r^2 on true birthdays are lower than the median within-r^2 for placebo birthdays, i.e. if the within-$r^2 < 0.17$. Note that I again check that I have at least 2 caseworkers per randomization cell after this last restriction.

Table A3: Jobseeker summary statistics

	mean	sd
Demographics[1]		
Age	40.231	(12.280)
Male	0.484	(0.500)
Immigrant	0.050	(0.218)
Descendant	0.002	(0.043)
Western	0.957	(0.203)
Married	0.422	(0.494)
Number of children	0.806	(1.039)
Education[2]		
0. Missing	0.013	(0.115)
15. Preparatory course	0.004	(0.062)
20. Upper secondary	0.047	(0.212)
30. Vocationa educ.	0.476	(0.499)
35. Qualifying program	0.000	(0.022)
40. Short cycle tertiary	0.047	(0.211)
50. Vocational bach.	0.125	(0.331)
60. Bachelor	0.014	(0.118)
70. Master	0.049	(0.215)
80. PhD	0.003	(0.053)
Labor market history[3]		
UI-benefits in year t-1	0.408	(0.491)
UI-benefits in year t-2	0.424	(0.494)
Any employment in year t-1	0.900	(0.300)
Any employment in year t-2	0.922	(0.268)
Employment rate in year t-1	0.642	(0.352)
Employment rate in year t-2	0.659	(0.348)
Wage earnings in 1,000 DKK in year t-1	212.235	(147.811)
Wage earnings in 1,000 DKK in year t-2	209.295	(141.285)
Number of employers in year t-1	1.398	(0.944)
Number of employers in year t-2	1.407	(0.907)
Public transfers in year t-1	0.627	(0.484)
Parental leave in year t-1	0.083	(0.276)
Education subsidy in year t-1	0.101	(0.302)
Previous industry[4]		
Real estate	0.011	(0.106)
Business services	0.115	(0.320)
Finance	0.010	(0.098)
Trade & transport	0.199	(0.400)
Manufacturing	0.123	(0.328)
Communication & it	0.017	(0.130)
Culture	0.036	(0.185)
Agriculture, forestry & fishing	0.026	(0.159)
Public administration, health, education	0.224	(0.417)
UI-association[4]		
Academics Association	0.064	(0.245)
Danish Trade Union Association	0.654	(0.476)
No association (yellow)	0.282	(0.450)
Obs	103,027	

Note: The table reports means and standard deviations of predetermined characteristics for jobseekers.
(1) *Demographics* rely on information from the population register (BEF and DREAM). Male, immigrant, descendant and married are dummies, while number of children is a count variable.
(2) *Education* rely on information from the education register (UDDA) and is based on the highest completed education. Omitted category is "10 Primary education".
(3) *Labor market history* variables rely on a register containing weekly information on UI-benefits and transfers (DREAM) and on the income register (Eindkomst). UI-benefits, any employment, public transfers, parental leave and education subsidy are all dummies. The employment rate, wages and number of employers are continuous and winsorized at the 99th percentile.
(4) *Previous industry* is based on the DREAM-register. It represents the dominating industry for the individual in the 12 months prior to the UI-spell start (the industry in which the individual had highest accumulated earnings). Omitted category is "Construction".
(5) *UI-association* are based on information about the individual's UI-fund membership from the DREAM-register. There are 25 UI-funds belonging to the Trade Union Association, the Academics Association or no association.

Table A4: How much can jobseeker characteristics explain?

	Obs	Mean	SD Total	SD Within	SD Within & controls
A: UI-spells					
Raw	103,027	27.0	23.9	23.3	22.7
Truncated at 52 weeks	103,027	23.6	16.0	15.6	15.2
B: Accumulated wages and hours by month 24					
Accumulated hours	89,249.0	1,857.2	1,167.3	1,132.6	1,075.1
Accumulated wages	89,249.0	339,052.7	238,786.2	231,202.7	208,060.8

Note: Column 1 shows the number of observations, column 2 the unconditional mean and column 3 the total standard deviation (across all jobseekers). Column 4 shows the standard deviation across jobseekers *within* the same jobcenter unit and year. Column 5 shows the standard deviation across jobseekers *within* the same jobcenter unit and year and *controlling* for all predetermined jobseeker characteristics in table A3 (as well as the jobseeker's UI-fund).

Table A5: Labor market states

	State in a given week 13	26	39	52	104
A: Relative to spell start					
UI-benefits	0.64	0.41	0.32	0.30	0.19
Other transfers	0.06	0.11	0.13	0.15	0.19
Education subsidies	0.01	0.02	0.03	0.03	0.04
Employed	0.24	0.42	0.48	0.48	0.53
Dropped out	0.04	0.04	0.04	0.04	0.05
B: Relative to spell end					
UI-benefits	0.15	0.18	0.20	0.19	0.14
Other transfers	0.18	0.18	0.19	0.19	0.20
Education subsidies	0.04	0.04	0.04	0.04	0.04
Employed	0.46	0.52	0.51	0.51	0.54
Dropped out	0.17	0.08	0.07	0.07	0.07

Note: The table shows the average labor market state for jobseekers in a given week measured relative to the start or end of the UI-spell (see definition in table 2). The labor market states are based on DREAM and EINDKOMST, respectively containing information on weekly payouts of transfers and wage earnings. Based on these registers, I define five mutually exclusive states. In a given week, the individual may receive UI-benefits, education subsidies, some other type of public transfers (e.g. cash benefits), she may be employed (positive wage earnings, no transfers) or she may have dropped out of the labor force (no wage earnings, no transfers). Panel A measures the labor market state relative to UI-spell start, while panel B measures the state relative to UI-spell start.

Table A6: Mechanisms

	Obs	Mean	SD Total	SD Within
A: Meetings (1)				
Weeks until 1st meeting	103,027	4.48	3.32	3.10
Meetings per week	103,027	0.17	0.12	0.11
B: Job search (2)				
Joblogs per week	39,621	1.71	0.84	0.83
Search methods	38,694	0.64	0.17	0.17
Search channels	38,694	0.59	0.17	0.17
Part time job	38,694	0.14	0.22	0.21
Commute distance (km)	21,176	30.77	25.22	22.91
B: Training assignments (3)				
Weeks until any training	44,822	15.96	12.71	10.68
Weeks until course training	35,031	17.04	14.87	12.58
Weeks until job training	22,340	23.89	18.39	16.59
Assignment to any training	103,027	0.44	0.50	0.48
Assignment to course training	103,027	0.34	0.47	0.45
Assignment to job training	103,027	0.22	0.41	0.40

Note: Column 1 shows the number of observations, column 2 the unconditional mean, column 3 the total standard deviation (across all jobseekers) and column 4 the standard deviation across jobseekers *within* the same jobcenter unit and year.
(1) *Weeks until first meeting* refers to the number of weeks between the UI-spell start and the first meeting in the jobcenter. *Meetings per week* refers to the number of caseworker meetings pr. week of unemployment. If the UI-spell is longer than 26 weeks, I only consider the first 26 weeks, because meeting frequency requirements change after 26 weeks.
(2) Job search variables represent joblog-behavior in the average week (not considering behavior in the first and last two weeks of the UI-spell, see(Fluchtmann et al., 2020)). *Joblogs per week* refers to the number of joblogs per week *Search methods* refers to the number of search methods per joblog, and *search channels* refers to the number of search channels per joblog. *Part time job* refers to the fraction of applied-for jobs that were part time. *Commute distance* refers to the average commute distance (in kilometers) of applied-for jobs.
(3) Training assignments come from the individual job plans. *Assignment to any training* is a dummy taking value one if any type of training is noted in the jobseeker's job plan durng her UI-spell. Likewise, for course training and job training. *Weeks until any training* shows the number of weeks between UI-spell start and the start of the training activity. Likewise for course and job training.

Table A7: Birthday test

	P-values	
	No controls	Control for non-western
Demographics		
Age	0.222	0.127
Male	0.380	0.381
Immigrant	0.022	0.073
Descendant	0.694	0.706
Western	0.000	.
Married	0.004	0.425
Number of children	0.027	0.462
Education		
0. Missing	0.149	0.360
15. Preparatory course	0.951	0.950
20. Upper secondary	0.251	0.298
30. Vocationa educ.	0.321	0.611
35. Qualifying program	0.165	0.183
40. Short cycle tertiary	0.456	0.466
50. Vocational bach.	0.789	0.908
60. Bachelor	0.486	0.484
70. Master	0.645	0.676
80. PhD	0.375	0.447
Labor market history		
UI-benefits in year t-1	0.314	0.307
UI-benefits in year t-2	0.235	0.230
Any employment in year t-1	0.027	0.047
Any employment in year t-2	0.270	0.418
Employment rate in year t-1	0.091	0.127
Employment rate in year t-2	0.061	0.164
Wage earnings in 1,000 DKK in year t-1	0.042	0.306
Wage earnings in 1,000 DKK in year t-2	0.038	0.241
Number of employers in year t-1	0.073	0.088
Number of employers in year t-2	0.366	0.388
Public transfers in year t-1	0.943	0.950
Parental leave in year t-1	0.594	0.559
Education subsidy in year t-1	0.074	0.137
Previous industry		
Real estate	0.394	0.388
Business services	0.221	0.282
Finance	0.451	0.507
Trade & transport	0.484	0.525
Manufacturing	0.160	0.162
Communication & it	0.080	0.136
Culture	0.352	0.379
Agriculture, forestry & fishing	0.027	0.026
Public administration, health, education	0.540	0.579
UI-fund association		
Academics Association	0.222	0.412
Danish Trade Union Association	0.672	0.685
No association (yellow)	0.221	0.342

Note: Level of observation in all regressions of the table is a job seeker (not a spell). This is to ensure that birthdays do not mechanically predict job seeker characteristics.

Table A8: Magnitude of caseworker effects
Non-truncated UI-spells

	UI-spells		Caseworker VA		Compare
	Mean	SD	SD	P-value	col 3 / col 2
FE	27.31	23.65	3.34	0.00	0.14
Empirical Bayes	27.31	23.65	1.19	0.00	0.05

Note: In contrast to table 4, the caseworker effects shown in this table is based on non-truncated UI-spells. Here, the censoring of UI-spells is handled by reducing the sample period by one year such that only UI-spells initiated from 2011-2017 are included in the estimation (instead of all UI-spells initiated from 2011-18). This reduces the sample to 90,050 UI-spells, 67,203 jobseekers and 419 caseworkers. Column 1 shows the unconditional mean for UI-spells truncated at 52 weeks, while column 2 shows the standard deviation within jobcenter unit and year and controlling for country of origin, quarter and age fixed effects. Column 3 shows the standard deviation on caseworker effects. The first row reports the standard deviation of the caseworker fixed effects opbtained by OLS estimation of (4). The second row reports the standard deviation directly obtained by restricted maximum likelihood estimation of (6) where the caseworker effects are treated as random. Column 4 shows the p-value on a test for joint significance of the caseworker effects (an F-test in row one, a likelihood ratio test in row two). Column 5 summarizes the magnitude of the caseworker effects by setting the SD of caseworker effects relative to the within-SD (col 3 / col 2).

Table A9: Out-of-sample prediction

	Truncated UI-spell	
	Weeks	SDs
Caseworker VA	-1.066***	-0.066***
	(0.137)	(0.009)
Obs	103027	103027
Dep var Mean	23.59	0.00
Dep var sd	16.04	1.00

Note: Both columns report the coefficient on caseworker value-added (leave-one-out empirical Bayes estimates). In column one, truncated UI-spells measured in weeks is regressed on caseworker value-added, while in column two, truncated and standardized UI-spells are regressed on caseworker value-added. All regressions include jobcenter $\times$ unit $\times$ year fixed effects, a control for country of origin, age and quarter fixed effects Standard errors are two-way clustered on caseworker and jobseekers. *p<0.10 ** p<0.05 *** p<0.01

Table A10: Rank-rank correlations of caseworker effects
Transitions, wages and hours

	Caseworker ranking		
	Transitions	Accum. wage month 12	Accum. hours month 12
Caseworker ranking			
Transitions	1		
Accum. wage month 12	0.559***	1	
Accum. hours month 12	0.611***	0.933***	1
Observations	467		

Note: The table shows correlations between caseworker ranking in terms of transitions out of unemployment against caseworker ranking in terms of accumulated wages or hours by month 12. Note that caseworkers always are ranked within a jobcenter unit and year, controlling for country of origin, quarter and age fixed effects. *Transition ranking* is based on truncated UI-spells and coded such that a caseworker ranks higher if the mean jobseeker has a shorter UI-spell. *Wage (hours) ranking* is based on accumulated wage (hours) by month 12 relative to unemployment start. It is coded such that a caseworker ranks higher if the mean jobseeker has higher wages (hours). *p<0.10 ** p<0.05 *** p<0.01

Table A11: Accumulated wages and hours since spell start

	Month since UI-spell start				
	Month 3	Month 6	Month 9	Month 12	Month 24
Accumulated hours	12.102***	22.928***	28.951***	32.747***	34.357***
	(1.722)	(3.439)	(4.481)	(5.419)	(8.825)
Accumulated wages	2321.091***	4685.928***	6135.867***	6749.145***	7371.113***
	(450.724)	(991.967)	(1349.305)	(1415.428)	(2067.419)
Obs	103027	103027	103027	101346	89241
Hours mean	104	301	540	791	1857
Wage empl. mean	19209	54768	97975	143649	339053

Note: Each cell in the first (second) row shows the coefficient on caseworker value-added from a regression where the dependent variable is a *accumulated* hours (wages) by a given month relative to unemployment start. All regressions include jobcenter × unit × year fixed effects, a control for country of origin, age and quarter fixed effects. Standard errors are two-way clustered on caseworker and jobseekers. *p<0.10 ** p<0.05 *** p<0.01

Table A12: Rank-rank correlations
Mean vs. percentiles

	Mean	10th pct	25th pct	50th pct	75th pct	90th pct
Mean	1					
10th pct	0.390***	1				
25th pct	0.559***	0.824***	1			
50th pct	0.817***	0.263***	0.492***	1		
75th pct	0.659***	-0.270***	-0.0748	0.515***	1	
90th pct	0.461***	0.308***	0.244***	0.136***	0.164***	1
Observations	467					

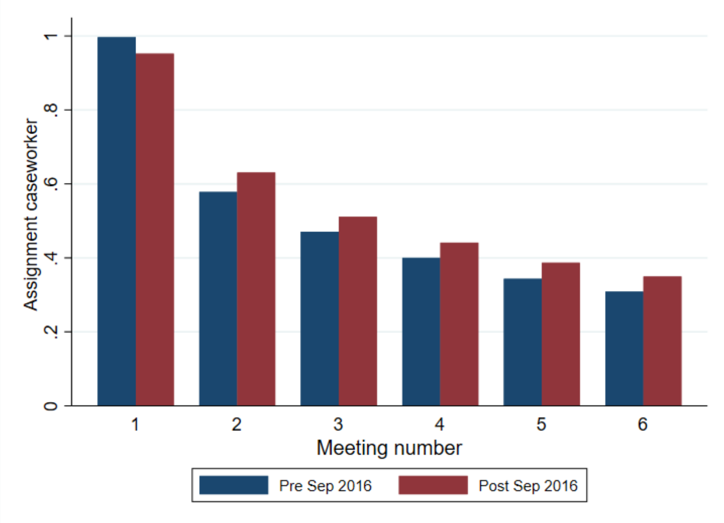
Note: The table shows correlations between caseworker ranking on the *mean* jobseeker's UI-spell against caseworker ranking on the UI-spell of some *percentile* of the jobseeker distribution. Three details of the caseworker rankings merit note. First, caseworkers always ranked within a jobcenter unit and year, controlling for country of origin, quarter and age fixed effects. Second, ranking is coded such that a caseworker ranks higher if the jobseeker of interest (the mean or a given percentile) has a shorter UI-spell. Third, to comply with data protection rules, all percentiles are based on at least 5 observations. *p<0.10 ** p<0.05 *** p<0.01

Table A13: Meetings by pre and post reform period

	Pre July 1st 2015		Post July 1st 2015	
	Weeks until 1st meeting	Meetings/week	Weeks until 1st meeting	Meetings/week
Caseworker VA	-0.424***	0.010***	-0.284***	0.011***
	(0.052)	(0.002)	(0.055)	(0.003)
Obs	58168	58168	44859	44859
Dep. var mean	4.839	0.144	4.006	0.206
Dep. var SD	3.635	0.114	2.785	0.119

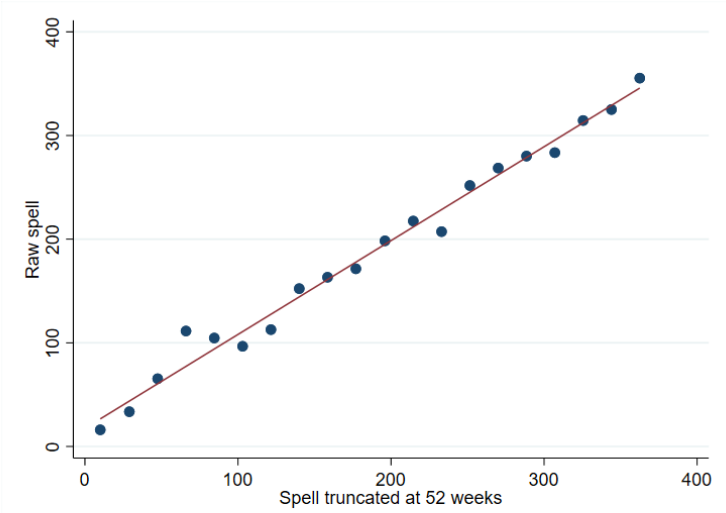
Note: *p<0.10 ** p<0.05 *** p<0.01

Figure A1: Caseworker survival
Pre and post self-booking



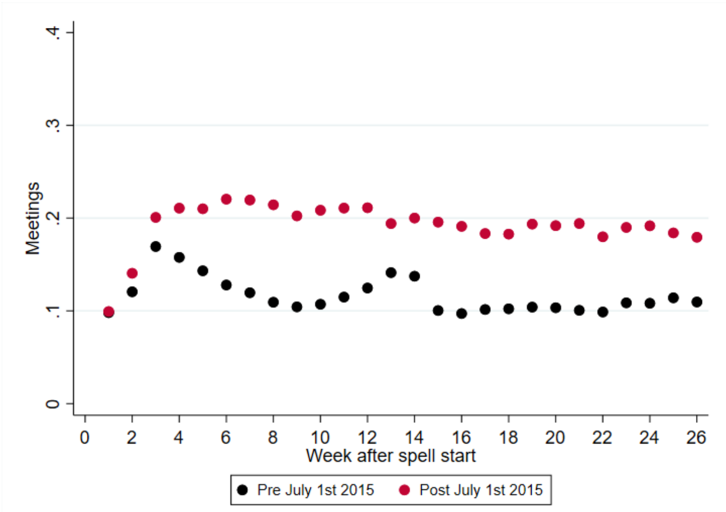
Note: The figure shows the fraction of jobseekers that have a given meeting in the UI-spell with their assigned caseworker (conditional on having the given meeting). Blue bars represent the period *before* September 2016, where all meetings were booked by the caseworker. Red bars represent the period *after* September 2016, where jobseekers had to book later meetings themselves (the first and one other meeting was still booked by the caseworker). The fraction is not one for the first meeting, since some jobseekers start with an information meeting and hence meet their assigned caseworker in the second meeting.

Figure A2: Rank-rank correlation of VA-measures
Truncated vs. non-truncated UI-spells



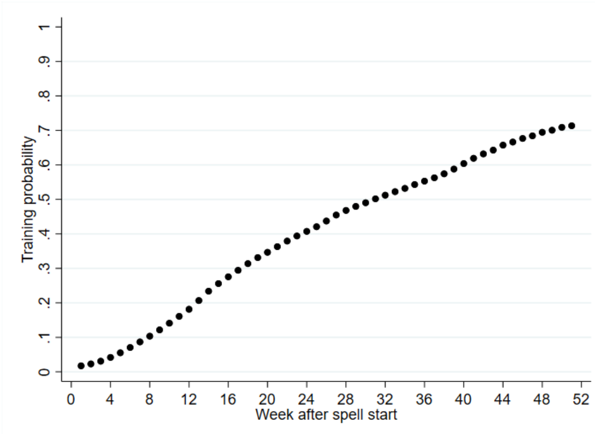
Note: The table compares the ranking of caseworkers based on empirical Bayes (EB) estimates using either truncated or non-truncated UI-spells as the outcome of interest. The EB estimates are obtained by restricted maximum likelihood estimation of (6). The EB estimates based on *truncated UI-spells* correspond to the baseline estimates of the paper. The truncation of UI-spells at 52 weeks makes it is possible to include all spells initiated from 2011-18. The EB estimates based on *non-truncated UI-spells* is an alternative to the baseline estimates. Here, the censoring of UI-spells is handled by reducing the sample period by one year such that only UI-spells initiated from 2011-2017 are included in the estimation. Caseworkers are ranked according to both EB estimates (ranking is coded such that a caseworker ranks higher if she is better at reducing UI-spell length). Note that the figure compares only the ranks of caseworkers working from 2011-17 (419 caseworkers), since these both will have an EB estimate based on the truncated and non-truncated UI-spells. The correlation between caseworker rankings on the two EB estimates is 0.882 (significant at the 99% level).

Figure A3: Meetings over the UI-spell

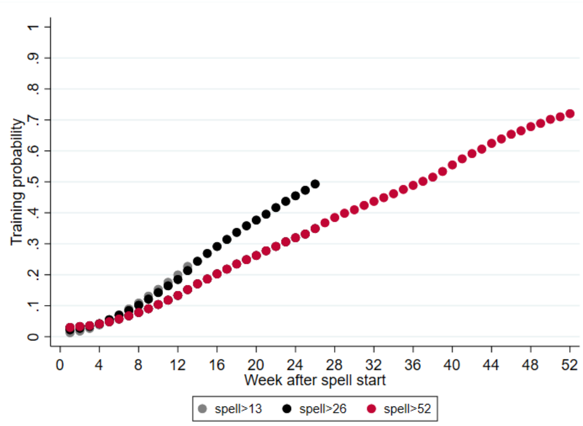


Note: The figure plots the average number of meetings in a given week of the UI-spell (for individuals surviving in unemployment to a given week). Black dots represent the average number of meetings in the pre-reform period (January 2011 - June 2015). Red dots represent the average number of meetings in the post-reform period (July 2015 - December 2018). The reform changed the frequency and timing of meetings over the first 6 months of unemployment.

Figure A4: Training probability over the UI-spell



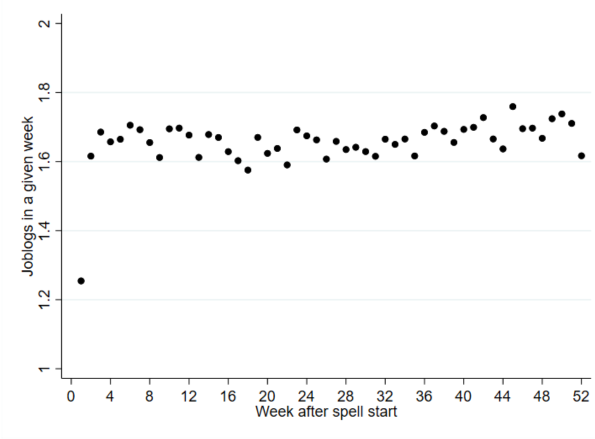
(a) Raw



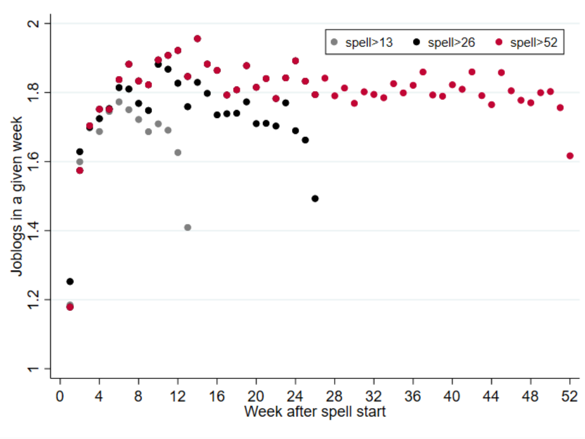
(b) Condition on spell length

Note: The figures plot the probability to be assigned to any type of training that *starts* in a given week of the UI-spell. This information is based on individual job plans. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

Figure A5: Joblogs over the UI-spell



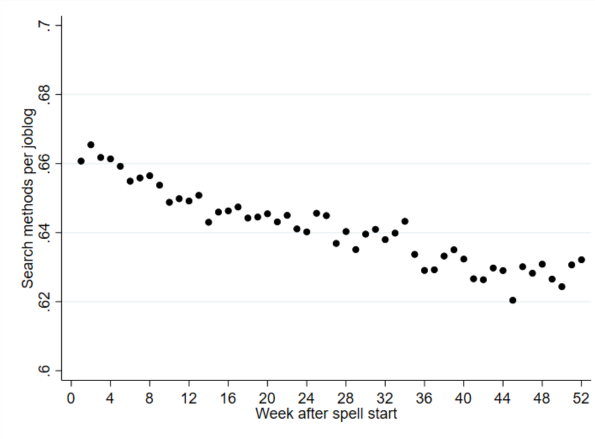
(a) Raw



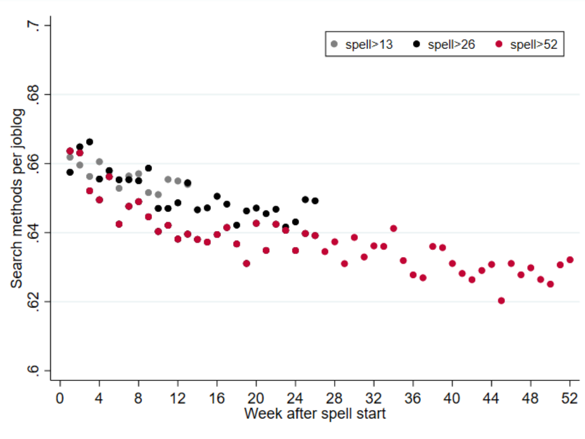
(b) Condition on spell length

Note: The figures plots the average number of applied-for jobs registered in Joblog in a given week of the UI-spell. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

Figure A6: Search methods over the UI-spell



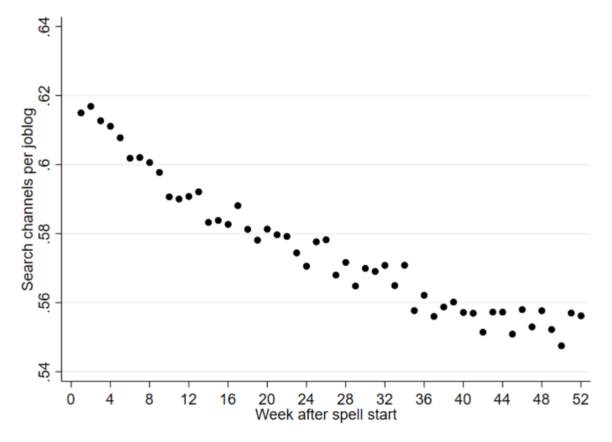
(a) Raw



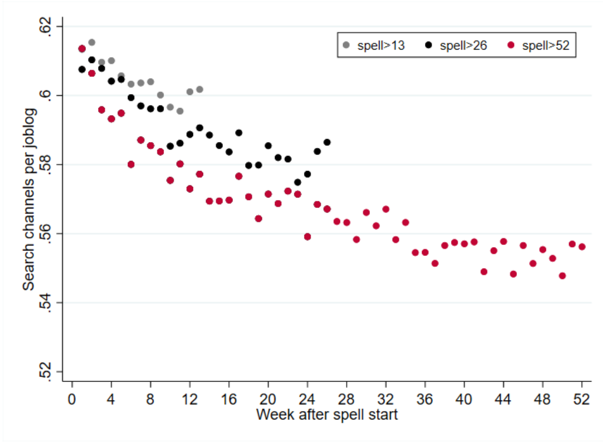
(b) Condition on spell length

Note: The figures plots the average number of search methods per joblog in a given week of the UI-spell. Examples of search methods are: ’via e-mail’, ’letter’, ’phone’, ’in person’. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

Figure A7: Search channels over the UI-spell



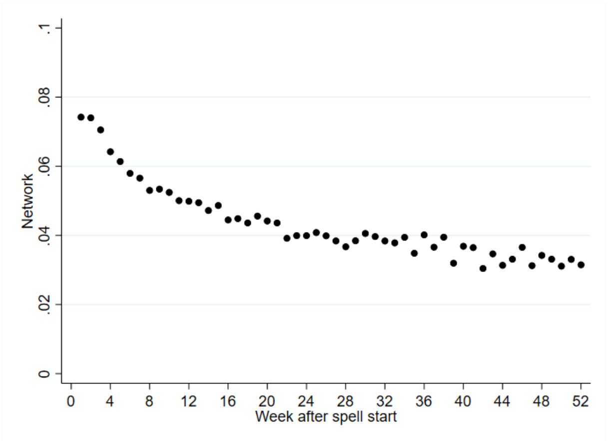
(a) Raw



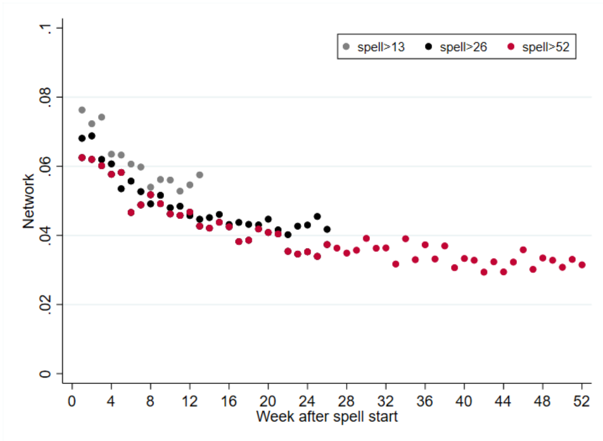
(b) Condition on spell length

Note: The figures plots the average number of search channels per joblog in a given week of the UI-spell. Examples of search channels are: 'unsolicited', 'advertised', 'through network'. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

Figure A8: Network over the UI-spell



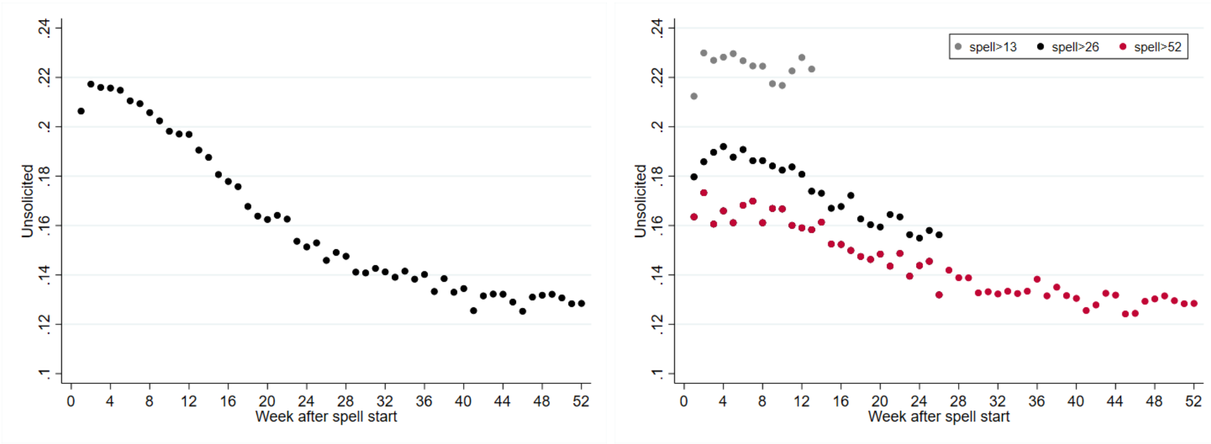
(a) Raw



(b) Condition on spell length

Note: The figures plots the fraction of applied-for jobs registered in the joblog in a given week that were found through the jobseeker's network. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

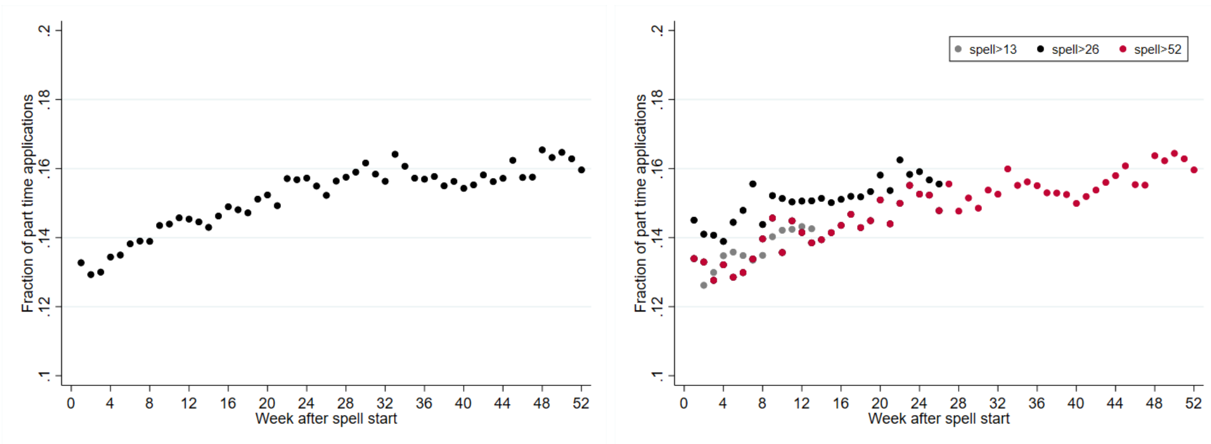
Figure A9: Unsolicited search over the UI-spell



(a) Raw (b) Condition on spell length

Note: The figures plots the fraction of applied-for jobs registered in the joblog in a given week that were sent unsolicited. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

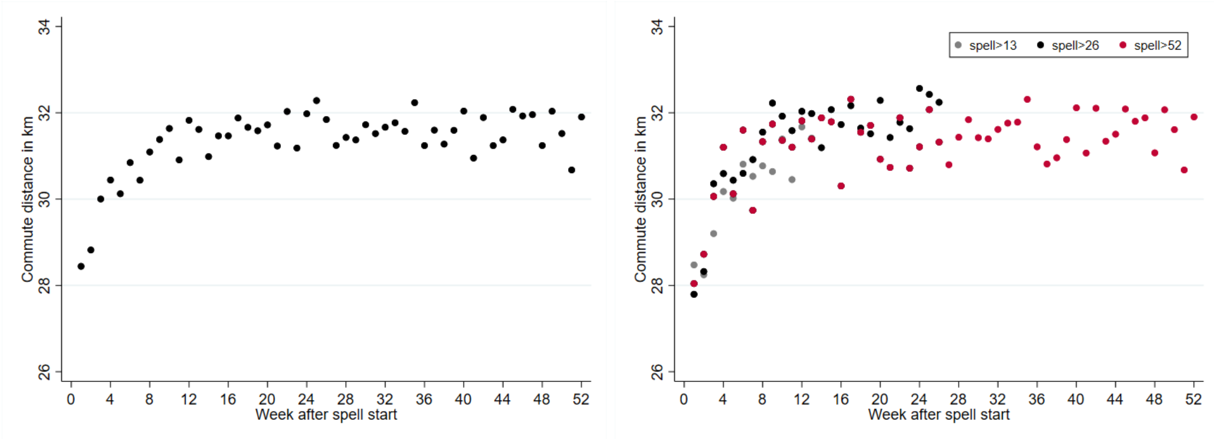
Figure A10: Part time job over the UI-spell



(a) Raw (b) Condition on spell length

Note: The figures plots the fraction of applied-for jobs registered in the joblog in a given week that were part time. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

Figure A11: Commute distance over the UI-spell



(a) Raw

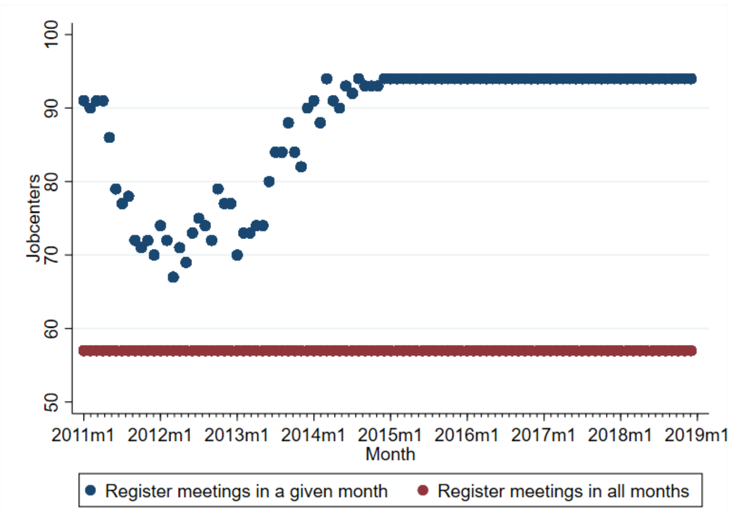
(b) Condition on spell length

Note: The figures plots the average commute distance (in kilometers) for applied-for jobs registered in the joblog in a given week of the UI-spell. This information is based on joblog data. Panel A plots the probability for all jobseekers, regardless of UI-spell length. Panel B conditions on the UI-spell length.

B Validation of Caseworker Data

This appendix validates the quality of the novel caseworker data described in section 4.1. Specifically, I focus on meetings registered between UI-benefit recipients[55] and their caseworkers in the Danish jobcenters from 2011-2018. Before turning to the validation exercise, a particular feature of the data set merits note. Figure B12 shows the number of jobcenters that register meetings with UI-benefit recipients in a given month from 2011-18. The figure shows that all 94 jobcenters continuously have been registering meetings since 2015. However, before 2015, there are months (years) in which some of the jobcenters do not register any meetings at all. This is not because there were no UI-benefit recipients in the municipality. Rather, it is likely explained by the fact that jobcenters are allowed to outsource tasks, e.g. meetings with UI-benefit recipients, to private entities.[56] Outsourcing of meetings with UI-benefit recipients would explain why some of the jobcenters do not register any meetings in a given year. It should be stressed that this unbalance in the caseworker data is no problem for the analysis conducted in this paper, since I always compare jobseekers assigned to different caseworkers *within* the same jobcenter unit and year. For the validation exercise, however, I focus solely on the 54 jobcenters that register meetings in all months from 2011-2018.[57]

Figure B12: Jobcenters registering meetings



Note: Blue dots represent the number of jobcenters that in a given month register (at least one) meeting with UI-benefit recipients. Red dots represents the number of jobcenters that register meetings with UI-benefits recipients in all months from January 2011 - December 2018.

[55]Besides restricting to jobseekers that are classified as UI-benefit recipients in the meeting registration data, I cross check that these individuals do in fact receive UI-benefit within 4 weeks from a meeting registration

[56]I interviewed one caseworker, who reported that the rules for reimbursement of the jobcenters' expenses - and hence also the incentive to outsource tasks - have changed over time.

[57]This is important when investigating data coverage on the extensive margin below.

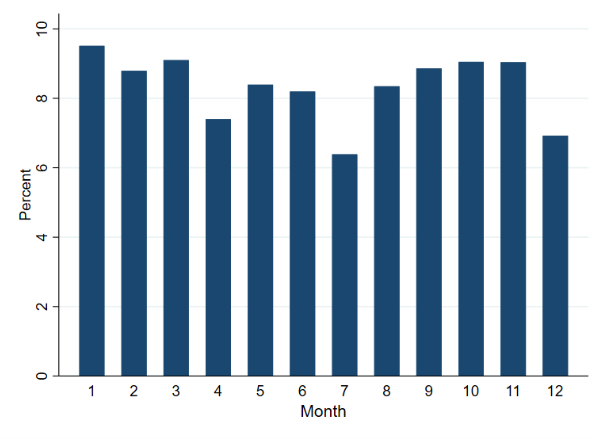
In the validation exercise, I address three concerns. The first concern regards *coverage* of the data. In particular, to what extent do jobseekers who become unemployed during the sample period also appear in the meeting registration data? To investigate this, I first identify all new UI-spells initiated from 2011-2018,[58] and I check how many of these UI-spells that have at least one meeting registration[59]. I find that 70 % of all new UI-spells appear with a meeting registration in the caseworker data. A likely explanation for why some jobseekers do *not* appear is that they exit unemployment before having a meeting: On average, the first meeting take place in week 5-6 of the spell, and at this point, around 60 % of jobseekers with no meeting registrations have exited, whereas only 7% of jobseekers with a meeting registration have exited. Thereafter, I check whether the number of meetings is in accordance with the official rules regarding meetings over the UI-spell. Here, it should be noted that the official rules have changed over the last decade: Before 2015, the requirement was 2-3 meetings during the first 6 months, and since 2015 the requirement has been 6 meetings. Overall, I find that meeting registrations are well in accordance with these rules: Before (after) 2015, jobseekers on average have 2.6 (5.4) meetings within the first 6 months. This suggests good coverage on both the extensive and intensive margin.

The second concern regards the *authenticity* of the meetings registered. There is no way to check whether the registered meetings actually took place, however, I can check for *misregistrations*. Namely, if the meeting registrations represent actual meetings, there should be no bunching of registrations on particular dates (e.g. December 31st), and the meetings should be registered to take place on days, where jobcenters are open, and during normal office hours. Appendix figure B13a and B13b indeed show that meeting registrations are smoothly distributed across day of the month and months, expectantly with fewer meetings during holiday seasons in December and July. Further, figure B13c and B13d show that 99 % of all meetings take place Monday-Friday and that more than 90 % of meetings taking place from 8am-5pm. I take this a evidence suggesting that meeting registrations do in fact represent real meetings.

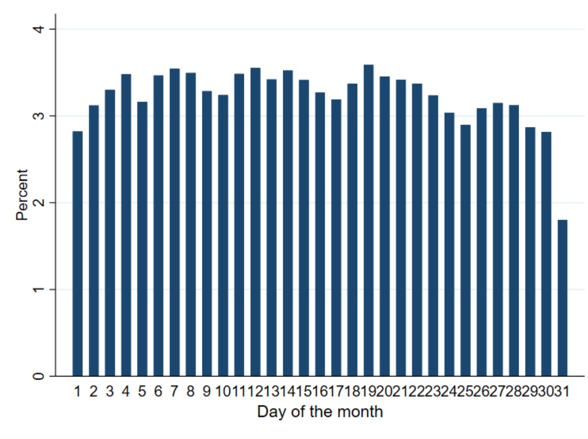
[58] For simplicity, I consider individuals who at UI-spell start are entitled to the maximum benefit duration. I identify these UI-spells using official 'placement dates' (indplaceringsdato) in the DREAM register. I exclude spells if there is no information about municipality

[59] I allow meeting registrations up to 31 days before spell start and until spell end

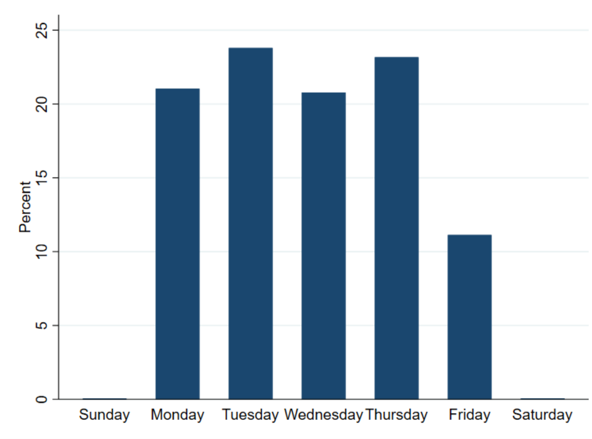
Figure B13: Meeting registrations



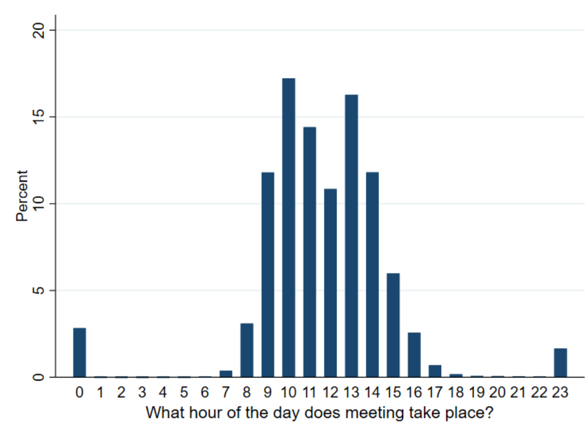
(a) Month



(b) Day of the month



(c) Day of the week

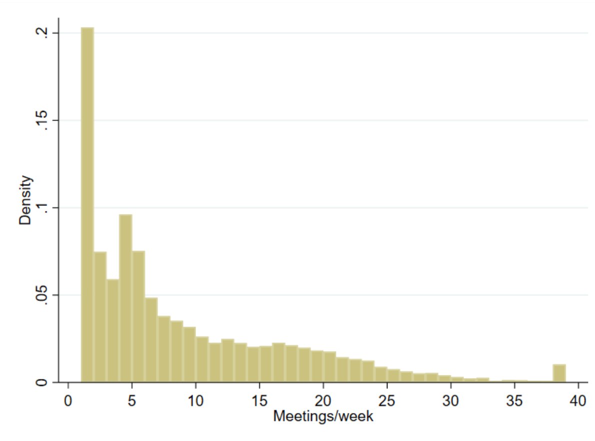


(d) Hour of the day

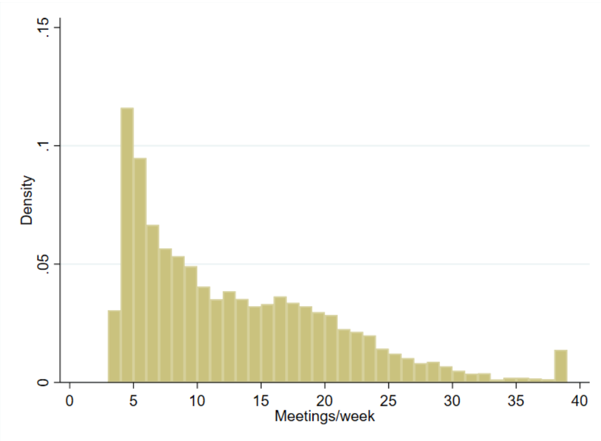
Note: Panel A shows the fraction of meetings registered to take place in a given month. Panel B shows the fraction of meetings registered to take place on a given day of the month Panel C shows the fraction of meetings registered to take place on a given day of the week. Panel D shows the fraction of meetings registered to take place during a given hours of the day.

A third concern regards the *caseworker identifiers* and in particular, whether the identifiers represent individual caseworkers or administrative employees registering meetings for a team of caseworkers. To investigate this, I rely on anecdotal evidence from jobcenters regarding the number of meetings that a caseworker has in a typical work week. The jobcenters generally say that they aim for caseworkers to have no more than 30 meetings per week. I compare this to the *registered* number of meetings/week per caseworker identifier in the data. Figure B14a shows the distribution of *registered* meetings/week for the caseworker identifiers. Two insights from this figure merit note. First, it is evident that less than 5 % of caseworkers have more than 30 meetings/week, which is reassuring given the jobcenter reports. It suggests that the caseworker identifiers do in fact represent individual caseworkers. Second, I find that caseworker identifiers on average have around 13 registered meetings/week if I condition on caseworkers that work in more than one week. Without conditioning, I find that the mean caseworker has less than 9 meetings/week. This seems low, but is driven by a large chunk of caseworkers (20 %) who only have (very few) meeting registrations. These likely represent 'test identifiers', e.g. when jobcenters adopt new IT-systems or new caseworkers are hired. These will not be part of the analysis sample of this paper, since I restrict to caseworkers that work for at least 3 months.

Figure B14: Meetings/week per caseworker



(a) All caseworkers



(b) Caseworkers with > 1 working week

Note: The figures plot the number of registered *meetings/week* for caseworker identifiers in a validation sample consisting of the 54 jobcenters that register meetings in all months from 2011-28. Figure A shows the distribution for all caseworker identifiers, while figure b shows that for caseworker identifiers with more than one working week. The *denominator* in the plotted metric, meetings/week, is the total number of unique *meeting slots* per caseworker. Note that I use unique meeting slots (time and date) instead of meeting identifiers, because some caseworkers hold information meetings for a *group* of jobseekers. Hence, an information meeting would count as multiple meetings if I use meeting identifiers instead of meeting slots. The *nominator* in the plotted metric, is the imputed number of working weeks for the caseworker. Note that I have no information about how many weeks the caseworkers were employed nor whether they were full or part time employed. However, I can impute the number of working weeks from the meeting registration dates. I do this by counting the number of days that a caseworker registers at least one meeting and divide by 5 (days in normal work week). This gives me an imputed number of working weeks.

Overall, I take this validation exercise as evidence suggesting that the caseworker data reflects real meetings between jobseekers and caseworkers in Danish jobcenters.

Chapter 2

Does Vocational Training Help Jobseekers? Evidence From Quasi-Random Caseworker Assignments

Does Vocational Training Help Jobseekers?

Evidence From Quasi-Random Caseworker Assignments *

Anders Humlum† **Mette Rasmussen**‡

March 1, 2021

Abstract

We estimate the effectiveness of vocational training for unemployed workers using a novel caseworker leniency instrument. Leveraging rich administrative data from Denmark, our identification strategy exploits that i) jobseekers are quasi-randomly assigned to caseworkers, and ii) caseworkers differ in their propensities to assign jobseekers to vocational training. Using caseworker assignment propensity as an instrumental variable, we cannot reject that training courses on average have zero impacts on labor market outcomes after one year. In contrast, OLS regressions show strong negative correlations between training and employment, indicating that it is workers with adverse job prospects who select into training. To investigate whether vocational training is more beneficial for workers who are exposed to rapid structural change, we zoom in on jobseekers whose previous jobs were in manufacturing. Although the estimates are noisy, we find economically significant long-run benefits to vocational training for former manufacturing workers.

*This paper builds on the sample defined and constructed in Rasmussen (2021). The section on data and institutional setting therefore repeats part of the text from Rasmussen (2021)

†University of Chicago; humlum@uchicago.edu

‡University of Copenhagen; mette.rasmussen@econ.ku.dk

1 Introduction

Mismatches between worker skills and employer demands cause structural unemployment. This is costly for the individual and inefficient for society. Looking into the future, such mismatches could become more prevalent as globalization and automation continue to disrupt local labor markets (see McKinsey, 2017; OECD, 2019). How can policymakers help workers who have lost their jobs to foreign competition or a robot?

In this paper, we evaluate whether vocational training courses can help displaced workers reskill and find jobs. To identify the causal effects of training, we develop a novel *caseworker leniency* instrument. The instrument exploits that jobseekers in Denmark are quasi-randomly assigned to caseworkers who in turn differ in their propensities to send jobseekers on training courses. Our instrumental variable (IV) strategy constitutes a methodological leap relative to existing studies of vocational training which mainly have relied on a 'selection on observables' assumption (McCall et al., 2016).

To facilitate a comparison to existing methods, we estimate the effects of training assignment using both Ordinary Least Squares (OLS, assuming selection on observables) and our IV strategy (caseworker leniency). The OLS estimates show strong negative effects of training on employment. Participating in course training is associated with 25 percent fewer hours of work in month 12 after start of unemployment. In contrast, the IV estimates are less than half the size and statistically insignificant. The stark difference between IV and OLS reveals a strong pattern of selection whereby workers with adverse job prospects opt into training programs.

Using our panel data, we estimate how the effects of training play out at different time horizons. After large locking-in effects in the first year, the IV estimates become statistically insignificant and remain so over the next year. That is, even after two years, we find no evidence of a positive post-program effect of assignment to training. The standard errors of our IV estimates are large, however, which could mask important heterogeneity in the effects of training across courses or workers.

To investigate heterogeneity in the effects of vocational training, we zoom in on jobseekers from the *manufacturing* sector. The idea here is to look for jobseekers who have higher exposure to globalization and robot adoption, and hence may be more likely to face mismatch problems in the labor market (Humlum and Munch, 2019). In line with this prediction, we find economically significant long-run benefits of vocational training for displaced manufacturing workers.

Although statistically insignificant, the estimates suggest that assignment to vocational training increases wage earnings by around 20,000 Danish Kroner (3,200 US dollars) and employment by 150 hours in month 24 after start of unemployment.

These estimated benefits to vocational training stand in contrast to the existing literature which concludes that training programs are a bad investment (Jespersen et al., 2008). Training is generally found to prolong the time spent in unemployment, and it is not clear that training has significant effects on wages nor on later employment (Jespersen et al., 2008; Munch and Skipper, 2008; Rosholm and Svarer, 2008). Taking the substantial administration costs into account, training generates a significant deficit (Jespersen et al., 2008). However, the existing studies suffer from two key shortcomings that we address in this paper.

First, the papers do not address the selection problem that workers opt into training based on unobservable job prospects (Ashenfelter, 1978).[1] Our findings show that solving this selection problem is crucial for uncovering the causal effects of training.

Second, due to data limitations, many of the existing studies lump training courses together, thus ignoring that different courses could have very different costs and benefits. We have data that allows us to distinguish between different types of training interventions. Furthermore, a unique feature of the *caseworker leniency* instrument is that it enable us to separately identify causal effects for the different *types* of training interventions. Combining this with unusually rich data on jobseekers from the Danish administrative registers allows us to shed light on the heterogeneous effects of training across courses and worker types.

In developing our instrument, we build on recent methodological advances on *judge leniency designs* (Dahl et al., 2014; Bhuller et al., 2020). In particular, 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 *exogeneity* assumption by validating that caseworker assignments do not correlate with observable characteristics of jobseekers. Second, we verify that caseworker leniencies are highly *relevant* for training assignments and that the effect likely is *monotone* across jobseekers. Finally, we support the *exclusion* restriction (that caseworker leniencies only impact jobseeker outcomes through training assignments) by showing that our conclusions are robust to controlling for the meeting frequencies of caseworkers.

In Section 2, we first lay out theoretical mechanisms through which vocational training

[1] The existing studies mainly rely on 'selection on observables'. McCall et al. (2016) conclude that "more rigorous evaluations of causal effects are needed".

may impact jobseekers. In Section 3, we describe the institutional context of unemployment and caseworker assignments in Denmark. In Section 4, we describe our data on caseworker assignments, training interventions, and jobseeker outcomes. In Sections 5 and 6, we specify our empirical strategy and test the IV assumptions. In Section 7, we estimate the causal effects of vocational training on jobseeker outcomes. Section 8 concludes.

2 How Could Vocational Training Impact Jobseekers?

This paper investigates the benefits of vocational training for unemployed jobseekers. It should be noted that it is not obvious how exactly to define the treatment of vocational training. It could be defined as assignment to, enrolment in or completion of a training program. As noted by McCall et al. (2016), *completion* has the virtue that we learn about the benefits of achieving the specific skills targeted by the program. On the other hand, *assignment* to training is likely the most policy relevant treatment, since this is what policy makers can do. The caseworker can assign the jobseeker to training, yet the jobseeker could leave unemployment before enrolling or completing the program (see discussion below). In this paper, we will define the treatment intervention as the *assignment* to training.[2]

We now briefly discuss how caseworker assignment to training courses potentially could affect employment, wages, and time spent in unemployment. We refer to the next section for details on the assignments, including rules for compliance. Our point of departure is the notion that *mismatches* between worker skills and employer demands cause structural unemployment. Vocational training may therefore reduce unemployment if the courses help workers reskill to meet demands in the labor market. This chain of reasoning rests on two premises. First, that assignments of training actually leads to enrollment and completion. Second, that the assigned courses produce skills that are in demand by employers.[3] We base the discussion of potential effects of training assignments on these two premises.

Assuming first that assignment leads to course completion, there are potentially two effects at play. The first is a well-known (and well-established) *locking-in effect*: Course trainees spend less time searching for jobs which tends to prolong their time spent unemployed (McCall et al.,

[2] We leave it for future work to investigate the effect of enrolment in and completion of training courses.

[3] 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).

2016; Jespersen et al., 2008; Munch and Skipper, 2008; Rosholm and Svarer, 2008; Rosholm and Skipper, 2009). However, over time, as the training courses end, the locking-in effects will dissipate. Whether the treated workers catch up to, or even overtake, the control groups depends on whether the course contents actually meet employer demands in the labor market.[4] We may expect this *post-program effect* to be more prominent for workers who are exposed to disruption in the labor market, such as production workers competing with robots and offshoring.[5]

Once we relax the assumption that course assignment leads to enrollment and participation, we open the possibility that these assignments could entail *threat effects*. Threat effects arise if the individual, upon assignment, exits unemployment merely to avoid participating in the training program. Several papers find sizable threat effects (Black et al., 2003; Rosholm and Svarer, 2008; Geerdsen, 2006; Geerdsen and Holm, 2007), which tend to generate positive effects on employment in the short run.

As our intervention of interest is *assignment* to training, we note that our estimates will reflect both threat, locking-in, and post-program effects. As the different effects play out at different time horizons, we expect the impacts of training to be non-stationary. To capture these dynamic effects, we follow workers over a two-year continuous period.

3 Institutional Setting

In Denmark, unemployed individuals may receive unemployment insurance (UI) benefits in up to two years. To be eligible, the individual must have paid contributions to one of the 24 different UI-funds in Denmark[6] for at least 12 months and have accumulated one year of full-time work over the previous three years.[7] The level of UI benefits constitute 90 percent of prior monthly wages up to a maximum of 18,866 DKK (3,075 USD). Around 85% of the Danish

[4]Rosholm and Skipper (2009) discuss why training potentially could have *negative* post program effects. They argue that since treated individuals obtain new skills through training, it might make them search more narrowly or increase the reservation wage. This would tend to decrease the employment probability even after participation in the program.

[5]Hummels et al. (2018) make the point that workers displaced due to offshoring or robot adoption have skills that are obsolete and may find it more difficult to reattach to the labor market without retraining compared to workers displaced to other events.

[6]The UI-funds administer the pay-out of UI-benefits, which are partly financed by membership contributions and partly by government subsidies. General eligibility rules are set by public policy. UI-funds typically target a specific profession, occupation or education group, and they are organized in two larger associations, the Trade Union Association and the Academic Association (except for a few UI-funds that are not organized).

[7]Full-time work can be in terms of hours (min 1,924 working hours) or earnings (minimum 238,512 Danish Kroners in 2020). Graduates are excepted from the latter requirement, but receive a lower level of UI benefits.

wage earners are members of a UI-fund, and UI-benefit recipients account for around 75% of the unemployment in Denmark.

Besides the eligibility criteria described above, there are a number of requirements that the individual must live up to in order to receive UI benefits. For instance, the individual must be able to take a job with one day's notice and must in general actively search for jobs.[8] Most importantly for this paper, the individual must attend regular meetings with a caseworker from the local jobcenter, and this caseworker may assign her to training programs. In this section, we describe rules regarding these meetings and how jobcenters practically organize themselves, including the assignment of caseworkers. Thereafter, we describe the system of vocational training courses that jobseekers in Denmark may be assigned to.

3.1 Caseworker Meetings in Danish Jobcenters

There are 94 jobcenters in Denmark, each of which are responsible for the public employment services for individuals residing in a given municipality.[9] When an individual becomes unemployed, she must register at her local jobcenter and attend regular meetings with a caseworker.[10] The required meeting frequency has increased over the last decade: Before 2015, jobseekers were supposed to meet with a caseworker at least every third month[11], while jobseekers since 2015 have been required to meet with a caseworker on a monthly basis.[12] Upon registration in the jobcenter, the jobseeker will be assigned a caseworker, who invites the jobseeker for the first of these meetings. The organization of caseworkers and the exact rule for assignment is key to our identification strategy and will be explained in detail.

Based on a survey among Danish jobcenters, Rasmussen (2021) documents that many jobcenters in Denmark organize their caseworkers in smaller units. Typically, the units are defined in terms of jobseeker age (above or below 30) or profession (academic or non-academic).[13]

[8] Since 2015, jobseekers have been required to document their job search in an online system called *joblog* (Lov nr. 548 af 07/05/2019 §20).

[9] There are 98 municipalities in total, and 94 jobcenters. Four of the jobcenters covers two municipalities.

[10] Although this paper relies on meetings with caseworkers from the jobcenter, we note that the jobseeker also meets with a caseworker from the UI-fund (but at a much lower frequency). Henceforth, we focus on the caseworkers from the jobcenters.

[11] Depending on age, the first meeting took place in the first or third month after unemployment start, and hereafter meetings took place every third month (LBK nr 990 af 12/09/2014, §17-19)

[12] Monthly meetings during the first 6 months of unemployment, thereafter meetings every third month. Note that the employment reform also implied more collaboration between jobcenters and UI-funds. For instance, the UI-fund may now participate in 2 of the meetings in the jobcenter (Lov nr 1486 af 23/12/2014, §16a)

[13] Anecdotally, the jobcenters interpret 'academic' broader than a college degree. Namely, it cold also reflect the

Hence, the jobseekers are first assigned to a unit, and hereafter, they are assigned a caseworker within the unit. As the jobcenter units therefore represent the level of randomization, we will include jobcenter unit and year indicators in all regressions. As in Rasmussen (2021), we divide all jobcenters into four units by fully interacting jobcenter, profession and age indicators. The profession indicator takes value one if the individual's UI-fund is a part of the Trade Union Association[14], and the age indicator takes value one if the individual is above 30.

Rasmussen (2021) further documents that a large subset of the jobcenters base caseworker assignment on jobseeker *birthdays*. That is, within a jobcenter unit, a jobseeker's birthday determines which caseworker she is assigned to. Some jobcenters have changed their assignment rules over time, e.g from an industry to a birthday rule or vice versa. In the average year from 2011-18, there are 24 jobcenters that report to use birthday rules for assignment (out of the 66 jobcenters that responded to the survey). Rasmussen (2021) further validates that i) the caseworker assignment is based on birth *day* of the month (1-31), and that ii) these birthdays (1-31) are *as-if random* in that they do not correlate with observables of the jobseeker.[15] These pieces of evidence suggest that caseworkers effectively are *quasi-randomly* assigned to jobseekers within a jobcenter unit. Our identification strategy will leverage this institutional feature.

It is important to note that the assigned caseworker invites and participates in the jobseeker's *first individual meeting* in the jobcenter. In the majority of the jobcenters, this corresponds to the first meeting in the jobseeker's UI-spell. However, in a few jobcenters, the meeting will be the second in the UI-spell, since the jobseeker first must participate in an information meeting with a group of other jobseekers. This knowledge allows us to identify the assigned caseworker from a data set containing information on the date and type of all meetings between jobseekers and caseworkers over the jobseeker's UI-spell.

Although the jobseeker is supposed to have all meetings with the assigned caseworker over her UI-spell, this need not be the case. Generally, there are two reasons for why caseworker switches might occur. First, it can be due to unplanned caseworker absences, where a col-

jobseeker's previous occupation or profession.

[14]The 24 UI-funds in Denmark typically target specific professions and/or education groups. The main part of the UI-funds are members of the Trade Union Association or the Academic Association. A few UI-funds are not members of any association, and in this analysis I will include them with jobseekers from the Academic Association.

[15]These tests are conducted conditional on country of origin. Danish immigration authorities assign immigrants with no birth certificate January 1st as their birthday. This predominantly happens for non-western immigrants and implies that they will be over-represented on January 1st.

league replaces the assigned caseworker. Second, the jobseeker could directly request a new caseworker. Here, it is relevant to note that jobseekers since 2016 have been required to book later meetings in an online booking system (except the first meeting, which is booked by the assigned caseworker).[16] One might worry that this increased the amount of caseworker switching, since some jobcenters allow the jobseeker to choose between multiple caseworkers when booking the meetings. However, Rasmussen (2021) documents that the jobcenters generally seek to minimize caseworker switching and that caseworker switching has not increased after the implementation of self-booking. In section 4.4 of this paper, we provide evidence on the amount of caseworker switching over a typical UI-spell. We show that jobseekers tend to stay with their assigned caseworker.

The aim of the caseworker meetings is to support the jobseeker in the transition into regular employment. This includes discussions about concrete jobs and job search strategies, as well as assignments to activation programs.[17] Already at the first meeting, the caseworker must prepare a *job plan* for the jobseeker. The job plan specifies activation programs, such as vocational training courses or specific job training, that the caseworker assigns the jobseeker to. The caseworker must update the job plan continuously as long as the jobseeker is unemployed.[18] Importantly, in order to continue to claim UI-benefits, the jobseeker *must* comply with the caseworker's assignments to activation programs.[19] If the jobseeker does not comply, e.g. fail to participate in the assigned activation program, she will be deprived her right to receive UI-benefits until she complies with the assignment. In Section 4.4, we provide evidence on the extent to which assignments (registrations in the job plan) result in enrollments.

3.2 Training for Jobseekers

Caseworkers have two classes of interventions at their disposal when designing an activation plan for the jobseeker: Course training and job training. Vocational training is the most common type of course training.[20] Other course training includes elementary courses, such as lan-

[16] Lov nr. 548 af 07/05/2019, §34.

[17] Lov nr 548 af 07/05/2019, §28-29.

[18] The job plan ("Min plan", in Danish) can be accessed on www.jobnet.dk (Lov nr 548 af 07/05/2019, §41-42).

[19] The law states that the jobseeker has the right and is obliged to participate in assigned activation programs (Lov nr. 548 af 07/05/2019, §100-103).

[20] Section 4.4 provides summary statistics on the volume and timing of these caseworker interventions. See Tables 2 and 3 in particular.

guage classes, or more advanced coursework, like a university class. Job training reflects wage subsidies and on-the-job training. As job training is not the focus of of this paper, we leave out further details.

Vocational training in Denmark consists of a more than 3,000 short courses which are often tailored to a specific occupation or industry. Training takes place at dedicated facilities, and is organized in classrooms or as open workshops. The goal of the training program is to help low and mid skilled workers keep their skill sets up to speed with new technologies or other recent changes in the labor market.[21] Table 1 shows the vocational training courses that were most frequently assigned by caseworkers in 2011.

Table 1: Popular Vocational Training Courses

Course Title	Duration in days	Price in USD	Share of activity in percent
Forklift certificate	7.5	1,064	4.2
Bus driver license	30	5,641	3.6
Truck driver license	30	5,641	3.2
Taxi driver license	30	3,759	3.2
Security guard	15	1,302	2.5

Note: The table shows the five training courses that most frequently got assigned by caseworkers in 2011 (weighted by duration of the course). Share of activity is calculated as the share in total training days by jobseekers that were assigned by caseworkers. Note that jobseekers do not pay for the courses themselves. The price refers to the cost for the municipality.

4 Data

In our analysis, we use the data sample defined and constructed in Rasmussen (2021). This sample is based on a novel caseworker data set and consists of jobseekers that were assigned to caseworkers based on their birthday from 2011 to 2018. We first briefly introduce the data and sample construction, while referring to Rasmussen (2021) for additional details. We then describe how we complement the sample with information on training assignments, jobseeker characteristics, and labor market outcomes. Finally, we present descriptive statistics for our analysis sample.

[21]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).

4.1 Unemployment Spells and Caseworker Meetings

Our analysis relies on a novel *caseworker data set* that originates from a legal requirement on jobcenters in Denmark. Since 2010, the jobcenters have been required to register all *meetings* held with unemployed jobseekers.[22] The meeting registrations are continuously collected by the Agency for Labor Market and Recruitment and are used for monitoring of the meeting activity. This has given rise to a data set containing information on date, location, time, type, and contact form of each meeting, as well as personal identifiers for jobseekers and caseworkers. While the jobseeker identifiers can be linked to Danish administrative data, the caseworker identifiers cannot. Given the novelty of the caseworker data - and its importance for our analysis - it is worth emphasizing that the data quality is validated in Rasmussen (2021). She provides clear evidence of the authenticity of the meeting registrations and the recorded caseworker identifiers.

To home in on the quasi-random assignment of caseworkers, we apply a set of sample restrictions. We outline these restrictions here, and refer to Appendix Table A. 1 for an overview of the sample reduction at each step. The starting point for our sample is all meetings registered to take place in Danish jobcenters between caseworkers and recipients of UI benefits from 2011 to 2018. By linking to a register on UI benefit payments (DREAM), we can construct UI-spells for all jobseekers in the sample. In the construction of UI-spells, we follow the convention by defining a UI-spell as consecutive weeks with UI-benefit payments while allowing for interruptions in payments of up to three weeks (Fluchtmann et al., 2020).[23] Thereby, all meetings are ordered relative to an entry date into unemployment.

Based on the institutional details laid out in section 3.1, we identify the *caseworker assigned* to a given jobseeker as the caseworker that attend the jobseeker's first individual meeting.[24] We make a couple of restrictions to ensure that we actually observe this meeting for all jobseekers in our sample.[25] First, we require that the first observed meeting for a given jobseeker takes place within three months of UI-spell start, which was required by law. Second, we drop UI-spells if they were initiated within 10 months from a previous UI-spell. Third, we require that the first individual meeting is coded as a 'regular meeting' that takes place 'in person'. Note that we

[22]BEK nr 418 af 23/04/2010, kap. 5

[23]We define an exit from unemployment as 4 consecutive weeks with UI-benefit payments. Likewise, an entry into unemployment is defined as the first time in which the individual receives UI-benefits after at least 4 consecutive weeks with no benefits. The UI-spell length is the number at weeks between an entry and exit week.

[24]We use the first meeting in the UI-spell that was not coded as an 'information meeting'.

[25]There could be different reasons why this might not be the case. E.g. if the caseworkers does not register the first meeting, or if the UI-spell we have identified actually is a continuation of a previous UI-spell.

drop the few jobseekers for whom we are unable to identify their assigned caseworker.[26]

Based on a survey conducted by and described in Rasmussen (2021), we restrict the sample to jobcenters and years in which the jobcenters *report* that they base caseworker assignments on birthdays. To further support our identification strategy, we restrict the sample to caseworkers for whom jobseeker birthdays can actually predict assignment.[27] Finally, to reduce noise in our instrument, we restrict the sample to caseworkers with at least 50 assigned jobseekers and who worked for at least one quarter. Since assignment to caseworkers occur within a jobcenter unit and year, we will include fully interacted jobcenter-unit-year fixed effects in all regressions. We therefore also require that at least two caseworkers are present in each fixed effect cell. Our final sample consist of 103,027 UI-spells spread over 75,811 jobseekers, 467 caseworkers, and 24 jobcenters.[28]

4.2 Training Assignments

To obtain information on the intervention of interest, *assignment to training*, we rely on the individual job plans that caseworkers prepare for each jobseeker. As described in Section 3.1, these job plans list activities, e.g. courses and job training, that the caseworker assigns the jobseeker to. Together with the meeting registrations, the job plan registrations are collected by the Agency for Labor Market and Recruitment. This has given rise to a comprehensive data set of all activities recorded in individual job plans from 2011 to 2018, including the start date and type of activities. *Activities* refer to job or course training, while *types* are subcategories for an activity. For example, two types of *course training* activities are 'vocational training courses' and 'language courses'.

Based on jobseeker identifiers and activity *start dates*, we can link all job plan activities to the UI-spells in our baseline sample. We then define the individual as being *assigned* to a given activity if that activity is set to start at any point during the UI-spell.[29] It should be noted that

[26] For example, if they exit unemployment before their first individual meeting with a caseworker.

[27] This sample restriction is based on a validation exercise in Rasmussen (2021), where assignment to a caseworker is predicted first based on the jobseekers' actual birthdays and hereafter based on placebo birthdays. The idea is to include only the caseworkers for whom jobseeker birthdays make a 'good prediction'. However, it is not obvious what the criterion for a 'good' prediction is, and choosing the criterion involves a trade-off between precision and power. In the end, we restrict the sample to all caseworkers for whom the prediction based on true birthdays is better than the median prediction based on placebo birthdays. We leave it for future work to test the robustness of the results to this criterion

[28] We note that any of the restrictions potentially could affect the results. We save it for future work to test the robustness of results to the sample restrictions.

[29] We require that the activity has start date up to one week before or after the UI-spell start and end date, respec-

although we include all activities set to start during the UI-spell, 99% of assigned jobseekers have their first activity starting *after* the first caseworker meeting in the jobcenter. Section 4.4 presents descriptive statistics for the different types of training assignments in our sample.

4.3 Jobseeker Characteristics and Labor Market Outcomes

We use the Danish administrative registers to enrich our sample with labor market outcomes. Since we aim at conducting a cost-benefit analysis of vocational training interventions in future versions of the paper, we will focus on outcomes that are relevant inputs to a cost-benefit analysis. A key component on the cost side is *UI benefit payments*. As discussed in Section 2, assignment to training could prolong unemployment due to locking-in effects. By linking to a register of government transfers (DREAM), we can see whether individuals receive UI benefits in a given week.[30]

On the benefit side, we draw on the E-Income Register (EINDKOMST) to obtain monthly wage earnings and working hours for all employees in Denmark. We measure the extensive margin of employment by constructing an *employment* dummy that takes a value of one if the individual had positive wage income in the given month.

We enrich the baseline sample with information on predetermined characteristics for all jobseekers. Through the Population Register (BEF), we obtain information on *demographics* such as gender, age, origin, marriage status, and number of children. We obtain information on the highest completed *education* from the Education Register (UDDA).[31] Finally, we obtain information on *labor market history* from the DREAM and EINDKOMST registers. From the transfer register (DREAM), we can see whether the individual received UI benefits, education subsidies, parental leave subsidies, UI-fund members, or other types of public transfers in the past. Based on the income register (EINDKOMST), we obtain information on wages, hours worked, and industry of past employment.

---

tively. This is to account for the fact that UI-spell start and end dates are constructed based on weekly information on benefit payments. We note that the extra week on either side of the UI-spell in practice is unimportant: 0.13 percent of UI-spells have the first activity starting before or after the UI-spell start and end date, respectively.

[30] In further work, we plan to also consider course expenditures, which is another important element on the cost side. We can study these course expenditures by combining the Course Participant Register (VEUV) with the course price catalogues from the Finance Acts.

[31] We use the highest education completed up to one month after UI-spell start. This is to ensure that we see the relevant education for graduates who may not have received their diploma yet.

4.4 Descriptive Statistics

In this subsection, we present descriptive statistics for training assignments, jobseekers, and caseworkers in our sample. First, we show the volume and timing of training activities. Second, we relate the timing of the activities to the likelihood of staying with the assigned caseworker over the unemployment spell. This is a first step to assess the relevance of our instrument. Third, we show the types of training assignments and how they relate to later course enrollments. Fourth, we present summary statistics on labor market outcomes for jobseekers which highlight the observable differences between jobseekers assigned to course training and those not assigned to any activity. These observable differences point to the identification challenge that we address in this paper.

Volume and Timing of Training Assignments

Table 2 summarizes the magnitude of training assignments in our sample. Panel A shows that 43.5 percent of all UI-spells are assigned to some type of training. Considering spells of at least 26 weeks, the probability of being assigned to some training increases to 76.7 percent. As revealed by Appendix Figure A. 1, this reflects a positive correlation between the probability to be assigned to training and time spent in unemployment. In Panel B, training assignments are split into three mutually exclusive categories; course training only, job training only, and assignment to both activities. Including those assigned to both activities, it shows that around a third of all spells in the sample are assigned to course training. In comparison, only around a fifth of all spells are assigned to job training.

Table 3 breaks course training into nine types of courses. The two largest types are 'Courses n.e.c.' and 'Vocational training courses'. Around 27 percent of all spells are assigned to the former and rather non-specific type of courses, while 7.6 percent of all spells are assigned to the latter type. The table also shows the timing of the different course types. In particular, it shows in what week of the UI-spell a given course type is set to start. For the two largest types, the mean (median) course starts in week 18.5 (13) and 21.9 (16) of the UI-spell. Although the standard deviations are large and hence suggest much heterogeneity in course start, the timing of the activities is interesting. Namely, it puts an upper bound on when the jobseekers on average are *assigned* to the different activities, since we expect the assignment to take place in weeks *prior* to the start date of the activity. For our instrument to be relevant, it must be the initial caseworker that assigns the jobseeker to training. In the next step, we therefore assess the extent to which jobseekers stay with their initial caseworker over the unemployment spell.

Table 2: Training Assignments in the Job Plan

	All spells		Spells $\geq$ 26 weeks		Spells $\geq$ 52 weeks	
	obs	pct	obs	pct	obs	pct
A: Assignment						
No assignment	58,205	56.5	8,872	23.3	1,011	6.9
At least one assignment	44,822	43.5	29,258	76.7	13,546	93.1
B: Types of assignments						
Course training only	22,482	21.8	12,481	32.7	4,684	32.2
Job training only	9,791	9.5	6,052	15.9	2,329	16.0
Course and job training	12,549	12.2	10,725	28.1	6,533	44.9
Obs	103,027		38,130		14,557	

Note: The table summarizes all activities registered in job plans for jobseekers in the sample. Activities noted in the job plan during a jobseeker's UI-spell are denoted *assignments*, and they are shown in levels as well as in percent. Column 1-2 show assignments for all spells, column 3-4 for spells of at least 26 weeks and column 5-6 for spells of at least 52 weeks. In panel A, spells are split into two mutually exclusive categories; spells with no assignments and spells with at least one assignment. In panel B, assigned spells are split into three mutually exclusive categories; spells assigned to course training only, spells assigned to job training only and spells assigned to course as well as job training.

Table 3: Course Training Types and Timing

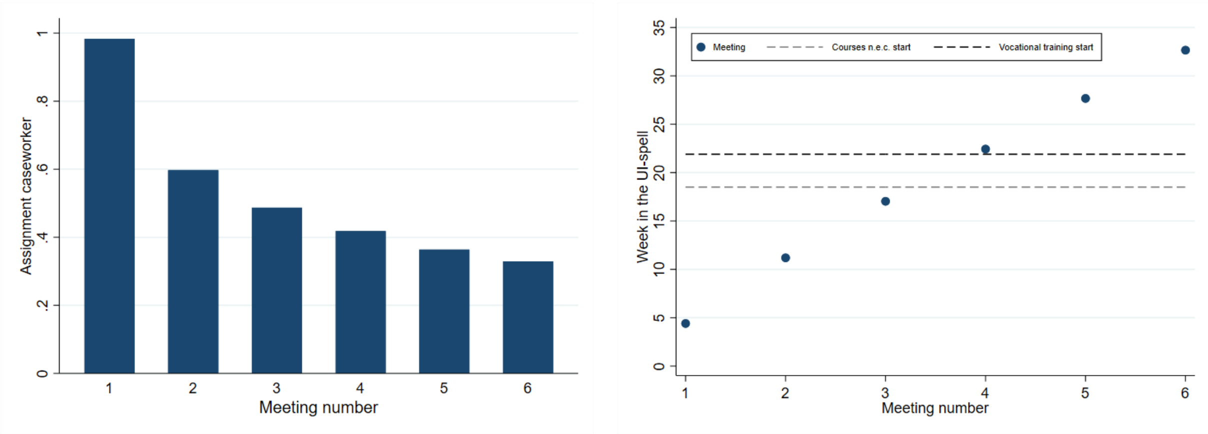
	Assignment rate		Timing of course in UI-spell						
	obs	pct	mean	sd	p10	p25	p50	p75	p90
1. Courses n.e.c.[(1)]	27,920	27.1	18.5	16.7	4.0	7.0	13.0	24.0	41.0
2. Vocational training	7,825	7.6	21.9	19.8	5.0	9.0	16.0	28.0	48.0
3. Other measure	1,728	1.7	23.6	20.3	5.0	10.0	17.0	30.0	54.0
4. Soc. health care and vocational educ.	1052	1.0	25.9	21.1	5.0	11.0	20.0	35.8	56.0
5. General adult education	875	0.8	23.3	19.9	5.0	9.0	17.0	32.8	51.0
6. Private courses	631	0.6	26.7	22.9	6.0	12.0	19.0	34.2	61.8
7. Higher education	493	0.5	26.4	21.0	6.0	11.4	19.0	36.6	56.6
8. Language training	488	0.5	18.0	17.2	4.0	7.0	12.0	22.6	46.4
9. Residual courses[(2)]	608	0.6	24.8	23.4	4.2	8.0	15.9	35.4	62.0

Note: The table summarizes the types of courses registered in the job plan (own translation from Danish) for jobseekers in the sample. Column 1 and 2 shows the assignment rate for a given course type, in levels and in pct respectively. I.e. a given row in column 1 (2) shows the number (percent) of jobseekers assigned to that course type during their spell. Note that a jobseeker may be assigned to multiple course types during the same UI-spell. Hence the same jobseeker may appear in multiple rows, which explains why the sum of the rows do not sum to the total number of jobseekers assigned to at least one course (22,482+12,549 = 35,031) shown in table 2. Column 3-9 summarize the timing of courses relative to UI-spell start. In particular, in what week of the spell a given course *starts*, not when it was assigned. Column 3 shows the average start week for a given course. Column 4 shows the SD on the start week. Column 5-9 reports given percentiles. To comply with data protection rules, all percentiles are based on at least 5 observations.
[(1)] Courses not elsewhere classified
[(2)] Residual category represents all other course types, e.g. reading, writing and math courses.

Survival with the Assigned Caseworker

Figure 1a plots the fraction of jobseekers that have a given meeting in the UI-spell with the initially assigned caseworker (conditional on staying unemployed up to that meeting). Since the vast majority of the jobseekers in the sample meet with their assigned caseworker in the first meeting of the UI-spell, the probability is close to 100 percent for the first meeting. As the figure shows, the probability of having a given meeting with the initially assigned caseworker decreases with time spent in unemployment. That said, caseworker survival rates are relatively high. For example, the probability of meeting the initial caseworker in the second meeting is 60 percent. In comparison, if jobseekers were randomly assigned between caseworkers within a jobcenter unit from one meeting to the next, this probability would have been around 12.5 percent.[32]

Figure 1: Survival with the Assigned Caseworker



(a) Caseworker survival (b) Timing of meeting and training

Note: Panel 1a shows the fraction of jobseekers that have a given meeting in the UI-spell with their assigned caseworker (conditional on having the given meeting). The fraction is not one for the first meeting, since 2% of jobseekers start with an information meeting and hence meet their assigned caseworker in the second meeting. Panel 1b shows the average timing of a meeting (conditional on having the meeting) relative to the UI-spell start. The dashed black (gray) line shows the mean timing of vocational training courses (courses n.e.c) in the UI-spell (see table 3). Note that this week refers to the start of the course, not the week of assignment.

Figure 1b plots the mean timing of meetings in the UI-spell (blue dots) as well as the mean timing of the two largest course types: Vocational training (black line) and Courses n.e.c. (gray line). As the jobseeker must be assigned *prior* to starting in a course, this figure suggests that the average jobseeker is assigned *before* the fourth meeting. This is interesting, as figure 1a shows that the probability to meet with the assigned caseworker is around 40 percent in the fourth meeting and 50 percent in the third meeting. This suggests that the assigned caseworker indeed *is* relevant for assignments to course training.

Training Assignment vs. Enrollment

To assess the link between training assignments and course enrollments, we link all jobseekers to the Course Participant Register (VEUV), which contains information on all courses that an

[32]The average number of caseworkers serving the same jobcenter unit and year cell is 8. Were the jobseeker randomly (and independently of previous meetings) assigned a caseworker from the same jobcenter unit and year cell, the probability of meeting the same caseworker would have been 1/8 (12.5 percent).

individual ever enrolls in, as well as the educational classification of these courses. This allows us to investigate, first, how many assignments result in actual enrollments. Second, it allows us to assess the course contents of the different assigned courses. Table 4 tabulates course assignments against course enrollments. The first takeaway from table is that assignment to a course does not necessarily lead to subsequent enrollment and participation. In particular, for jobseekers assigned to 'courses n.e.c.', less than half end up enrolling in some training course during the course of their UI-spell. The enrollment rate is much higher for jobseekers assigned to 'vocational training courses'. Among these jobseekers, more than 80 percent also enroll in a course during their UI-spell. Note that the fact that some jobseekers are assigned but never enrol in the training activity need *not* imply that they did not comply with the training assignment. Rather it may simply reflect that some jobseekers leave unemployment after being assigned but before enrolling in an activity, potentially due to the threat effect discussed in section 2.

Second, the table shows that the assigned course types are well in accordance with the courses that jobseekers actually enroll in. For instance, it is reassuring to see that, conditional on enrollment, more than 80 percent of the jobseekers *assigned* to vocational training actually *enroll* in courses classified as vocational education. This tabulation also sheds some light on the largest, and rather non-specific, types of courses that jobseekers are assigned to ('course n.e.c.'). Namely, conditional on enrollment, 65 percent of the jobseekers assigned to 'courses n.e.c.' enroll in courses classified as vocational. This suggests that most assignments in the 'course n.e.c.' category actually reflects vocational training courses. That said, it is also clear that substantial shares of the 'course n.e.c.' assignments fall within the primary, secondary, and post-secondary educational levels. As these education types are substantively different from vocational training courses, both in terms of duration, costs and job targets, we should expect them to have very different locking-in and post program effects.

Motivated by this, we will focus on two treatments in our analysis. We will first investigate the effect of assignment to *any course*. Second, we will zoom in on the assignment to *vocational training courses* in particular.[33] Appendix Table A. 3 summarizes how these treatments differ in terms of enrollment rates and education enrollment types. Out of all individuals assigned to *any course*, around a half actually end up enrolling in some course, and close to 70 percent of the

[33]To investigate the former, we instrument *any courses* with the caseworker's propensity to assign jobseekers to any course. To investigate the latter, we adapt our instrument. In particular, we instrument *vocational training* with the caseworker's propensity to assign jobseekers to vocational training.

enrollments are in vocational training. In comparison, for individuals assigned to *vocational training*, more than 80 percent enroll in some course and the vast majority of these (also 80 percent) enroll in vocational education.

Table 4: Assignment vs. Enrollment

	Enrolment		Enrolment education types		
	No	Yes	Prim./sec.	Vocational	Post sec.
Panel A: Levels					
1. Courses n.e.c.	16,409	11,511	2,498	8,302	2,024
2. Vocational training	1,378	6,447	684	5,906	620
3. Other measure	1,021	707	218	438	133
4. Soc. health care, vocational educ.	421	631	107	530	83
5. General adult education	148	727	429	410	70
6. Private courses	383	248	26	149	99
7. Higher education	194	299	25	114	198
8. Language training	165	323	294	-	-
9. Residual courses	216	392	187	275	23
Panel B: Percentages					
1. Courses n.e.c.	59	41	19	65	16
2. Vocational training	18	82	9	82	9
3. Other measure	59	41	28	56	17
4. Soc. health care, vocational educ.	40	60	15	74	12
5. General adult education	17	83	47	45	8
6. Private courses	61	39	9	54	36
7. Higher education	39	61	7	34	59
8. Language training	34	66	72	-	-
9. Residual courses	36	64	39	57	5

Note: The table relates course *assignments* (based on registrations in individual job plans to) course *enrollments* (based on the Danish course participant register, VEUV). All courses enrolled in during the UI-spell are included. In panel A, the first two columns show how many jobseekers with a given course assignment that actually end up enrolling in a course. The next three columns show what type of courses these jobseekers enrol in. There are overall three types: Primary and secondary education (column 3), Vocational training (column 4) and Post secondary education (column 5). Since jobseekers can enrol in *multiple* courses of different types, the enrolment types (column 3-5) do not sum to the number of jobseekers that enrol in *at least one* course (column 2). In panel B, the first two columns show the percent of jobseekers with a given course type assignment that actually end up enrolling in a course. The next three columns shows the percent of jobseekers that enrol in a given course type (conditional on enrollment). To comply with data protection rules, cells with less than 4 observations are not shown.

Jobseeker Characteristics and Labor Market Outcomes

Appendix Table A. 2 presents summary statistics for the jobseekers in our sample. Columns 1 and 2 present statistics for all jobseekers, and Columns 3-4 and 5-6 focus, respectively, on jobseekers assigned to course training and jobseekers not assigned to any activity at all.[34] Column

[34]The object of interest in this paper is the causal effect of assignment to *course training* as opposed to no training assignment. We therefore focus our descriptive analysis on selection into course training in particular. Note however, that jobseekers assigned to job training are included in column 1 and 2.

7 highlights differences between these two groups of jobseekers. Jobseekers assigned to course training are less likely to hold a vocational education or a bachelor's degree. This suggest that course training participants primarily includes low skilled workers who are looking to upskill through vocational training. Although jobseekers that are assigned to course training were no more likely to be unemployed in the past, their labor market attachments tend to be weaker with lower past rates of employment and wages. Further, the assigned jobseekers are more likely to have worked in trade and transportation services and less likely to have worked in public administration. These stark differences in observable characteristics highlight the key challenge for identifying the causal effects of course training: jobseekers who get assigned to training are a selected group of individuals with weaker labor market attachments.

Table 5 shows labor market outcomes one and two years after unemployment start. The last column is particularly interesting as it shows the difference in means between jobseekers assigned to course training and jobseekers not assigned to any activity at all. It is evident that jobseekers assigned to course training tend to do worse than jobseekers not assigned to any activity at all. One year after unemployment start, assigned jobseekers earn 5,000 DKK less, work fewer hours, and are more likely to remain unemployed. Two years after unemployment start, the jobseekers assigned to course training still do worse than those not assigned to any activity. Appendix Table A. 4 shows that jobseekers assigned to course training not only do worse *on average*; the entire wage distribution is shifted downward relative to jobseekers that do not get assigned to any activity.

Although these observable differences in labor market outcomes are stark, they could just reflect an adverse selection of jobseekers into course training. To investigate the *causal* effects of being assigned to course training, we adopt an instrumental variable (IV) strategy that exploits that caseworkers, who differ in their propensities to use course training, are quasi-randomly assigned to individual jobseekers.

In order to precisely estimate these caseworker propensities, we must observe a sufficient number of jobseekers being assigned to each caseworker. Appendix Table A. 5 shows that the mean (median) caseworker in our sample is assigned 220 (152) jobseekers over a period of approximately 1.5 years. We note that these figures are in the range of the average caseload sizes for judges in Bhuller et al. (2020) who applies a similar identification strategy.[35]

[35] In Bhuller et al. (2020), the average judge is assigned 258 court cases.

Table 5: Labor Market Outcomes by Course Assignment

	Full sample		Course training		No assignment		Diff in means
	mean	sd	mean	sd	mean	sd	col 3 - 5
A: Outcomes in month 12							
Wage	15,164.4	(14,159.1)	12,185.4	(13,771.7)	17,200.5	(14,282.6)	-5015.1***
Hours	82.4	(70.2)	67.9	(70.9)	91.1	(68.3)	-23.2***
Employed	0.7	(0.5)	0.6	(0.5)	0.7	(0.4)	-0.2***
UI-benefits	0.3	(0.5)	0.4	(0.5)	0.2	(0.4)	0.2***
B: Outcomes in month 24							
Wage	17,149.0	(14,709.1)	15,346.0	(14,485.7)	18,485.5	(14,885.2)	-3139.5***
Hours	91.2	(69.9)	84.2	(71.6)	95.7	(68.3)	-11.5***
Employed	0.7	(0.5)	0.6	(0.5)	0.7	(0.4)	-0.1***
UI-benefits	0.2	(0.4)	0.2	(0.4)	0.2	(0.4)	0.1***
Obs in month 12	103,027		35,031		58,205		93,236
Obs in month 24	90,959		30,105		52,059		82,164

Note: In panel A, the dependent variables are measured in month 12 after UI-spell start, whereas the dependent variables in panel B are measured in month 24 after UI-spell start. *Wages* and *hours* refer to labor earnings measured in DKK and hours worked in a given month (with non-employment coded as zero hours). *Employed* and *UI benefits* are dummies taking value one if the individual is employed or receives UI benefits. Column 1-2 show the mean and SD for the full sample. Column 3-4 show mean and SD for spells with a least one course assignment (including spells with a course assignment only *and* spells with a course as well as a job training assignment). Column 5-6 show the mean and SD for spells with no assignment (to job nor to course training). Column 7 shows the difference in means for spells assigned to course training relative to spells on passive UI. *p<0.10 ** p<0.05 *** p<0.01.

5 Empirical Strategy

In this section, we present our empirical strategy. We first discuss the challenge to identification, and then turn to describe a novel research design that allows us to overcome this challenge. Similarly to Bhuller et al. (2020), we employ an instrumental variables (IV) strategy with identifying assumptions that we discuss by the end of the section.

The object of interest in our analysis is the causal effect of assigning jobseekers to course training on their subsequent labor market outcomes. In particular, we are interested in the causal effect of assigning jobseekers to course training *as opposed to* not assigning them to any type of training. Given that a jobseeker can be assigned to course training, assigned to job training, or not assigned at all, our structural equation of interest reads as follows

$$Y_{i,t} = \beta_t D_i^{course} + \delta_t D_i^{job} + X_i' \theta_t + u_{i,t} \tag{1}$$

In the regression, $Y_{i,t}$ represents the labor market outcome of jobseeker i in period t relative to UI-spell start, and X_i is a vector of control variables. D_i^{course} and D_i^{job} are indicators for whether the jobseeker is assigned to course or job training at some point during her UI-spell. Hence, the parameter of interest in the regression is β_t.

The main challenge to identification of β_t is that jobseekers with certain labor market prospects opt into course training. Note that since treatment in this paper is *assignment to* *training*, the selection could in principle occur both on the jobseeker and caseworker side. Caseworkers could be assigning jobseekers with certain labor market prospects to training, however, jobseekers with certain prospects could also lobby for training. For simplicity, we refer to the identification challenge in the classical way; as selection on the jobseeker side. In Section 4.4, we presented clear evidence of such non-random selection.

As noted in a handbook chapter by McCall et al. (2016), the vast majority of studies in the training evaluation literature rely on a 'selection on observables' strategy to solve this identification challenge. However, there is no reason to believe that selection is based solely on observables. Jobseekers likely also select into training based on *un*observables $u_{i,t}$, e.g. future job opportunities, in which case Ordinary Least Squares (OLS) estimation of Equation (1) will yield a biased estimate of β_t.

We use a novel research design to overcome this challenge to identification. In particular, we exploit i) that caseworkers within a jobcenter unit and year are assigned to jobseekers based on jobseeker birthdays, and ii) that caseworkers vary in their propensities to assign jobseekers to course training. As noted earlier, Rasmussen (2021) validates that caseworkers in our sample are assigned based on birthday of the month (1-31), and further that birthdays are as-if random. These findings support the assumption that caseworkers effectively are *quasi-randomly* assigned to jobseekers. The quasi-random assignment of caseworkers combined with ii) provides us with *exogenous variation* in the probability that a jobseeker is assigned to course training.

We exploit this in an instrumental variables (IV) strategy to identify the causal effect of course training. In particular, we instrument the assignment to course training for jobseeker i with her caseworker's overall propensity to assign jobseekers to course training. We measure the caseworker's assignment propensity as a leave-out mean, i.e. as the mean assignment rate for all other jobseekers that were also assigned to the caseworker (except i).[36] We then use a two-stage least squares (2SLS) set-up to estimate β_t in Equation (1), where we account for the fact that jobseekers may not only be assigned to course training but also to job training. That is, besides instrumenting assignment to course training, we also instrument assignment to job training.[37] This means that we have two first stage equations as specified below.

$$D_i^{course} = \alpha_1 Z_{j(i)}^{course} + \alpha_2 Z_{j(i)}^{job} + X_i'\gamma + \eta_i \tag{2}$$

[36] As in Bhuller et al. (2020), this includes former and future jobseekers assigned to the caseworker.

[37] Similarly to course training, we instrument job training with the caseworker's propensity to assign jobseekers to job training. We measure this as a leave-out mean.

$$D_i^{job} = \mu_1 Z_{j(i)}^{course} + \mu_2 Z_{j(i)}^{job} + X_i' \omega + \epsilon_i \tag{3}$$

Here, $Z_{j(i)}^{course}$ and $Z_{j(i)}^{job}$ represent caseworker j's assignment propensities for course and job training (leave-out means). Importantly, X_i includes a set of controls that ensures quasi-random assignment of jobseekers to caseworkers. As the birthday-based assignment of caseworkers occurs within a jobcenter unit and year, we include fully interacted jobcenter $\times$ unit $\times$ year fixed effects. Further, as shown in Rasmussen (2021), birthdays are as-if random *conditional* on the country of origin of the jobseeker, and we therefore control for non-western origin in all regressions.[38] We also include quarter fixed effects to control for potential business cycles. Finally, since we cannot rule out that some jobcenters might use birth years for assignment, corresponding to age-based assignment, we include age fixed effects.

We interpret the IV-estimate as in Bhuller et al. (2020). Given instrument exogeneity, relevance, exclusion, and monotonicity, β_t^{IV} can be interpreted as a local average treatment effect (LATE) of assignment to course training. In particular, we can interpret it as the causal effect of assignment to course training for jobseekers who would have received a different course training assignment decision, had they been assigned to a different caseworker in the same jobcenter unit and year.

Instrument *exogeneity* is satisfied if caseworkers are quasi-randomly assigned to jobseekers. In Section 6.1, we show that the course training assignment propensity of the caseworker indeed cannot be predicted by job seeker characteristics, which supports the exogeneity assumption. In Section 6.2, we show that the instrument is indeed also *relevant*. Namely, we can predict jobseeker i's assignment to course training with the caseworker's general course assignment propensity (leave-out mean). *Monotonicity* is needed in case of heterogeneous effects across jobseekers. It implies that jobseekers assigned to course training by a lenient caseworker (low assignment propensity) would also need to be assigned to course training by a strict caseworker (high assignment propensity). Likewise, jobseekers *not* assigned by a strict caseworker would also need *not* to be assigned by a lenient caseworker. We perform two tests in Section 6.3 that strongly support this assumption. Finally, the *exclusion restriction* allows us to interpret LATE as the causal effect of assignment to course training. It requires that the instrument only affects outcomes through assignment to course training. An obvious threat in this regard is that

[38]Immigrants who lack a birth certificate when they arrive to Denmark are assigned a birthday of January 1 by the immigration authorities. This predominantly occurs for non-western immigrants, who are therefore over-represented on January 1st.

caseworkers have multiple tasks. For example, besides assigning jobseekers to course training, caseworkers also meet with jobseekers to discuss job search strategies. Hence, if caseworker assignment propensities correlate with their meeting frequency, it could violate the exclusion restriction. In Section 7.1, we present evidence that clearly supports the exclusion restriction.

6 Assessing the Caseworker Leniency Instrument

In this section, we assess the validity of the caseworker leniency instrument. We first show evidence that clearly suggests that caseworkers are quasi-randomly assigned to jobseekers. Hereafter, we show that our instrument is highly relevant and that monotonicity likely is satisfied.

6.1 Conditional Independence

Table 6 presents a randomization test that clearly suggests that caseworkers are quasi-randomly assigned to jobseekers within a jobcenter unit and year. The idea in this table is to show the selection problem with the endogenous treatment variable, and then assess whether the instrumental variable solves the problem. The first column in the table shows the coefficients from a regression of the *endogenous* treatment variable, assignment to course training, on jobseeker predetermined characteristics. It shows that the majority of predetermined characteristics strongly predict whether a jobseeker is assigned to course training. This again points to the non-random selection of jobseekers into course training. The third column in the table shows the coefficients from a regression of the *instrumental* variable, caseworker course assignment propensity, on the same set of predetermined characteristics. It is evident that we can*not* predict the caseworker's course assignment propensity with the jobseeker's predetermined characteristics. Individually, the vast majority of coefficients on the predetermined characteristics are small and statistically insignificant. The F-stat and corresponding p-value in the bottom of the table show that the predetermined characteristics jointly are insignificant as well. We take these findings as supportive evidence that caseworker assignments are quasi-random and that our instrument thus satisfied the exogeneity assumption.

Table 6: Randomization Test

	Pr(Course Training)		Course Instrument		Covariates	
	b	se	b	se	mean	sd
Demographics(1)						
Male	0.002	(0.008)	0.002	(0.004)	0.484	0.500
Immigrant	0.143***	(0.016)	0.009	(0.009)	0.050	0.218
Descendant	0.094	(0.074)	-0.059	(0.040)	0.002	0.043
Married	-0.041***	(0.008)	-0.007*	(0.004)	0.422	0.494
Number of children	-0.004	(0.003)	0.001	(0.002)	0.806	1.039
Education(2)						
0. Missing	0.008	(0.029)	-0.014	(0.013)	0.013	0.115
15. Preparatory course	0.128**	(0.055)	0.031	(0.028)	0.004	0.062
20. Upper secondary	0.018	(0.015)	-0.016**	(0.008)	0.047	0.212
30. Vocationa educ.	-0.043***	(0.009)	-0.002	(0.005)	0.476	0.499
35. Qualifying program	-0.038	(0.132)	-0.089	(0.071)	0.000	0.022
40. Short cycle tertiary	0.063***	(0.017)	0.016	(0.010)	0.047	0.211
50. Vocational bach.	-0.040***	(0.013)	0.004	(0.007)	0.125	0.331
60. Bachelor	0.074***	(0.028)	-0.006	(0.017)	0.014	0.118
70. Master	0.036*	(0.020)	0.012	(0.014)	0.049	0.215
80. PhD	0.030	(0.058)	0.008	(0.053)	0.003	0.053
Labor market history(3)						
UI-benefits in year t-1	-0.072***	(0.010)	-0.004	(0.007)	0.408	0.491
UI-benefits in year t-2	-0.089***	(0.009)	-0.001	(0.007)	0.424	0.494
Any employment in year t-1	-0.128***	(0.017)	0.009	(0.009)	0.900	0.300
Any employment in year t-2	0.000	(0.015)	0.006	(0.008)	0.922	0.268
Employment rate in year t-1	-0.158***	(0.026)	-0.021	(0.015)	0.642	0.352
Employment rate in year t-2	0.058***	(0.020)	-0.016	(0.014)	0.659	0.348
Wage earnings in 1,000 DKK in year t-1	-0.000**	(0.000)	0.000	(0.000)	212.235	147.811
Wage earnings in 1,000 DKK in year t-2	-0.000***	(0.000)	-0.000	(0.000)	209.295	141.285
Number of employers in year t-1	-0.038***	(0.004)	0.001	(0.002)	1.398	0.944
Number of employers in year t-2	-0.005	(0.004)	-0.002	(0.002)	1.407	0.907
Public transfers in year t-1	0.048***	(0.009)	-0.007	(0.007)	0.627	0.484
Parental leave in year t-1	-0.008	(0.012)	-0.002	(0.007)	0.083	0.276
Education subsidy in year t-1	-0.160***	(0.014)	-0.005	(0.009)	0.101	0.302
Previous industry(4)						
Real estate	0.127***	(0.030)	0.019	(0.016)	0.011	0.106
Business services	0.138***	(0.013)	0.003	(0.007)	0.115	0.320
Finance	0.272***	(0.035)	0.032*	(0.017)	0.010	0.098
Trade & transport	0.144***	(0.013)	0.010	(0.006)	0.199	0.400
Manufacturing	0.193***	(0.014)	-0.010	(0.013)	0.123	0.328
Communication & it	0.183***	(0.027)	-0.005	(0.019)	0.017	0.130
Culture	0.045**	(0.019)	-0.000	(0.009)	0.036	0.185
Agriculture, forestry & fishing	0.062***	(0.021)	-0.001	(0.010)	0.026	0.159
Public administration, health, education	0.044***	(0.012)	0.004	(0.007)	0.224	0.417
Obs	103027		103027		103027	
Dep var Mean	-0.000		-0.000			
Dep var sd	1.000		1.000			
Number of FE's	677		677			
F-stat	43.679		1.121			
P-value	0.000		0.291			

Note: *p<0.10 ** p<0.05 *** p<0.01. The first four columns two report coefficients and standard errors from a regression including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. In column 1, the dependent variable is a dummy capturing whether the individual is assigned to course training, while it is the leave-out mean course assignment probability in column 3. In both regressions, the dependent variable has been standardized to make comparison easier. Standard errors are two-way clustered at the caseworker and jobseeker level. Column 5-6 reports means and standard deviations of the covariates.

(1) *Demographics* rely on information from the population register (BEF and DREAM). Male, immigrant, descendant and married are dummies, while number of children is a count variable.

(2) *Education* rely on information from the education register (UDDA) and is based on the highest completed education (education completed up to 1 month after spell start is included). Omitted category is "10 Primary education".

(3) *Labor market history* variables rely on a register containing weekly information on UI benefits and transfers (DREAM) and on the income register (Eindkomst). UI benefits, any employment, public transfers, parental leave and education subsidy are all dummies. The employment rate, wages and number of employers are continuous and winsorized at the 99th percentile.

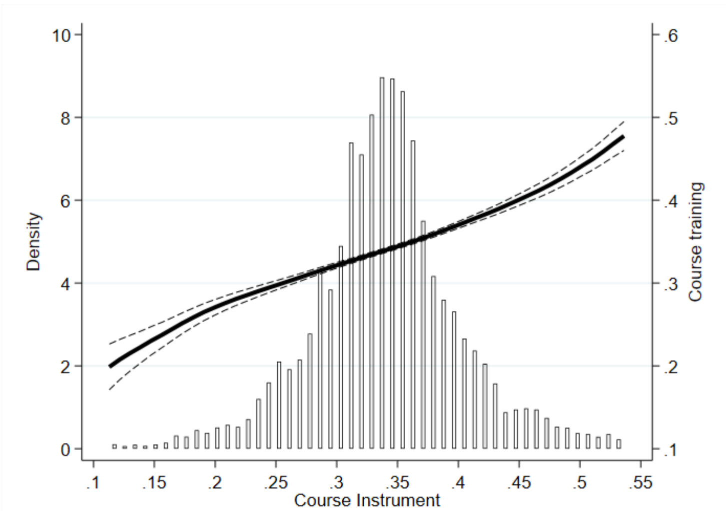
(4) *Previous industry* is based on the DREAM-register. It represents the dominating industry for the individual in the 12 months prior to the UI-spell start (the industry in which the individual had highest accumulated earnings). Omitted category is "Construction".

6.2 Relevance

Figure 2 shows graphically the relevance of our caseworker leniency instrument. All plotted values in this figure have been residualized by fully interacted jobcenter unit and year fixed effects and controlling for quarter fixed effects as well as jobseeker country of origin and age. The bars show the density of the course instrument. It has a mean of 0.34 and a standard deviation of 0.07. As depicted in the figure, there is large variation in caseworker course assignment propensities within a jobcenter unit and year. A caseworker at the bottom 5 percent of the distribution assigns less than one in four of her jobseekers to course training, while a caseworker in the top 5 percent of the distribution assigns around a half of her jobseekers to course training. Appendix Table A. 6 show percentiles of the caseworker assignment propensities.

The solid line in Figure 2 represents a local linear regression of the endogenous variable, assignment to course training, on the instrument, caseworker course assignment propensity. The dashed lines represent 95 percent confidence intervals. The regression line shows that a caseworker's course training assignment propensity (calculated excluding the jobseeker in consideration) is highly predictive of whether the jobseeker is assigned to course training.

Figure 2: Course Training Instrument



Note: The figure plots the probability of assignment to course training against the caseworker's course assignment propensity (leave-out mean). Plotted values are the residuals (with the unconditional mean added) from regressions on jobcenter × unit × year, non-western origin, age and quarter fixed effects. The line represents a local linear regression of course training on the leave-out mean (degree 1, bandwidth 0.2), and the dashed lines represent 95 percent confidence intervals. Bars represent the density of the leave-out mean. Top and bottom 1 percent are excluded.

Table 7 shows the coefficients from our first-stage regressions specified in Equations (2) and (3). The table shows that the course instrument is a highly significant predictor of whether the jobseeker is assigned to course training. Appendix Table A. 7 shows that this conclusion remains even if we include all predetermined characteristics from Table 6 in the first-stage regressions. The table shows that, holding fixed the job training instrument, assignment to a caseworker with one standard deviation higher course assignment propensity increases the probability that the jobseeker is assigned to course training by almost five percentage points.[39] The F-stat in the bottom of the table shows that the instruments are highly predictive of assignment to course and job training, and weak instruments are therefore not a concern for our IV analysis.

Table 7: First Stage of the Training Instruments

	(1) Course	(2) Job training
Course instrument	0.656***	-0.013
	(0.032)	(0.020)
Job training instrument	-0.191***	0.263***
	(0.036)	(0.045)
Obs	103027	103027
Dep var Mean	0.340	0.217
Dep var sd	0.474	0.412
Covariates	No	No
Number of FE's	677	677
F-stat (instruments)	224.345	17.950
P-value (F-stat)	0.000	0.000

Note: The columns report coefficients from a regression including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. The F-stat represents a test for joint significance of instruments (with corresponding p-value). *p<0.10 ** p<0.05 *** p<0.01.

6.3 Monotonicity

We perform two tests of monotonicity adopted from Bhuller et al. (2020). Recall from the previous section that the course training instrument is positively related to course training assignment. The first test of monotonicity relates to the requirement that this first-stage relationship is non-negative in each sub-sample of the population. We test this prediction by partitioning our sample into four quartiles based on the predicted probability of assignment to course training. This prediction is based on all the jobseeker characteristics from Table 6. For each quartile group, we estimate the first-stage regression using our baseline instruments. Table 8 implement

[39]The mean-residualized course instrument has a standard deviation of 0.07. Multiplying the first-stage coefficient of 0.656 by 0.07 gives an effect of 0.046.

this test, showing that the course instrument is indeed positive and highly significant in each of the quartiles.

Table 8: Test of Monotonicity: Course Training and Baseline Instruments

	Quartiles			
	(1)	(2)	(3)	(4)
	1	2	3	4
Course instrument	0.598***	0.734***	0.649***	0.610***
	(0.048)	(0.044)	(0.047)	(0.044)
Job training instrument	-0.158**	-0.312***	-0.127*	-0.262***
	(0.074)	(0.071)	(0.066)	(0.075)
Obs	25747	25748	25749	25752
Dep var Mean	0.246	0.316	0.360	0.438
Dep var sd	0.431	0.465	0.480	0.496
Number of FE's	656	665	666	668
F-stat (instruments)	76.356	141.288	97.266	98.510
P-value (F-stat)	0.000	0.000	0.000	0.000

Note: The sample is partitioned into four quartiles based on predicted probability of course training assignment (prediction is based on predetermined characteristics from table 6). Each column represents the coefficients from a quartile-specific first-stage regression. The first stage is based on the baseline and including jobcenter $\times$ unit $\times$ year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. The F-stat represents a test for joint significance of instruments (with corresponding p-value). *p<0.10 ** p<0.05 *** p<0.01.

Our second test of monotonicity concerns the prediction that caseworkers who are stricter toward jobseekers from one subgroup of the population must also be stricter toward jobseekers from another subgroup. We test this prediction by relying on the same quartiles as before but create 'reversed' instruments for each quartile. These instruments are measured as the average training assignment propensity for jobseekers assigned to the same caseworker but belonging to one of the other three quartiles. We then estimate the first-stage regression for each quartile using the reversed instruments. The first-stage coefficients on the reversed course training instrument are shown in Table 9. As the table shows, the coefficients are positive in all quartiles as implied by the monotonicity assumption.

In summary, we take the results of the two tests above as evidence in support of the monotonicity assumption in our IV setup.

Table 9: Test of Monotonicity: Course Training and Reversed Instruments

	Quartiles			
	(1)	(2)	(3)	(4)
	1	2	3	4
Rev. course instrument	0.561***	0.733***	0.646***	0.580***
	(0.049)	(0.047)	(0.050)	(0.046)
Rev. job training instrument	-0.107	-0.293***	-0.127*	-0.223***
	(0.072)	(0.073)	(0.070)	(0.077)
Obs	25747	25748	25749	25676
Dep var Mean	0.246	0.316	0.360	0.438
Dep var sd	0.431	0.465	0.480	0.496
Number of FE's	656	665	666	667
F-stat (instruments)	66.004	120.834	84.597	77.879
P-value (F-stat)	0.000	0.000	0.000	0.000

Note: The sample is partitioned into four quartiles based on predicted probability of course training assignment (prediction is based on predetermined characteristics from table 6). For each quartile a 'reversed' instrument is constructed: Using the average training assignment probability for jobseekers assigned to the same caseworker but belonging to one of the other three quartiles. Each column represents the coefficients from a quartile-specific first-stage regression with the 'reversed' instruments and including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. The F-stat represents a test for joint significance of instruments (with corresponding p-value). *p<0.10 ** p<0.05 *** p<0.01.

7 Causal Effects of Training on Jobseeker Outcomes

In this section, we estimate the causal impact of course training assignment on jobseeker outcomes. We first study the effects of assignment to *any type of course* on labor market outcomes 12 months after UI-spell start. We show the effect using conventional OLS, and then turn to our IV estimation approach. Hereafter, we investigate whether the effect of assignment to courses changes over time. To investigate treatment effect heterogeneity, we then zoom in on *vocational training courses*. Finally, we zoom in on jobseekers whose previous jobs were in the *manufacturing* sector. From the rise of China in global trade to the diffusion of industrial robotics, manufacturing workers have experienced rapid structural change in recent decades. From a skill mismatch perspective, these worker could therefore benefit from vocational training.

7.1 Assignment to Any Training Course

Table 10 presents the effect of assignment to *any type of course training* on individual labor market outcomes measured 12 month after UI-spell start. The first four columns of the table show coefficients from a simple OLS estimation of Equation (1). Recall that equation (1) also controls for assignment to job training, however, as the object of interest in this paper is the effect of course training, we only report this coefficient in the table.

Table 10: Effect of Assignment to Course Training on Outcomes in Month 12

	OLS				IV			
	Wage	Hours	Employed	UI-benefits	Wage	Hours	Employed	UI-benefits
A: No covariates								
Course	-4,185.510***	-21.056***	-0.143***	0.171***	-1,907.876	-7.709	-0.081**	-0.005
	(145.373)	(0.725)	(0.005)	(0.005)	(1,412.483)	(6.764)	(0.041)	(0.034)
B: With covariates								
Course	-3,625.529***	-18.680***	-0.125***	0.166***	-1,705.398	-6.647	-0.070*	-0.009
	(129.235)	(0.668)	(0.005)	(0.005)	(1,122.722)	(5.980)	(0.038)	(0.034)
Obs	101,346	101,346	103,027	103,027	101,346	101,346	103,027	103,027
Dep var Mean	15,163.857	82.418	0.659	0.297	15,163.857	82.418	0.659	0.297
Dep var sd	14,159.092	70.196	0.474	0.457	14,159.092	70.196	0.474	0.457

Note: All regressions include jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. Panel A has no further controls, while panel B controls for all predetermined characteristics in table 6. Column 1-4 represent OLS estimates from regressions with dummies for assignment to course and job training as the independent variables (equation 1). Column 5-8 represent IV estimates from regressions where assignment to course and job training are instrumented with caseworker assignment propensities (leave-out means). The dependent variables are all measured in month 12 after UI-spell start. *Wages* and *hours* refer to labor earnings measured in DKK and hours worked (non-employment coded as zero). *Employed* and *UI benefits* are dummies taking value one if the individual is employed or receives UI benefits. The number of observations is slightly lower for wages and hours in month 12 compared to employment and UI-benefits in month 12. This is because information on wages and hours only was available up until November 2019, and hence, we don't have this information for individuals that became unemployed in December 2018. *p<0.10 ** p<0.05 *** p<0.01.

The OLS estimates in panel A are well in line with the training evaluation literature as they support the notion that assignment to course training yields negative returns. The estimates are statistically significant and suggest that individuals assigned to course training earn about 4,000 DKK (640 USD) less and work about 20 hours less in month 12. Further, the OLS shows evidence of large locking-in effects, as assigned individuals are 17 percentage points more likely to receive UI benefits in month 12. Panel B shows the OLS estimates when controlling for the full set of predetermined jobseeker characteristics from Table 6. The inclusion of these control moderates the negative effects only slightly.

The last four columns in Table 10 show the coefficients from a 2SLS estimation of Equation (1) where we have instrumented assignment to training by the caseworker's training assignment propensity (leave-out mean). The IV-estimates are about half the size of the OLS estimates and statistically insignificant. This suggests that OLS suffers from a substantial negative bias caused by jobseekers with bad employment prospects selecting into training.

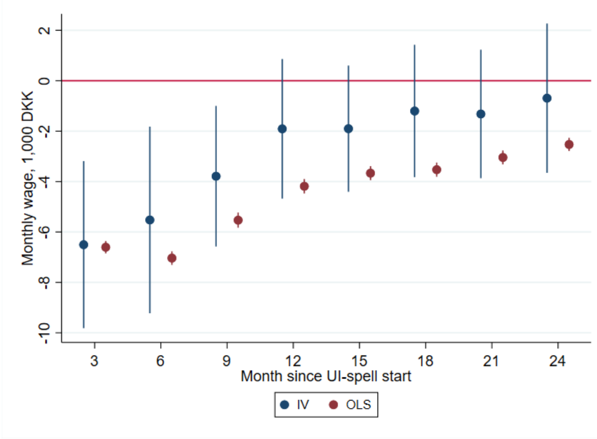
Still, taken at face value, the IV point estimates suggest that assignment to course training yield negative returns after one year. They suggest that individuals assigned to course training earn about 2,000 DKK less and work seven hours less in month 12 compared to individuals that do not get assigned. However, there is no evidence of locking-in effects on UI benefits in month 12, and the table suggests that these negative effects are driven by a negative extensive margin effect on employment. Nevertheless, these effect should be interpreted with caution as the IV-estimates (except for employment) are in fact statistically insignificant.

Effects Over Time

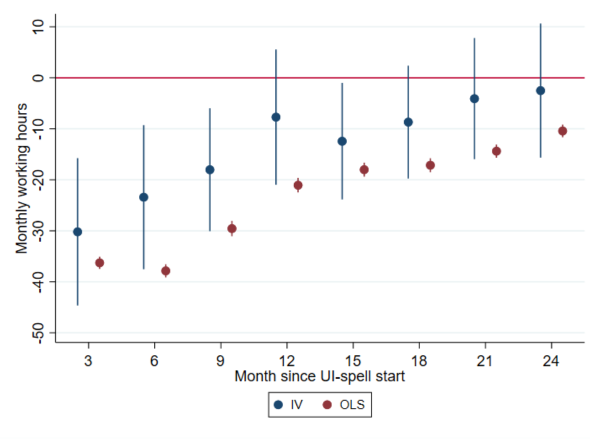
We now investigate how the effect of assignment to course training evolves over time. Figure 3 shows the effect on labor market outcomes evaluated every third month after UI-spell start and up to month 24. The blue dots (and 95 percent confidence intervals) represent the IV-estimates, while the red dots (and 95 percent confidence intervals) represent the OLS-estimates. Interestingly, the OLS and IV-estimates share the same profile, yet there is a significant difference in the levels of the point estimates. This again confirms the notion that OLS is negatively biased.

Both the OLS and IV estimates show that assignment to course training are associated with large *locking-in effects*. The locking-in effects are particularly severe around month 3, which is in line with table 3. Namely, this table shows that for the two largest course types, half of all courses start around month 3-4 (week 13 and 16) of the UI-spell. At this point in time, the OLS and IV estimates are of similar size. It suggests that all jobseekers, despite potential labor market outcomes are locked-in. The locking-in effects decay over time but at very different speeds: While the locking-in effects have vanished completely in month 12 for the IV, they still have not vanished by month 24 for the OLS. In sum, the OLS estimates suggest large, negative, and statistically significant effects on wage earnings, hours, and employment over the entire period considered. Although the OLS estimates do get closer to zero over time, they are still negative and statistically significant after two years. The IV-estimates are also negative and statistically significant in the beginning of the period. However, they get closer to zero over time and from month 12 to 24, the IV-estimates are statistically insignificant. In fact, they barely move over the last 12 months considered but stay rather constant. Hence, there is no evidence of a positive post program effect: the IV suggests that the treatment group catches up but never overtakes the control group. We finally note that the confidence bands on all IV-estimates are quite large, which could mask substantial heterogeneity in effects across courses or workers.

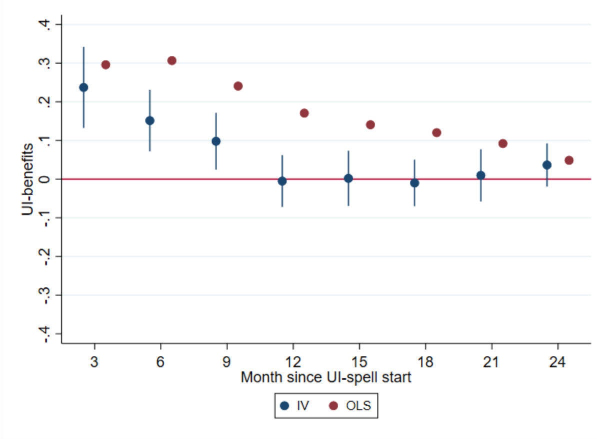
Figure 3: Effect of Course Training Over Time



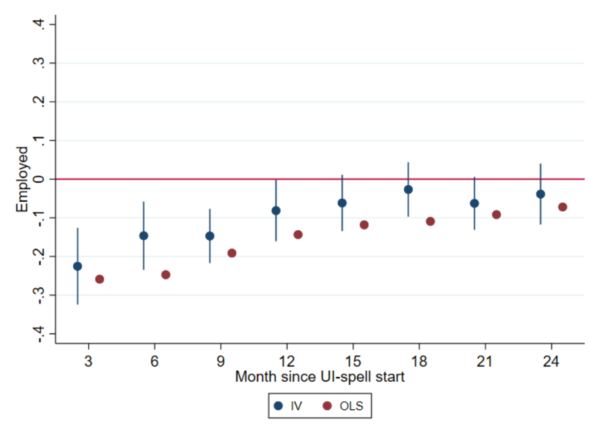
(a) Wage earnings



(b) Hours



(c) UI benefits



(d) Employed

Note: The figure plots the OLS and IV estimates of the effect of assignment to course training on labor market outcomes measured in a given month relative to UI-spell start. Note that the IV-estimates comes from regressions where assignment to course and job training are instrumented with caseworker assignment propensities (leave-out means). All regressions include jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. 95 percent confidence intervals are added to the plot. The dependent variables are measured in a given month after after UI-spell start.

Exclusion Restriction

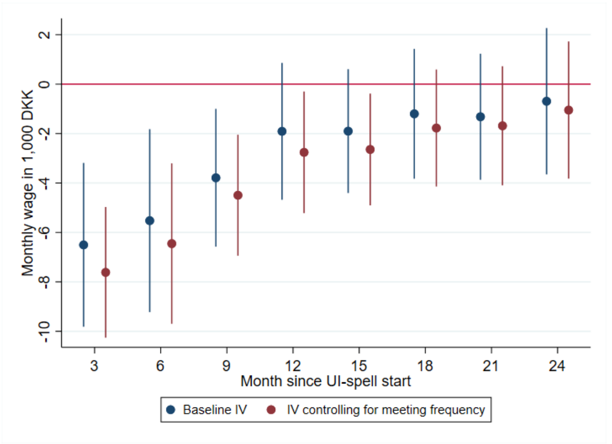
In order to interpret the IV estimates as the local average treatment effect of assignment to courses we need an exclusion restriction. In particular, the instrument can only affect labor market outcomes through the assignment to course training. We already noted that an obvious threat to this assumption is that caseworkers serve multiple purposes. For example, besides assigning jobseekers to training, the caseworkers meet with jobseekers and give advice on job search. There is evidence showing that caseworker meetings are important for transitions out of unemployment and subsequent employment (Van den Berg et al., 2012; Maibom et al., 2017; Schiprowski, 2020). Hence, a potential threat to our identification would be if caseworkers with a high training assignment propensity also meet more frequently with jobseekers. The exclusion restriction would in that case be violated, and we would need to interpret our estimates as the joint effect of assignment to course training and frequent caseworker meetings.

To test the exclusion restriction, we re-estimate our main regressions while this time controlling for the frequency with which the jobseeker meet with her caseworker.[40] As this is endogenous, we instrument the meeting frequency of jobseeker i with the caseworker's general meeting frequency. Again, we measure this as a leave-out mean.[41] Figure 4 presents our test of the exclusion restriction. In this figure, the blue dots (and 95 percent confidence intervals) represent the baseline IV estimates, and the red dots (and 95 percent confidence intervals) represent the IV estimates where we control for caseworker meeting frequency. As the figure shows, the IV estimates do not change much when we control for the caseworker meeting frequency. The level and profile of the IV estimates are very similar which we take this as evidence in support of the exclusion restriction.

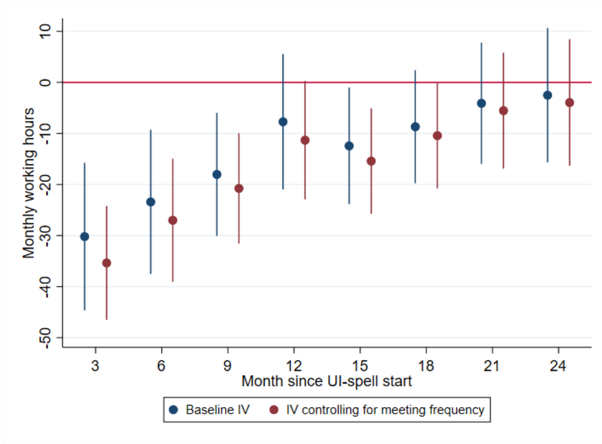
[40] We define meeting frequency for jobseeker i as the number of caseworker meetings per week of unemployment. If the jobseeker's UI-spell is longer than 26 weeks, we only consider the first 26 weeks, because meeting frequency requirements change after 26 weeks. Note that this measure includes all caseworker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

[41] For jobseeker i assigned to caseworker j, we use the average meeting frequency for all other jobseekers that were assigned to caseworker j.

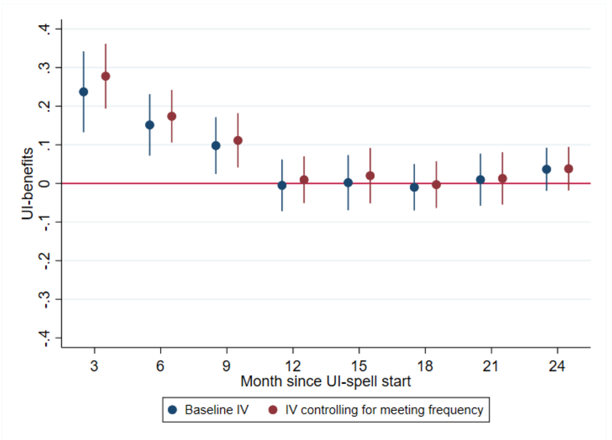
Figure 4: Controlling for Caseworker Meeting Frequency



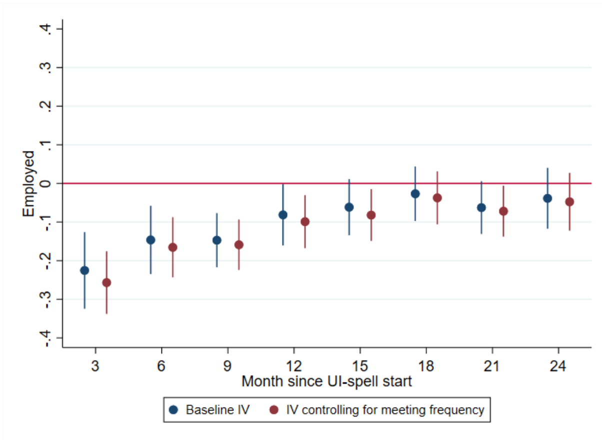
(a) Wage earnings



(b) Hours



(c) UI benefits



(d) Employed

Note: The plot presents our test of the exclusion restriction. Blue dots represent our baseline IV-estimates, while red dots represent the IV-estimates from regressions that further control for caseworker meeting frequency by instrumenting the jobseeker i's number of caseworker meetings per week with the caseworkers general meeting propensity (leave-out mean). All regressions include jobcenter $\times$ unit $\times$ year fixed effects, a dummy for non-western origin, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. 95 percent confidence intervals are added to the plot. The dependent variables are measured in a given month after after UI-spell start.

7.2 Zooming in on Vocational Training Courses

We now zoom in on *vocational training courses* and investigate the effect of assignment to such courses as opposed to not being assigned to any activity. To do this, we define a dummy for assignment to vocational training, D_i^{Voc}, and a dummy for assignment to any other courses, D_i^{Other}. Given that the individual can either be assigned to no activity, vocational training, other courses or job training, our structural equation of interest reads as follows

$$Y_{i,t} = \beta_t D_i^{Voc} + \gamma_t D_i^{Other} + \delta_t D_i^{job} + X_i'\theta_t + u_{i,t} \tag{4}$$

Again, the parameter of interest is β_t, and since the regressors, D_i^{Voc}, D_i^{Other} and D_i^{Job}, are endogenous, we instrument each of them with the caseworker's training-specific assignment propensity, measured as leave-out means.[42] We estimate β_t in a 2SLS estimation setup, where we have a first-stage regression for each of the endogenous variables.[43] We refer to the appendix for evidence of instrument relevance. In particular, Figure A. 3 plots the assignment probability against the instruments, and Table A. 8 and A. 9 confirm that the instruments indeed are relevant.

Figure 5 shows the causal effects of assignment to vocational training over time. It is evident that the locking-in effect of vocational training courses is more profound and decays at a slower rate compared to assignment to any course. Although statistically insignificant, a good 10 percent of jobseekers assigned to vocational training are still unemployed in month 12. Note that a likely explanation for why the locking-in effect is more profound for assignments to vocational training is that the enrollment rate is much higher.[44]

The profoundness of the locking-in effects likely explains why the IV estimates on employment, wage earnings, and working hours are negative and - statistically as well as economically significant - even 12 months after UI-spell start. However, around 18 months after UI-spell start,

[42]For example, the individual's assignment to vocational training, D_i^{Voc}, is instrumented with the mean vocational training assignment rate for all other jobseekers assigned to the same caseworker, $Z_{j(i)}^{Voc}$. Likewise for assignment to other courses and job training.

[43]The first-stage regressions are:

$$D_i^{voc} = \alpha_1 Z_{j(i)}^{voc} + \alpha_2 Z_{j(i)}^{other} + \alpha_3 Z_{j(i)}^{job} + X_i'\gamma + \eta_i$$

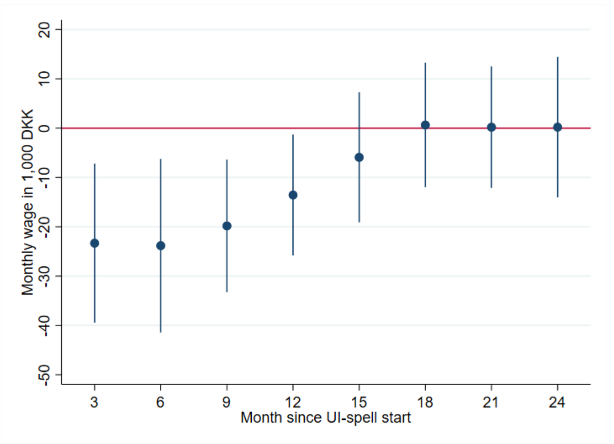
$$D_i^{other} = \xi_1 Z_{j(i)}^{voc} + \xi_2 Z_{j(i)}^{other} + \xi_3 Z_{j(i)}^{job} + X_i'\delta + e_i$$

$$D_i^{job} = \mu_1 Z_{j(i)}^{voc} + \mu_2 Z_{j(i)}^{other} + \mu_3 Z_{j(i)}^{job} + X_i'\eta + \epsilon_i$$

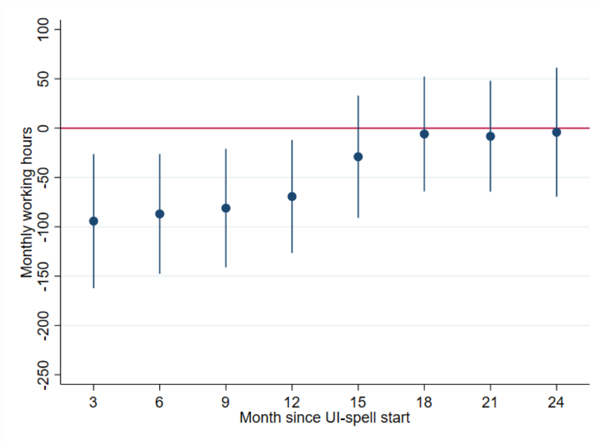
[44]More than 80 percent of jobseekers assigned to vocational training courses end up enrolling in a training course. In comparison, only around 50 percent of jobseekers assigned to any course end up enroling in some course; see Table 4.

the treatment group has completely caught up with the control group, and the point estimates suggest null effects on all outcomes. However, we again note that the confidence intervals are large. In the next subsection, we turn to investigate whether this variability reflects heterogeneity in the effectiveness of vocational training across worker types.

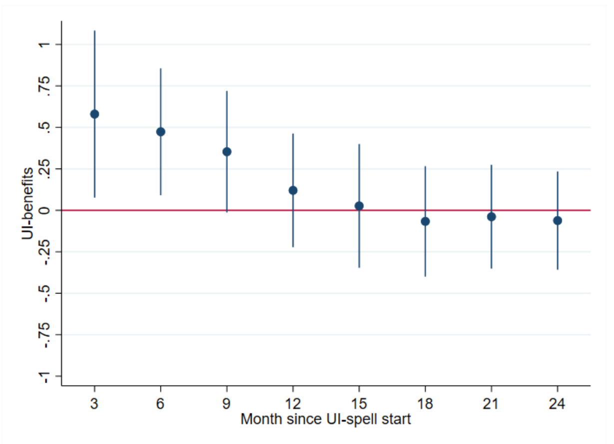
Figure 5: Effect of Vocational Training Over Time



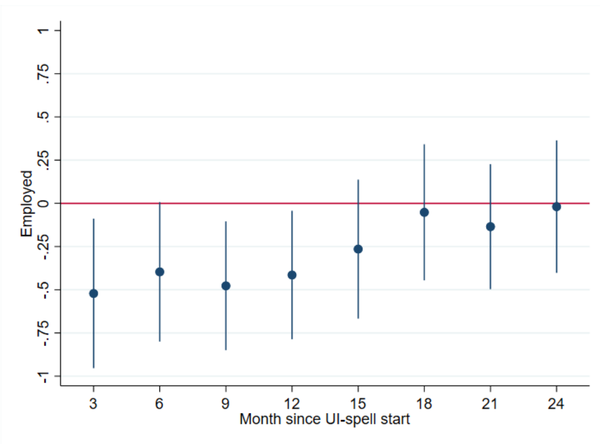
(a) Wage earnings



(b) Hours



(c) UI benefits



(d) Employed

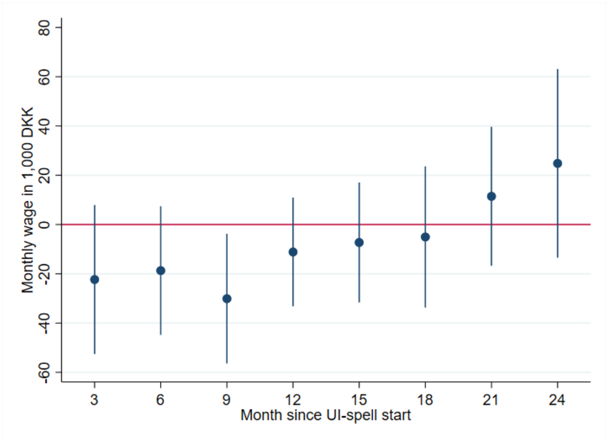
Note: Plotted values represent IV-coefficients on vocational training. All regressions including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. 95 percent confidence intervals are added to the plot. The dependent variables are measured in a given month after after UI-spell start.

7.3 Zooming in on Jobseekers From Manufacturing

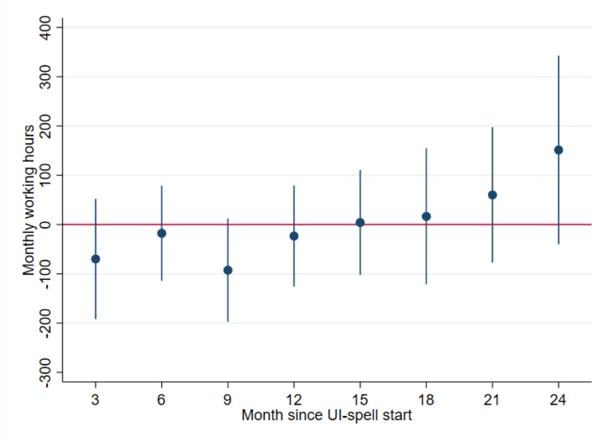
We now investigate whether vocational training is particularly beneficial for jobseekers who were previously employed in manufacturing. The idea behind this analysis is to look at a group of jobseekers who have higher exposure to globalization and robot adoption, and hence may be more likely to face mismatch problems in the labor market. By zooming in on jobseekers who were previously employed in manufacturing, our sample reduces to 12,626 UI-spells (12.3 percent of the initial sample), of which 1,412 were assigned to vocational training.

Figure 6 shows the effects of assignment to vocational training for these former manufacturing workers. The jobseekers from manufacturing also experience profound locking-in effects as well as negative and economically significant effects on earnings and hours in the beginning of the UI-spell. However, these locking-in effects vanish over time, as the IV-estimates start increasing steadily from month 9. Most interestingly, the IV estimates keep increasing such that the treatment group not only catches up but actually overtakes the control group. Although not statistically significant, the estimates suggest that a positive post-program effect sets in. Taken at face value, the IV estimates suggest that assignment to vocational training increases wage earnings by around 20,000 DKK (around 3,200 USD) and working hours by 150 hours in month 24. It should be noted that the confidence bands are very wide. This is not surprising, since we have restricted the sample considerably. Nevertheless, the positive trend in the treatment effects is evident in all sub figures (negative trend for UI-benefits), and the sizable point estimates suggest that vocational training deliver economically significant long-run benefits to jobseekers from manufacturing.

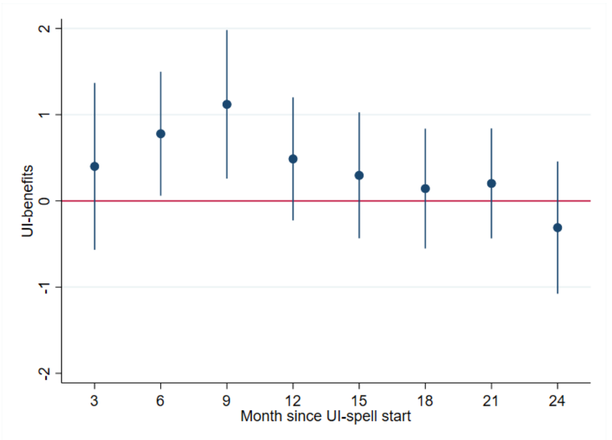
Figure 6: Effect of Vocational Training Over Time:
Jobseekers From Manufacturing



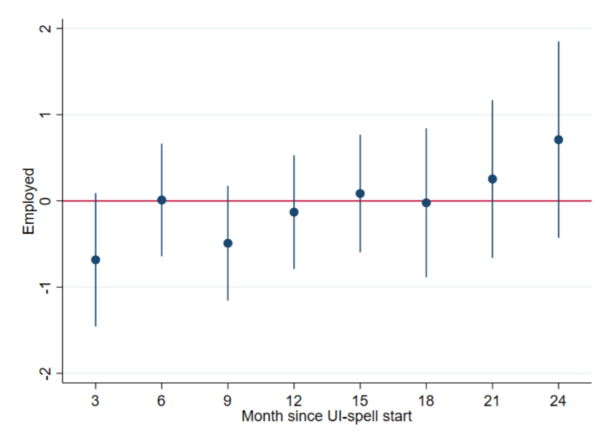
(a) Wage earnings



(b) Hours



(c) UI benefits



(d) Employed

Note: Plotted values represent IV-coefficients on vocational training from regressions including only jobseekers from manufacturing. All regressions including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. 95 percent confidence intervals are added to the plot. The dependent variables are measured in a given month after after UI-spell start.

8 Conclusion

In this paper, we provide causal evidence on the effectiveness of vocational training for jobseekers. Our key contribution is to develop a caseworker leniency instrument that solves the selection problem that jobseekers opt into training based on unobserved job prospects (Ashenfelter, 1978).

A takeaway from our analysis is that solving the selection problem is crucial for uncovering the causal effects of training on employment. While OLS estimates of training assignment on employment are strongly negative, the IV estimates are half the size and insignificant.

The confidence bands of our IV estimates are wide, however, which could hint at important heterogeneity in the effects of training. In particular, we zoom in on displaced manufacturing workers to find substantial long-run benefits of vocational training. These findings suggest that training may be more beneficial for workers who have lost their job to foreign competition or industrial robots. In this sense, our estimates provide empirical support for targeted training interventions like the Trade Adjustment Assistance (TAA) program in the United States.

To inform cost-benefit analyses of training programs, we estimate effects on earnings and employment (measuring the benefit side), as well as unemployment benefit claims (a key component on the cost side).[45] In further work, we plan to plug these estimates into a cost-benefit model to evaluate the net benefits of a counterfactual system with more targeted training.[46]

Evaluating counterfactual training systems is tricky, however, as it involves a different course assignment than the one that generated our local average treatment effects (LATEs). We plan here to build on extrapolation methods developed in Mogstad et al. (2018) to convert the caseworker-induced LATEs into the policy-relevant treatment parameters of interest.

[45] In further work, we plan to also consider course expenditures.

[46] We are interested in answering questions like *"how many jobs could be created by better targeting of existing training expenditures?"*.

References

Ashenfelter, O. (1978), 'Estimating the effect of training programs on earnings', *The Review of Economics and Statistics* pp. 47–57.

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

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.

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., Harmon, N. and Maibom, J. (2020), 'The dynamics of job search in unemployment: Beyond search effort and reservation wages', *mimeo* .

Geerdsen, L. P. (2006), 'Is there a threat effect of labour market programmes? a study of almp in the danish ui system', *The Economic Journal* **116**(513), 738–750.

Geerdsen, L. P. and Holm, A. (2007), 'Duration of ui periods and the perceived threat effect from labour market programmes', *Labour Economics* **14**(3), 639–652.

Humlum, A. and Munch, J. R. (2019), 'Globalization, flexicurity and adult vocational training in denmark', *Making Globalization More Inclusive* p. 51.

Hummels, D., Munch, J. R. and Xiang, C. (2018), 'Offshoring and labor markets', *Journal of Economic Literature* **56**(3), 981–1028.

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.

Maibom, J., Rosholm, M. and Svarer, M. (2017), 'Experimental evidence on the effects of early meetings and activation', *The Scandinavian Journal of Economics* **119**(3), 541–570.

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.

McKinsey (2017), Jobs lost, jobs gained: Workforce transitions in a time of automation, Report.

Mogstad, M., Santos, A. and Torgovitsky, A. (2018), 'Using instrumental variables for inference about policy relevant treatment parameters', *Econometrica* **86**(5), 1589–1619.

Munch, J. R. and Skipper, L. (2008), 'Program participation, labor force dynamics, and accepted wage rates', *Advances in Econometrics* **21**, 197–262.

OECD (2019), Oecd employment outlook 2019: The future of work, Report.

Rasmussen, M. (2021), 'Caseworker quality and transitions out of unemployment: Evidence and policy implications.', *mimeo* .

Rosholm, M. and Skipper, L. (2009), 'Is labour market training a curse for the unemployed? evidence from a social experiment', *Journal of Applied Econometrics* **24**(2), 338–365.

Rosholm, M. and Svarer, M. (2008), 'The threat effect of active labour market programmes', *scandinavian Journal of Economics* **110**(2), 385–401.

Schiprowski, A. (2020), 'The role of caseworkers in unemployment insurance: Evidence from unplanned absences', *Journal of Labor Economics* .

Van den Berg, G., Kjærsgaard, K. and Rosholm, M. (2012), 'To meet or not to meet (your case worker) - that is the question', *IZA Discussion Paper 6476* .

Appendix

Table A. 1: Sampling

	Meetings	Spells	Jobseekers	Caseworkers	Jobcenters
Caseworker data	**4,401,504**	**1,110,742**	**651,089**	**14,043**	**94**
No missing covariates$^{(1)}$	4,395,767	1,109,387	650,373	14,036	94
Timing of first meeting$^{(2)}$	4,002,681	1,000,816	606,610	13,628	94
Distance to previous spell$^{(3)}$	3,509,206	865,551	584,912	13,292	94
Identify assignment caseworker and meeting $^{(4)}$	3,506,162	864,036	584,148	13,290	94
Only the assignment meeting$^{(5)}$	864,036	864,036	584,148	10,088	94
Meeting type$^{(6)}$	803,372	803,372	557,007	7,632	94
Meeting contact$^{(7)}$	787,073	787,073	548,383	7,468	94
Assignment mechanism$^{(8)}$	151,953	151,953	106,782	1,486	29
Caseload size$^{(9)}$	141,998	141,998	101,479	612	25
Caseworker work experience $^{(10)}$	141,088	141,088	100,930	601	25
Min. 2 caseworkers per randomization cell$^{(11)}$	140,781	140,781	100,698	601	25
Sample	**140,781**	**140,781**	**100,698**	**601**	**25**
Sample with r^2 restriction $^{(12)}$	**103,027**	**103,027**	**75,811**	**467**	**24**

Note: The caseworker data consists of all jobseekers who entered unemployment from 2011-2018 and has at least one meeting registration in a jobcenter.

$^{(1)}$ I drop observations if predetermined characteristics are missing. This includes demographics (age, gender, immigrant, decendent, western, married, children), labor market history (past unemployment, employment, receipt of public transfers) and UI-fund association. Note that I allow for missing education and/or industry, and code missings as a category.

$^{(2)}$ I require that the first meeting takes place in the same week or up to 13 weeks after spell start. This is based on the rule sin PES specifying that the first meeting had to take place within 3 months (1 month) until 2015 (since 2015).

$^{(3)}$ I require min 10 weeks between the current and any potential previous UI-spell. This is to be sure that we are observing new UI-spells and not a continuation of old spells.

$^{(4)}$ I identify the *assigned* caseworker based on the jobseeker's *first individual meeting*. In most jobcenters, jobseekers have the first meeting with their assigned caseworker. In a few jobcenters, jobseekers first go to a group meeting (an information meeting with a group of other jobseekers) and only in the second meeting meet their assigned caseworker. I therefore define the *assigned caseworker* as the caseworker from the first meeting if this is not an information meeting (coded as "Informationsmøde"), and otherwise as the caseworker from the second meeting. Note that some jobseekers leave unemployment before making it to the second meeting, and I cannot define an assignment caseworker for these jobseekers. Hence, they are dropped from the sample.

$^{(5)}$ I restrict the sample to the first individual meeting only. Hence the level of observation goes from spell$\times$meeting to spells.

$^{(6)}$ I require that the first individual meeting is coded as a "job interview" ("jobsamtale" or "jobsamtale med deltagelse af a-kasse")

$^{(7)}$ I require that the first individual meeting takes place in person ("personlig kontakt")

$^{(8)}$ Based on the jobcenter survey, I know the yearly assignment mechanism for all responding jobcenters over the sample period. I keep all jobcenter $\times$ years in which they used report to use "birthday" assignment.

$^{(9)}$ I restrict the sample to caseworkers that have at least 50 cases

$^{(10)}$ I restrict the sample to caseworkers who work in at least 10 weeks

$^{(11)}$ Randomization occurs at the jobcenter $\times$ unit $\times$ year level, and hence I require at least two caseworkers per randomization cell.

$^{(12)}$ For each caseworker, I test how well jobseeker birth day of the month (1-31) predict assignment to a particular caseworker. In particular, for each caseworker, I run a regression with a dummy for assignment to casewrker j on birthday $\times$ quarter an jobcenter $\times$ unit $\times$ year FE's. I then save the within-r^2 from each of these regressions. To have a benchmark, I run a similar set of regressions including placebo birthdays instead of true birthday fixed effects (again interacted with quarters). Finally, I drop caseworkers if the within-r^2 on true birthdays are lower than the median within-r^2 for placebo birthdays, i.e. if the within-$r^2 < 0.17$. Note that I again check that I have at least 2 caseworkers per randomization cell after this last restriction.

Table A. 2: Summary Statistics

	Full sample		Course training		No assignment		Diff in means
	mean	sd	mean	sd	mean	sd	col 3-5
Demographics[(1)]							
Age	40.231	(12.280)	40.413	(12.501)	40.346	(12.064)	0.066
Male	0.484	(0.500)	0.475	(0.499)	0.504	(0.500)	-0.029***
Immigrant	0.050	(0.218)	0.066	(0.248)	0.042	(0.200)	0.024***
Descendant	0.002	(0.043)	0.002	(0.048)	0.002	(0.040)	0.001**
Western	0.957	(0.203)	0.942	(0.234)	0.965	(0.184)	-0.023***
Married	0.422	(0.494)	0.418	(0.493)	0.429	(0.495)	-0.011***
Number of children	0.806	(1.039)	0.787	(1.034)	0.821	(1.045)	-0.034***
Education[(2)]							
0. Missing	0.013	(0.115)	0.016	(0.126)	0.012	(0.109)	0.004***
15. Preparatory course	0.004	(0.062)	0.005	(0.069)	0.003	(0.053)	0.002***
20. Upper secondary	0.047	(0.212)	0.054	(0.226)	0.044	(0.206)	0.010***
30. Vocationa educ.	0.476	(0.499)	0.443	(0.497)	0.496	(0.500)	-0.053***
35. Qualifying program	0.000	(0.022)	0.001	(0.023)	0.000	(0.022)	0.000
40. Short cycle tertiary	0.047	(0.211)	0.054	(0.227)	0.043	(0.202)	0.012***
50. Vocational bach.	0.125	(0.331)	0.118	(0.323)	0.127	(0.333)	-0.008***
60. Bachelor	0.014	(0.118)	0.018	(0.134)	0.012	(0.107)	0.007***
70. Master	0.049	(0.215)	0.060	(0.237)	0.041	(0.199)	0.019***
80. PhD	0.003	(0.053)	0.004	(0.060)	0.002	(0.050)	0.001***
Labor market history[(3)]							
UI-benefits in year t-1	0.408	(0.491)	0.378	(0.485)	0.429	(0.495)	-0.052***
UI-benefits in year t-2	0.424	(0.494)	0.391	(0.488)	0.446	(0.497)	-0.055***
Any employment in year t-1	0.900	(0.300)	0.869	(0.337)	0.926	(0.262)	-0.057***
Any employment in year t-2	0.922	(0.268)	0.903	(0.295)	0.937	(0.244)	-0.033***
Employment rate in year t-1	0.642	(0.352)	0.609	(0.371)	0.673	(0.331)	-0.065***
Employment rate in year t-2	0.659	(0.348)	0.642	(0.363)	0.676	(0.335)	-0.033***
Wage earnings in 1,000 DKK in year t-1	212.235	(147.811)	201.451	(155.240)	224.702	(142.889)	-23.251***
Wage earnings in 1,000 DKK in year t-2	209.295	(141.285)	203.096	(146.294)	217.343	(138.672)	-14.247***
Number of employers in year t-1	1.398	(0.944)	1.300	(0.925)	1.479	(0.952)	-0.180***
Number of employers in year t-2	1.407	(0.907)	1.342	(0.894)	1.457	(0.915)	-0.116***
Public transfers in year t-1	0.627	(0.484)	0.610	(0.488)	0.637	(0.481)	-0.027***
Parental leave in year t-1	0.083	(0.276)	0.089	(0.284)	0.077	(0.266)	0.012***
Education subsidy in year t-1	0.101	(0.302)	0.102	(0.303)	0.097	(0.296)	0.005***
Previous industry[(4)]							
Real estate	0.011	(0.106)	0.012	(0.109)	0.011	(0.103)	0.001*
Business services	0.115	(0.320)	0.124	(0.330)	0.113	(0.317)	0.011***
Finance	0.010	(0.098)	0.012	(0.110)	0.008	(0.090)	0.004***
Trade & transport	0.199	(0.400)	0.215	(0.411)	0.187	(0.390)	0.028***
Manufacturing	0.123	(0.328)	0.128	(0.334)	0.122	(0.328)	0.005**
Communication & it	0.017	(0.130)	0.021	(0.142)	0.015	(0.123)	0.005***
Culture	0.036	(0.185)	0.032	(0.176)	0.039	(0.192)	-0.007***
Agriculture, forestry & fishing	0.026	(0.159)	0.022	(0.147)	0.029	(0.167)	-0.006***
Public administration, health, education	0.224	(0.417)	0.204	(0.403)	0.232	(0.422)	-0.028***
UI-association [(5)]							
Academics Association	0.064	(0.245)	0.078	(0.269)	0.056	(0.229)	0.023***
Danish Trade Union Association	0.654	(0.476)	0.625	(0.484)	0.672	(0.469)	-0.047***
No association (yellow)	0.282	(0.450)	0.297	(0.457)	0.272	(0.445)	0.025***
Obs	103027		35031		58205		93236

Note: Column 1-2 shows the mean and SD of characteristics for all spells in the sample. Column 3-4 shows the mean and SD for spells with a least one course assignment (including spells with a course assignment only *and* spells with a course as well as a job training assignment). Column 5-6 shows the mean and SD for spells with no assignment (to job nor to course training).

(1) *Demographics* rely on information from the population register (BEF and DREAM). Male, immigrant, descendant and married are dummies, while number of children is a count variable.

(2) *Education* rely on information from the education register (UDDA) and is based on the highest completed education (education completed up to 1 month after spell start is included). Omitted category is "10 Primary education".

(3) *Labor market history* variables rely on a register containing weekly information on UI benefits and transfers (DREAM) and on the income register (Eindkomst). UI benefits, any employment, public transfers, parental leave and education subsidy are all dummies. The employment rate, wages and number of employers are continuous and winsorized at the 99th percentile.

(4) *Previous industry* is based on the DREAM-register. It represents the dominating industry for the individual in the 12 months prior to the UI-spell start (the industry in which the individual had highest accumulated earnings). Omitted category is "Construction".

(5) *UI-association* are based on information about the individual's UI-fund memebership from the DREAM-register. There are 25 UI-funds; these may be belong to the Trade Union Association (FH), the Academics Association (AC) or to no association.

Table A. 3: Treatment Dummies
Assignment vs. Enrollment

	Enrolment		Enrolment education types		
	No	Yes	Prim./sec.	Vocational	Post sec.
Panel A: Levels					
Any course training [(1)]	18,357	16,674	3,395	12,605	2,573
Vocational course [(2)]	1,378	6,447	684	5,906	620
Other courses [(3)]	17,695	13,149	3,134	9,278	2,345
Panel B: Percentages					
Any course training [(1)]	52	48	18	68	14
Vocational course [(2)]	18	82	9	82	9
Other courses [(3)]	57	43	21	63	16

Note: The table relates treatment dummies, *assignment to some course category*, to course *enrollments* from the Danish course participant register (VEUV). All courses enrolled in during the UI-spell are included. In panel A, the first two columns show how many jobseekers with a given course assignment that actually end up enrolling in a course. The next three columns show what type of courses these jobseekers enrol in. There are overall three types: Primary and secondary education (column 3), Vocational training (column 4) and Post secondary education (column 5). Since jobseekers can enrol in multiple courses of different types, the enrolment types (column 3-5) do not sum to the number of jobseekers that enrol in at least one course (column 2). In panel B, the first two columns show the percent of jobseekers with a given course type assignment that actually end up enrolling in a course. The next three columns shows the percent of jobseekers that enrol in a given course type (conditional on enrollment).
[(1)] Any course training is a dummy taking value one if the individual is assigned to any course (1-9) listed in table 4
[(2)] Vocational training is a dummy taking value one if the individual is assigned to vocational training courses (2) from table 4
[(3)] Other courses is a dummy taking value one if the individual is assigned to any other course than vocational training from table 4

Table A. 4: Distribution of Labor Market Outcomes

				Percentile				
	Obs	Mean	SD	10	25	50	75	90
A: No assignment								
UI-spell	58,205	15.62	11.86	4.00	8.00	13.00	20.00	30.00
Wages in month 12	57,308	17,200.50	14,282.55	0.00	0.00	19,196.72	27,125.91	34,373.52
Hours in month 12	57,308	91.10	68.28	0.00	0.00	119.00	160.00	160.00
Wages in month 24	51,060	18,485.48	14,885.18	0.00	0.00	20,726.56	28,277.97	36,044.48
Hours in month 24	51,060	95.67	68.32	0.00	0.00	129.00	160.00	160.00
B: Assigned to course training								
UI-spell	35,031	43.09	27.93	13.00	21.00	35.00	60.00	90.00
Wages in month 12	34,387	12,185.41	13,771.73	0.00	0.00	6,056.29	23,166.93	30,870.81
Hours in month 12	34,387	67.91	70.90	0.00	0.00	35.30	150.00	160.00
Wages in month 24	29,503	15,345.96	14,485.72	0.00	0.00	16,640.27	25,850.18	33,522.96
Hours in month 24	29,503	84.15	71.63	0.00	0.00	108.00	160.00	160.00

Note: Panel A shows the distribution of outcomes for spells with no assignment (to job nor to course training). Panel B shows the distribution of outcomes for spells with a least one course assignment (including spells with a course assignment only *and* spells with a course as well as a job training assignment). UI-spells are constructed based on the DREAM register, a 4-week exit criterion and meeting registrations. The UI-spell *end* is identified as the first week (following a meeting registration) in which the individual in the 4 following weeks do not receive UI benefits. The spell start is identified by going back in time from the UI-spell end until the individual in 4 consecutive weeks did not receive UI benefits. UI-spells are capped at 104 weeks (max benefit duration). For the last year of the sample period (2018), UI-spells above 52 weeks are censored. Hours and wages are based on the income register collected by the Danish tax authorities (Eindkomst). *Wages* refer to labor earnings in a given month and are measured in DKK. *Hours* refer to hours worked in a given month (with non-employment coded as zero hours). To comply with data protection rules, all percentiles are based on at least 5 observations.

Table A. 5: Caseworker Characteristics

	Obs	Mean	SD	Percentiles				
				10	25	50	75	90
Caseload size	467	220.61	203.33	62.60	82.80	151.60	286.40	419.80
Meetings	467	985.46	814.36	226.00	367.40	765.60	1377.40	1990.40
Working weeks	467	84.53	53.12	27.80	42.00	74.00	106.00	174.80
Meetings / week	467	11.62	5.28	3.87	7.89	11.82	15.00	18.32

Note: To comply with data protection rules, all percentiles are based on at least 5 observations. Caseload size is the number of UI-spells assigned to the caseworker. This statistic involves only UI-spells from the final sample. The latter three statistics - meetings, working weeks and meetings/week - are based on all UI-spells in the caseworker data described in section x. Meetings refer to the total number of meeting slots (unique time and date) registered for the caseworker. Working weeks refer to the total number of weeks in which the caseworker registered at least one meeting. Meetings/week is total meetings divided by working weeks for the caseworker.

Table A. 6: Distribution of Course Instrument

	Obs	Mean	SD	Percentile						
				5	10	25	50	75	90	95
Course instrument, raw	103,027	0.34	0.14	0.16	0.20	0.24	0.32	0.42	0.51	0.57
Course instrument, residualized	103,027	0.34	0.07	0.24	0.26	0.31	0.34	0.37	0.42	0.46

Note: The table shows the distribution of course instrument. The first row shows the raw distribution, while the second row shows the distribution of the mean-residualized instrument (residualized by fully interacted jobcenter unit and year fixed effects, non-western origin, quarter and age fixed effects). To comply with data protection rules, all percentiles are based on at least 5 observations.

Table A. 7: First stage with Covariates
Course Training and Job Training

	(1)	(2)
	Course	Job training
Course instrument	0.652***	-0.019
	(0.032)	(0.019)
Job training instrument	-0.210***	0.251***
	(0.035)	(0.045)
Obs	103027	103027
Dep var Mean	0.340	0.217
Dep var sd	0.474	0.412
Covariates	Yes	Yes
Number of FE's	677	677
F-stat (instruments)	55.934	68.535
P-value (F-stat)	0.000	0.000

Note: The columns report coefficients from a regression including jobcenter $\times$ unit $\times$ year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. The F-stat represents a test for joint significance of instruments (with corresponding p-value). *p<0.10 ** p<0.05 *** p<0.01.

Table A. 8: First stage
Vocational training, Other Courses, and Job Training

	(1) VEU	(2) Other courses	(3) Job training
Vocational training instrument	0.352***	0.054	0.005
	(0.043)	(0.053)	(0.048)
Other courses instrument	0.026**	0.662***	-0.015
	(0.012)	(0.034)	(0.019)
Job training instrument	-0.035	-0.178***	0.264***
	(0.024)	(0.035)	(0.045)
Obs	103027	103027	103027
Dep var Mean	0.076	0.299	0.217
Dep var sd	0.265	0.458	0.412
Covariates	No	No	No
Number of FE's	677	677	677
F-stat (instruments)	27.855	139.711	12.038
P-value (F-stat)	0.000	0.000	0.000

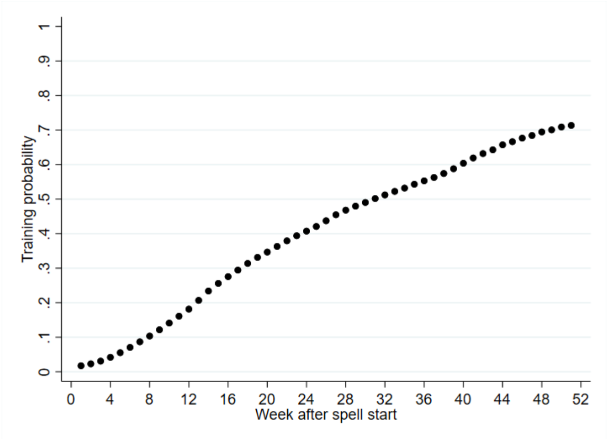
The columns report coefficients from a regression including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. The F-stat represents a test for joint significance of instruments (with corresponding p-value). *p<0.10 ** p<0.05 *** p<0.01.

Table A. 9: First Stage with Covariates
Vocational Training, Other Courses, and Job Training

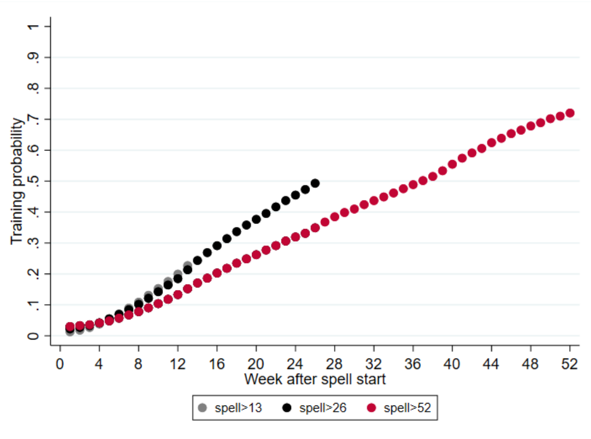
	(1) VEU	(2) Other courses	(3) Job training
Vocational training instrument	0.347***	0.061	0.020
	(0.045)	(0.053)	(0.046)
Other courses instrument	0.029**	0.655***	-0.024
	(0.012)	(0.034)	(0.018)
Job training instrument	-0.038	-0.195***	0.252***
	(0.024)	(0.034)	(0.045)
Obs	103027	103027	103027
Dep var Mean	0.076	0.299	0.217
Dep var sd	0.265	0.458	0.412
Covariates	Yes	Yes	Yes
Number of FE's	677	677	677
F-stat (instruments)	16.463	59.964	66.860
P-value (F-stat)	0.000	0.000	0.000

The columns report coefficients from a regression including jobcenter × unit × year fixed effects, a dummy for non-western, age and quarter fixed effects. Standard errors are two-way clustered at the caseworker and jobseeker level. The F-stat represents a test for joint significance of instruments (with corresponding p-value). *p<0.10 ** p<0.05 *** p<0.01.

Figure A. 1: Training Probability Over the Spell



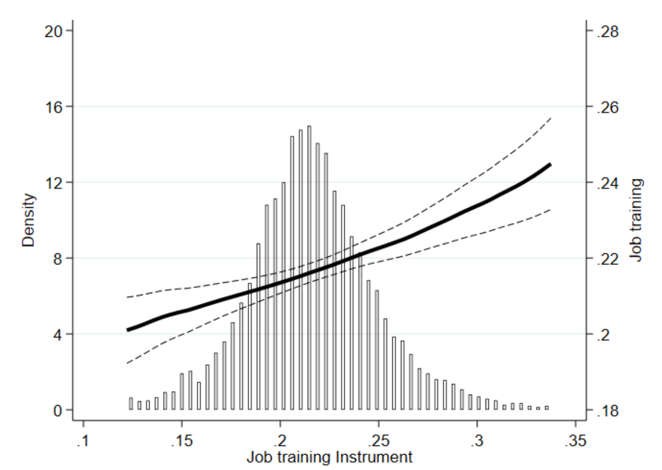
(a) Raw



(b) Condition on Spell Length

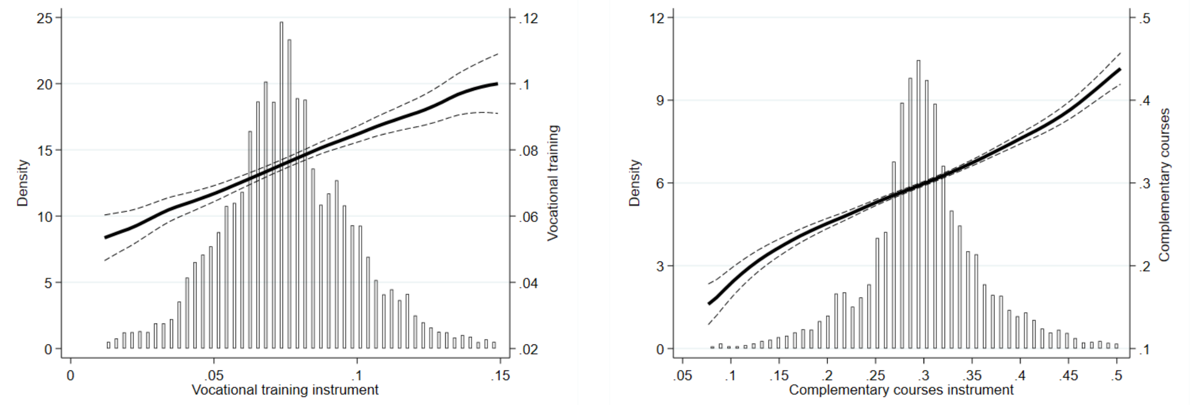
Note: The panels plot the probability to be assigned to any type of training that *starts* in a given week of the UI-spell. Panel A plots the raw probability (all spells, regardless of length), while panel B seek to sort out dynamic selection by conditioning on spell length. E.g. the black dots in panel shows the probability of assignment in a given week, conditional on having a spell of at least 26 weeks.

Figure A. 2: Job Training Instrument



Note: The figure plots the probability of assignment to job training against the case-worker's job training assignment propensity (leave-out mean). Plotted values are the residuals (with the unconditional mean added) from regressions on jobcenter × unit × year, non-western origin, age and quarter fixed effects, The line represents a local linear regression of the course category on the leave-out mean (degree 1, bandwidth 0.2), and the dashed lines represent 95 percent confidence intervals. Bars represent the density of the leave-out mean. Top and bottom 1 percent are excluded.

Figure A. 3: First Stage: Vocational Training and Other Courses



(a) Vocational Training

(b) Other Courses

Note: The panels plot the probability of assignment to a given course category against the caseworker's assignment propensity for that particular category (leave-out mean). Plotted values are the residuals (with the unconditional mean added) from regressions on jobcenter × unit × year, non-western origin, age and quarter fixed effects, The line represents a local linear regression of the course category on the leave-out mean (degree 1, bandwidth 0.1 in panel a and 0.2 in panel b), and the dashed lines represent 95 percent confidence intervals. Bars represent the density of the leave-out mean. Top and bottom 1 percent are excluded.

Chapter 3

Helping the Unemployed Through Statistical Prediction?

Helping the Unemployed Through Statistical Prediction?

Nikolaj A. Harmon* **Robert Mahlstedt†**
Mette Rasmussen‡

March 1, 2021

Abstract

We evaluate the effect of using statistical profiling tools to inform workers about their individual risk of long-term unemployment. In Denmark, a Machine Learning algorithm informs both newly unemployed workers and their case workers of whether they belong to a group with a high risk of remaining unemployed for more than six months. Leveraging age discontinuities in the Machine Learning algorithm, we use a regression discontinuity design to estimate the effect of being reported as being in high risk. We estimate that workers marginally flagged as high-risk are between 5-14 percent less likely to be unemployed after 6 months. After 12 months however this difference has disappeared. Unfortunately, standard validity checks suggest that the identifying assumptions of our regression discontinuity design may not hold. While our results thus points to statistical prediction tools as a promising way to speed up unemployment exit, more evaluation is necessary to reach any firm conclusion.

*University of Copenhagen; `nikolaj.harmon@econ.ku.dk`
†University of Copenhagen, IZA and DFI; `robert.mahlstedt@econ.ku.dk`
‡University of Copenhagen; `mette.rasmussen@econ.ku.dk`

Acknowledgement

This paper builds on and repeats text from the master's thesis of Mette Rasmussen, "Behavioral Responses to Information about Individual Employment Prospects: Evidence from Denmark" (University of Copenhagen, 2019). The master's thesis was mainly concerned with the reconstruction of the historical version of the profiling algorithm, and included only very introductory RDD estimations. The project has been extended in multiple and important dimensions since Mette defended her master's thesis. For example, we have worked extensively on the actual RDD analysis. This includes refining the initial strategy and carefully constructing samples, which allow us to explore heterogeneity in treatment effects. Further, we have done a thorough examination of the identifying assumptions and investigated the dynamics of the treatment effects. Finally, the project has been substantially rewritten, and the overlap with the text in the master's thesis is therefore limited.

1 Introduction

It is often argued that we are in the middle of a *Big Data* revolution: In combination with recent developments in Machine Learning and related techniques, large and detailed data sets on individual behavior are increasingly available and offer a range of new possibilities. For public policy, a particular promise is the possibility to use statistical prediction to improve the decisions of both policy makers, administrators and citizens. During the last years, we saw an increased use of such prediction tools in many areas, such as law enforcement (Santtila et al., 2004; Persico and Todd, 2005; Bjerk, 2007; Phillips and Pohl, 2012), health care (Burgess Jr et al., 2000) or the labor market (Black et al., 2007).

In this paper, we evaluate a particular example of such a *Big Data* approach to labor market policy, specifically in relation to combating long term unemployment. The Danish Employment Agency has developed a Machine Learning algorithm that uses rich survey and administrative data to predict the risk of long-term unemployment for newly unemployed individuals - a so-called *profiling tool*. Many countries have implemented such statistical profiling of unemployed workers (see e.g. Desiere and Langenbucher, 2018, for an overview). In conjunction with a survey distributed to newly unemployed job seekers, the Danish profiling tool relies on a range of individual characteristics to determine whether the job seeker is at-risk of becoming long-term unemployed or not. Afterwards, both the individual and her caseworker receive this information, which should give them the opportunity to adjust the job search strategy, counseling and the targeting of active labor market programs. In this paper, we examine the effects of being flagged at-risk on the unemployed individual's reemployment prospects.

In an attempt to identify causal effects, we leverage that the profiling tool lends itself to a Regression Discontinuity Design (RDD) using age as the running variable. The profiling tool is based on a decision tree that includes age as a key splitting variable. Accordingly, the predicted risk of long-term unemployment for many individuals changes discontinuously at certain age cutoffs. This suggest that a comparison of individuals just above and below the relevant age cutoff should provide as-good-as-random variation in whether the individual is flagged as being at risk of long-term unemployment. For each job seeker who fills out the survey, we thus use information on her other characteristics to determine whether she faces an age cutoff in the regression tree underlying the prediction tool and at what value the cutoff falls. The resulting sample of job seekers subject to age cutoff falls naturally in two groups: A group of young workers face cutoffs at either age 28 or 29, while a group of older workers face cutoffs at either 54 or 56 years. We analyze these two groups separately in the RDD.

As our main outcome we analyze whether the individual is still unemployed six months after (potentially) getting flagged as having a high risk of long-term unemployment. We find important effects for both the young and old group of workers. Getting flagged and receiving the message about being high risk lowers the likelihood of remaining unemployed after six months by 14 percentage points for the young sample and by 5 percentage points for the old sample. Analyzing employment outcomes suggests that the lower likelihood of being unemployed at least to some extent is explained by job seekers switching to self-support or other forms of public benefits: Point estimates are small and we see no significant effects on employment for either group. At the same time, however, confidence intervals are wide enough to explain most of the effect on unemployment. We also see no significant effects on working hours or wages.

Examining the effect of the high risk message over time shows that the positive effect on unemployment exit is focused around months 4-6 after potentially receiving the message but appears to die out in the longer run. After 12 months, we see no differences between individuals who marginally get flagged as high-risk and those who marginally do not.

The validity of our estimates relies on the usual RDD assumption that unemployed individuals just above and below the relevant age cutoffs are not systematically different except for the treatment status. Ex ante it appears very likely that this assumption holds. Although job seekers themselves choose when to participate in the survey that generates the statistical prediction, they have little scope for knowing how their exact age will affect the prediction. Similarly, there are no other known age-related mechanisms that should generate differences across the two groups.

As a check of the identifying assumption, we conduct the usual tests based on the distribution of individuals around the age cutoff and on the distribution of other predetermined covariates above and below the cutoff. Surprisingly, we see indications that the identifying assumption may be violated. In particular, we see substantial discontinuities in several predetermined covariates. Although all results are robust to controlling for these predetermined covariates, this finding raises severe concerns about the possibility of additional unobserved differences.

Overall, the results of our analysis appear promising for the possibility of using profiling tools to help move workers out of unemployment faster. At the same time, the potential violations of the identifying assumptions preclude us from making any solid conclusions. Additional evaluation of profiling tools are necessary, preferably using an explicit randomized controlled trial.

This paper contributes to a literature that relies on profiling tools in the labor market. Typically, a person's individual characteristics are utilized as input for a statistical model to identify common patterns and to predict the risk of, e.g., becoming long-term unemployed. This information is then used to assign specific treatments, such as active labor market policies (Frölich, 2008; Behncke et al., 2009; Staghøj et al., 2010). However, in contrast to our case, the outcome of the profiling process is usually not revealed to the individual itself. Our paper provides a first attempt to utilize the outcome of the statistical profiling for the purpose of an econometric policy evaluation.

The aim of the Danish profiling tool is to better inform unemployed job seekers and their case workers about the risks of long-term unemployment. Accordingly, our paper also relates to the literature on biased-beliefs among job seekers and on case workers. Several recent papers have shown that job seekers may have biased beliefs about their job prospects (see e.g. Spinnewijn, 2015; Kassenboehmer and Schatz, 2017; Mueller et al., 2021) and that this can have implications for job finding. For recent work on the role of caseworkers for the reintegration of unemployed workers into the labor market and in targeting active labor market policies, see e.g. Lechner and Smith (2007), Behncke et al. (2010), Schiprowski (2020) or Schmieder and Trenkle (2020).

Moreover, we add to the growing literature on online labor market interventions that aim to target the specific needs of the individual job seeker (see e.g. Horton, 2017; Belot et al., 2018). In particular, preventing at-risk job seekers from ending up in long-term unemployment is primary concern of public policy since the societal and individual costs arising for individuals who do not manage to leave unemployment within a few months are potentially very high. This include direct transfer payments and lower tax revenues, but also indirect costs like the individual's risk of losing human capital (see e.g. Jacobson et al., 1993; Neal, 1995) or suffering from health issues (e.g. Sullivan and von Wachter, 2009) in the long-run.

The remainder of the paper is organized as follows. In the next section, we discuss the institutional setting and the details of the statistical profiling tool. Section 3 presents the data, while Section 4 discusses the empirical strategy. Finally, Section 5 shows the empirical results, while Section 6 concludes.

2 Institutional Setting

Unemployment insurance (UI) benefits in Denmark are organized in a voluntary opt-in system, where unemployed workers are eligible to receive UI benefits for a period of up to two years, if they have paid contributions for at least 12 months within the last three years. The level of monthly benefits is fixed at 90% of prior wage income up to a ceiling of 18,866DKK (3,075 USD) per month before taxes. Around 85% of the Danish wage-earners participate in the system and pay contributions, while around 75% of the actual benefit recipients receive the maximum amount of UI benefits.

2.1 Profiling Tool and Information Treatment

The aim of this paper is to evaluate the effect of an information treatment, which informs job seekers and their case workers about the predicted individual risk of long-term unemployment. When becoming unemployed, all job seekers in Denmark are required to register at the online portal of the Danish Employment Agency (*jobnet.dk*) to receive unemployment insurance benefits. Upon registering, they are asked to answer a voluntary online survey.[1] The survey is advertised as preparation for the first meeting with the caseworker and covers 12 questions related to the field of education, planned job search strategies, job preferences and expectations about employment prospects. About 40 pct. of newly unemployed job seekers answer the survey[2], and the majority answers during the first few weeks of the unemployment spell (see figure A1 in appendix A).

Upon completion of the survey, the individual's survey answers are combined with additional administrative data and are then fed into a statistical profiling tool. As described further below, the output of the profiling tool is a prediction for whether the individual faces a high risk of becoming long-term unemployed as defined by being unemployed for six months or more.

UI benefit recipients are immediately informed about the prediction from the profiling tool. For all job seekers, the survey ends with a message on the screen thanking them for their participation and briefly mentioning how their survey answers will be useful for an eventual meeting with their case worker. For individuals flagged as high risk, however, the screen contains an extra paragraph. This additional paragraph informs the job seeker about the increased risk of

[1]The survey is available for the first twelve weeks after registering as unemployed for all individuals who had not received unemployment benefits within the last 180 days.

[2]Considering new unemployment entries from July 2015 to August 2017.

becoming long-term unemployed, with the following statement:[3]

> *Your characteristics indicate that it could be challenging for you to get back into employment. Our analysis shows that persons with similar characteristics as you, who were previously unemployed, had more difficulties to find new employment compared to other unemployment benefit recipients. Therefore, we recommend you to discuss and plan which measures to take with your caseworker in order to bring you back to employment quickly.*

In addition to this message to the job seeker, the outcome of the profiling tool, i.e. whether the individual is classified as high risk, is also revealed digitally to the caseworker who is supporting the unemployed during the job search process. Any effect of the profiling tool on job outcomes may thus come either from the job seeker and/or the case worker responding to the information provided by the profiling tool.

2.2 Details of the Profiling Tool

The tool, developed by the Danish Employment Agency (STAR), is based on a simple Machine Learning algorithm that was trained to predict long-term unemployment from observable characteristics. The set of observable characteristics used in the (final) algorithm are age, ethnicity, former employment, former industry, education, as well as the past receipt of unemployment and welfare benefits. The specific algorithm underlying the Profiling Tool is based on a decision-tree that determines whether a UI benefit recipient has a high risk by sequentially checking the value of the different variables used in the algorithm. The specific structure of the tree - cutoff values and which variables to examine when - was determined using a standard training procedure applied to administrative data on UI recipients for the years 2011 to 2013.

For our empirical approach, it is crucial that we are able to replicate the decision tree underlying the profiling tool. Figure 1 shows a representation of the actual decision tree underlying the Danish profiling tool.[4] To classify whether a given UI benefit recipient is classified as being of high risk, we first check whether the individual was employed at some point during the past

[3]The full messages sent to low and high risk individuals are shown in Appendix figure A2.

[4]Note that a given decision tree does not have a unique graphical representation. A tree that first splits observations into two groups by gender and then splits both groups by whether they have a college degree will be equivalent to a tree that first splits by college degree and then by gender. As we shall see below, the specific representation shown in Figure 1 is particularly useful for the purpose of our empirical analysis.

year[5]. If not, we proceed along the left branch and next check the person's field of education. If yes, we proceed along the right branch and check the previous industry. Focusing further on an individual with past employment along this right branch, we next see that if the individual was employed in the construction industry (category 3), we move along a branch with no further variables to be checked: a so-called terminal node or leaf. When reaching such a point, the classification is completed and the individual is either classified as high risk or not, depending on the specific terminal node that was reached. In the specific case of a UI recipient with previous employment in the construction industry, we can see that the individual is not classified as high risk (as signified by (L)ow at the end of the relevant branch).

2.3 Age Discontinuities in the Profiling Tool

As noted previously, our empirical analysis will be based on a regression discontinuity design (RDD), leveraging age discontinuities in the profiling tool. With this in mind, we constructed the particular representation of the decision tree shown in Figure 1. It can be seen that as we progress through the tree, we either end up at a terminal node without considering the individual's age (see blue nodes) or we are required to examine the individual's age against one or more cut-offs as the last variable before reaching a terminal node (see red nodes).[6]

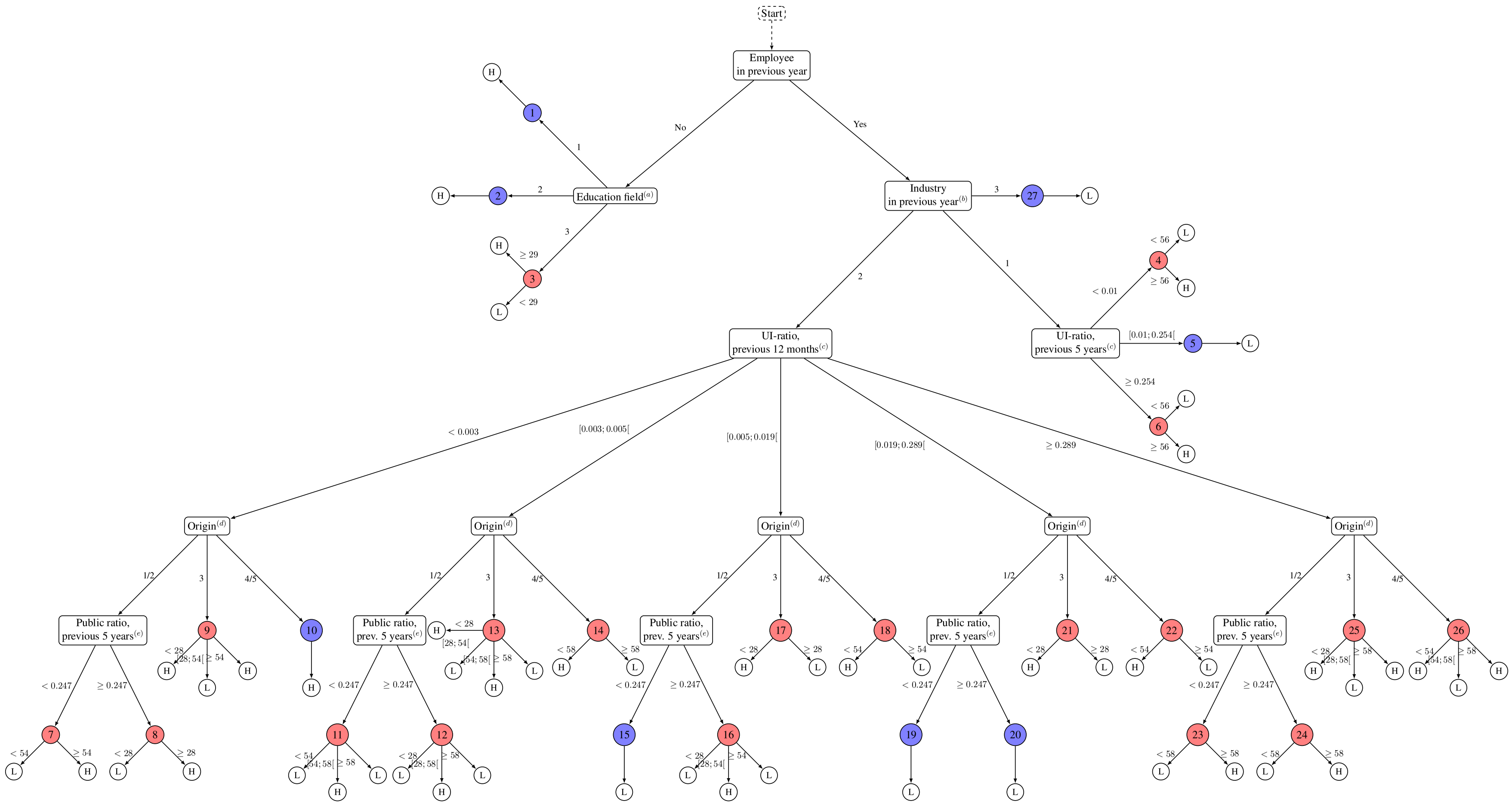
As it can be seen from Figure 1, with data on all input characteristics for a given individual, we can determine how they will progress through the tree. This allows us to partition the full sample into 27 subgroups based on their relevant input characteristics except for age (indicated by the red and blue circles numbered from one to 27). For eight of the 27 subgroups, age is not identified as a relevant input factor (blue nodes), while for the other 19 subgroups (red nodes), the treatment status is determined by examining the individual's age against at least one cutoff. As such, each of the 19 red subgroups lends itself to one (or more) RDD designs.[7] This is the basis for our identification strategy. To ensure a sufficient number of observations around the cutoff, we do not analyze the 19 different subgroups individually, but pool those with similar

[5]Defined by whether the jobseeker had registered wage income in any of the previous 12 months (Eindkomst)

[6]As noted, there are multiple representations of the same decision-tree (see footnote 4). The training procedure used to construct the decision tree provides a simple representation in which the same variable is typically not to be checked across many different branches (see the original representation of the tree in appendix table A1). In Figure 1, we inverted this original representation of the tree such that age is always the last variable examined along all branches whenever possible.

[7]Because some of the red circles involve checking age against two cutoffs, some of them can be used naturally for two separate RDDs.

Figure 1: Partition into subgroups



Note: We partition the sample into 27 subgroups, each represented by a colored dot. Blue dots represent subgroups with no age cut-off, red dots represent subgroups with at least one age cut-off. Each of the red subgroups are further partitioned into risk groups represented by white dots, where "H" denotes high-risk and "L" denotes low-risk.

$^{(a)}$*Education categories:* 1 - humanities, religion, aesthetic or missing education; 2 - social work, office, non-commercial, or pedagogical training, fishery, agriculture or food, scientific education 3 - Manufacturing and crafts, health, transportation and communications

$^{(b)}$*Industry categories:* 1 - public administration, health, teaching, employed in unknown activity, manufacturing, mining and quarrying, utilities, agriculture, forestry or fishery; 2 - trade, logistics, business services, culture, leisure or other services, real estate, information and communication, financial or insurance services. 3 - Construction

$^{(c)}$ *UI-ratio* defined as the fraction of days spent on UI-benefits in the previous 12 months or 5 years.

$^{(d)}$*Origin categories:* 1 - Danish, 2 - descendent (Western country), 3 - immigrant (Western country), 4 - descendent (non-Western country), 5 - immigrant (non-Western country)

$^{(e)}$ *Public ratio* defined as the fraction of days spent on public transfers in the previous 5 years.

See a description of each of the 27 subgroups in Appendix Table A2 age cutoffs (see Section 4.1).

3 Data

The empirical analysis builds on a sample of all newly registered UI benefit recipients between July 2015 and August 2017 who answered the online survey discussed in Section 2.1. Note that only survey participants potentially receive information about their individual risk of becoming long-term unemployed. The sample period starts in July 2015 when the information treatment was implemented and ends just before the underlying algorithm was revised in September 2017.[8]

If individuals respond to the survey more than once within our observation period (because they enter unemployment multiple times), we only consider the first incidence. This results in an overall sample of 168,262 individuals. We directly observe their answers to the survey questions and the outcome of the profiling model, i.e. whether they are flagged as high risk and receive the information treatment. We then link the survey information to additional administrative data, which allows us to reconstruct all relevant input variables of the profiling model (see appendix B.1 for an overview) and to observe realized labor market outcomes. The next few sections expand on these data.

3.1 Input Variables for the Statistical Model

Reconstructing the input variables that are used to predict the job seeker's risk of long-term unemployment is crucial for our analysis as it allows us to determine whether any of the potential age discontinuities is relevant for the individual job seeker (i.e. it allows us to identify the 27 subgroups discussed in Section 2.3).

Some of the input variables are straightforward to obtain. For instance, the field of education is directly observed in the survey and the individual's origin can be obtained from the population register. However, some of the input variables vary over time and are not contained in the survey, which impedes the possibility to measure them accurately. Note that we need to determine the value of each variable *at the exact point in time when the individual answers the survey*. These time-varying input variables include the UI-benefit ratios (fraction of days receiving UI benefits

[8]For the purpose of our identification strategy, it is crucial to consider a sample of job seekers whose treatment status was determined by the same underlying algorithm (see Section 4.2).

in the previous 12 months or 5 years), the public-benefit ratio (fraction of days spent on any type of public benefits in the previous 5 years), the job seeker's employment status in the previous year, the corresponding industry and age.

We use information regarding the survey date and the full individual history in the administrative data to reconstruct the value of the these time-varying input variables at the time of the survey. For example, given the survey date, detailed information regarding UI benefit payments allow us to calculate the fraction of days receiving UI benefits in the 12-month period prior to the survey. Similarly, data on the exact birthday of each individual allows us to reconstruct the individual's age at the time of the survey.

Two problems, however, imply that our reconstructed measures of time-varying input variables will be subject to some measurement error. First, some of the input variables rely on administrative data that may have been updated since the time of the survey. For example, delayed reports from municipalities and UI funds imply that the number of past days receiving UI benefits that we observe in the data may differ from the corresponding number observed in the same data set at the time of the survey. Since historic versions of the registers are not stored, this source of measurement error seems inevitable. The second problem arises from the fact that, when implementing the statistical model, information about input variables were not updated on the corresponding server on a daily basis. Rather they were only updated weekly. For instance, when considering an individual answering the survey five days after an update, the prediction computed in the profiling tool would be based on the value of her input variables as of five days ago. To overcome this issue and to create accurate measures for the input variables, we attempt to infer the dates at which the input variables were updated on the server (see Appendix B.2). Since this procedure is not exact, some measurement error remains when actual server updates are off by a few days relative to our procedure.

The measurement error with respect to the input variables has two implications for our RDD analysis. First, measurement error in the non-age input variables implies that we will incorrectly assign some individuals to the subgroups shown in Figure 1. This introduces some degree of imperfect compliance in our RDD. Individuals in our analysis who were in fact assigned to a subgroup without an age cutoff (or with a different age cutoff) will be treated (or untreated) regardless of whether their age passes the threshold we are analyzing. Since our data contains information about whether each individual was *actually* flagged as high risk, however, we can correct for this imperfect compliance in the usual way. We simply use the information about being flagged as high risk as a treatment indicator in a Fuzzy RDD (see Section 4.2).

Second, measurement error regarding an individual's age at the time of the survey in principle introduces a very small amount of measurement error in our running variable. In particular, this precludes us from measuring age relative to the cutoff more finely than at the weekly level. As the measurement error is more prevalent in week 0, we simply drop this week and estimate a so-called donut RDD (see Barreca et al., 2011; Almond and Doyle, 2011).

3.2 Other Covariates and Labor Market Outcomes

In addition to the specific input variables used in the profiling tool, our data contains information on a rich set of covariates. In particular, from the survey and the administrative data we construct measures of ethnicity, gender, fertility, the level and field of education, the previous industry, the number of employers over the past year, the UI fund membership as well as working hours and labor earnings in the previous year. We particularly use these pre-determined characteristics when testing the validity of the identifying assumptions.

Our administrative data also allows us to measure a range of relevant outcomes that may be affected by receiving the information from the profiling tool. Since the profiling tool aims to predict the risk of remaining unemployed for 26 weeks, we take (un)employment status after 26 weeks as our natural main outcome. In addition, however, we present results for each of the first 52 weeks to examine the dynamic pattern of effects.

4 Empirical Strategy

Our empirical strategy exploits the age discontinuities in the decision tree underlying the profiling tool. As discussed in Section 2.3 and 3.1, the reconstructed non-age input variables used in the decision tree allow us to split the sample into 19 subgroups, in which the treatment status (i.e. whether the job seeker is classified as of high risk) depends only on the individual age relative to some cutoff. In the following sections, we discuss how this allows us to implement the RDD.

4.1 Pooling of Subgroups with Different Cutoffs

In principle, we could conduct separate RDD analyses for each of the 19 subgroups (and each of the relevant cutoffs). To ensure a sufficient number of observations around the cut-off in our RDD, however, we opt to pool multiple subgroups and estimate an average effect across these pooled groups. The pooling is based on two objectives: First we want to ensure that there is a sufficiently large number of observations in each group. Second, to make the average effects meaningful, we want to avoid pooling too many groups where the profiling tool may have very different effects.

Given these two goals, two natural pooled samples arise. First, pooling subgroups 3 and 8 gives rise to a sample of young native Danes who face cutoffs in the decision tree in their late-twenties (at age 28 or 29). We refer to this as the *young* sample. Second, pooling subgroups 4, 6, 7 and 9, gives rise to a sample of older Danes, who face cutoffs at age 54 or 56. We refer to this as the *old* sample. We conduct our RDD analysis separately for the young and old sample and discard all other subgroups. The discarded subgroups either include very few individuals around the cutoff and/or contain a highly selected set of individuals that may respond very differently to the information provided by the profiling tool (subgroup 9 for example include exclusively descendants of western immigrants, see Appendix Table A2).

4.2 Implementation of the Regression Discontinuity Design

In both the young and the old sample, we estimate the treatment effect by comparing individuals just below and above an age cut-off. To take account of the fact that different individuals in each sample contain different age cut-offs (28, 29, 54 or 56 years), we create a running variable that measures the distance to the relevant age cutoff. Specifically, we define RV_i as the difference between the individual's actual birthday and the first potential birthday what would have caused them to be flagged as high risk. That is, we take the survey date as given and ask when the individual should have been born in order to receive treatment.[9] We refer to Appendix B.3 for additional details of how we calculate the running variable.

[9]Because the day at which the individual's age is calculated differ from the day at which the survey is taken, there are in principle two ways to measure the RV. We take the survey date as given and ask when the individual should have been born in order to be treated. However, one could also take the birthday as given and ask when the individual should have filled out the survey to be treated. The difference between these two distance measures is minor, since there only a few days between the survey and the calculation day. Note that we will measure the running variable in weeks, exactly because the input variables were calculated on a weekly basis and we have imperfect information about the exact day of the week that they were calculated.

In both samples, all individuals above an age cut-off are *predicted* to be treated, whereas all individuals below an age cut-off are predicted not to be treated. For some individuals, however, the predicted treatment will deviate from *received* treatment. This measurement error regarding the treatment status is caused by the measurement error in historical input variables described in Section 3. Had the reconstruction of input variables been perfect, there would be no measurement error regarding the treatment status. All individuals would have been placed in a subgroup with a cutoff perfectly corresponding to the one they were profiled according to. Hence, the predicted treatment would correspond to the received treatment. Due to the imperfect reconstruction, however, we place some individuals in the wrong subgroup and therefore also assign them to an incorrect cutoff. Depending on their age, this can cause received treatment to differ from predicted treatment.[10]

It appears plausible that such a measurement error arising from the calculation of the historical input variables is non-systematic as it is basically created by the frequency of database updates on the corresponding server of the Employment Agency. Therefore, we can rely on a fuzzy RD design in order to estimate the causal effect of the information treatment on labor market outcomes, where the assignment rule can be used as instrument for the actual received treatment. In particular, we are interested in two effects: 1) the intention-to-treat effect (ITT) and the local average treatment effect (LATE). We obtain the ITT by estimating the following regression equation on the (weighted) sample of individuals with RV_i close to 0 (see next subsection):

$$Y_i = \beta_0 + \beta_1 RV_i + \beta_2 RV_i \times T_i + \tau_{ITT} T_i + \beta_3 X_i + \varepsilon_i, \tag{1}$$

where T_i is an indicator for the predicted treatment (based on constructed input factors), X_i denotes a vector of control variables (see below) and RV_i characterizes the running variable. Note that the running variable is measured in weeks, and that we exclude individuals who fill out the survey in week 0 (the week in which they would actually cross the cutoff) since the measurement error may be particularly pronounced for these observations.

The ITT is the coefficient on predicted - or assigned - treatment, τ_{ITT}. However, the ITT characterizes the average effect in the sample without taking into account that not all assigned

[10]For instance, consider an individual of age 55 who, according to her historical input variables, belongs into subgroup 9. In subgroup 9, all individuals above 54 are treated and hence the individual would actually receive the treatment. If we, however, miscalculate the UI-benefit ratio in the previous 12 months, e.g. find it to be above 0.005 instead of below 0.003, we would place her in subgroup 17. In this subgroup, individuals below age 28 are treated and hence we would predict that she is not treated, even though she did receive the treatment.

individuals actually receive the treatment. Therefore, we also estimate the LATE, which can be interpreted as the effect on those who actually receive the information treatment. We do this by applying the following two-stage least squares (2SLS) estimator applied to a (weighted) sample of individuals with RV_i close to 0 (see next subsection):

$$D_i = \gamma_0 + \gamma_1 RV_i + \gamma_2 RV_i \times T_i + \tau_{FS} T_i + \gamma_3 X_i + u_i \tag{2}$$

$$Y_i = \delta_0 + \delta_1 RV_i + \delta_2 RV_i \times D_i + \tau_{LATE} T_i + \delta_3 X_i + v_i, \tag{3}$$

where τ_{FS} represents the jump in the probability of actually receiving the treatment D_i at the cutoff and τ_{LATE} can be interpreted as the LATE for the affected population around the cutoff.

4.3 Optimal Bandwidth, Estimation and Inference

As laid out in the previous section, our implementation of the RDD is based on estimating a linear regression, locally in a neighborhood around the cutoff (for individuals with RV_i close to 0). This requires us to choose a bandwidth for which individuals get included, as well as weights for each individual as a function of the distance to the cutoff. Following Cattaneo et al. (2017), in our main specification, we use a triangular kernel and select our bandwidth by relying on the optimal bandwidth selector that reduces the mean squared error. Moreover, we conduct inference using robust confidence intervals (see Calonico et al., 2014; Cattaneo et al., 2019) that take into account the bias-variance trade-off when choosing the optimal bandwidth. As we show in the appendix (Table A5 and A6), none of our main results are sensitive to alternative estimation choices in the implementation of the RDD.

4.4 Ex-ante Validity of Identifying Assumptions

To ensure that the RDD approach indeed identifies the causal effect of the information treatment, we have to assume that individuals on both sides of the cut-off are similar with respect to all relevant characteristics apart from the treatment assignment. This implies that individuals should not perfectly manipulate their value of the running variable, i.e. their age when they take the survey, in order to receive or avoid the treatment. Although job seekers cannot manipulate their date of birth, one could be potentially concerned that job seekers choose to fill out the survey on a particular day that would place them on one side of the cut-off or the other. However, this seems to be very unlikely since job seekers are not informed that they are profiled and the associated consequences in advanced. Note that we exclude all re-entries into unemployment, i.e. job seekers who have been profiled before.

Moreover, even if they would know that the profiling takes place, the exact algorithm is not common knowledge. Its complexity in combination with the fact that the underlying data sources are updated with a time lag makes it nearly impossible that strategic behavior could allow them to deliberately receive or avoid the treatment. Finally, we are not aware of any other policies involving discontinuities at the relevant age cutoffs. In summary, it seems reasonable to assume that our RD approach allows us to identify the causal effect of the information treatment.

5 Results

We now present our empirical results. We preface our RDD analysis by presenting descriptive statistics for the two samples involved (the young and old sample). We then proceed by presenting the first stage for the Fuzzy RDD analysis and validity checks of the identifying assumptions. Finally, we show the estimated effects of the information treatment on job seekers' subsequent labor market outcomes.

5.1 Descriptive Statistics

Table 1 shows summary statistics for the young and the old sample considering only individuals within the optimal bandwidth used in our main specification. Besides the age difference of about 30 years, which is a consequence of the sample construction, we observe some further differences regarding individuals' background characteristics. Job seekers in the young sample tend to have a higher level of education and tend to be educated in the fields health or manufacturing more often, whereas individuals in the old sample tend to hold a pedagogic education or an education in the social sciences, administration or trade. Moreover, individuals in the young sample tend to have lower labor market attachment, e.g. received public transfers more often and a smaller share of job seekers has been regularly employed before, compared to the old sample. Likewise, individuals in the old sample have a higher wage income and worked more hours in the previous year.

Table 1: Summary Statistics

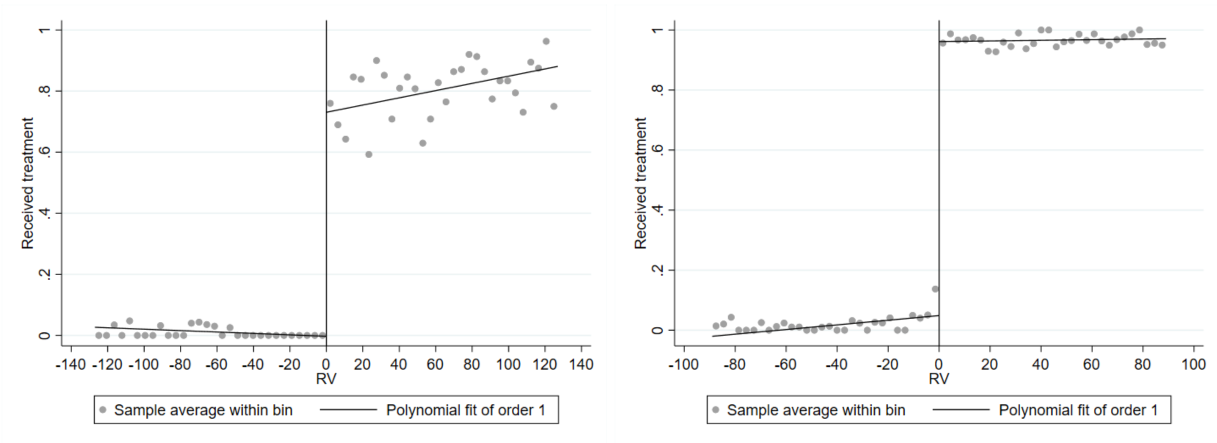
	Young		Old	
	Mean	SD	Mean	SD
Demographics				
Age	28.026	(1.484)	54.609	(1.473)
Male	0.382	(0.486)	0.399	(0.490)
Migrant	0.134	(0.341)	0.067	(0.250)
Western	0.933	(0.251)	0.975	(0.156)
Copenhagen	0.274	(0.446)	0.221	(0.415)
Pregnant at time of survey	0.081	(0.273)	0.001	(0.025)
Married	0.161	(0.368)	0.810	(0.393)
No children	0.699	(0.459)	0.638	(0.481)
Education level				
0 Qualifying or no information	0.013	(0.113)	0.001	(0.025)
1 Primary	0.094	(0.292)	0.185	(0.389)
2 Upper secondary	0.079	(0.270)	0.058	(0.234)
4 Bachelor programmes	0.075	(0.263)	0.015	(0.120)
3 Vocational educ. & training	0.251	(0.434)	0.424	(0.494)
5 Short cycle education	0.092	(0.290)	0.065	(0.246)
6 Vocational bachelors education	0.256	(0.437)	0.184	(0.388)
7 Higher education	0.140	(0.347)	0.068	(0.252)
Education field, survey				
Missing field	0.256	(0.436)	0.540	(0.498)
Humanities	0.016	(0.125)	0.029	(0.168)
Non-vocational	0.011	(0.104)	0.028	(0.165)
Manufacturing, craft	0.197	(0.398)	0.039	(0.194)
Natural science	0.007	(0.085)	0.039	(0.193)
Pedagogic	0.007	(0.082)	0.114	(0.318)
Social science, administration, trade	0.055	(0.228)	0.151	(0.358)
Health	0.376	(0.485)	0.040	(0.196)
Transport & communication	0.073	(0.261)	0.010	(0.099)
Agrarian	0.002	(0.043)	0.010	(0.101)
Transfers				
UI-ratio prev. 12 months	0.018	(0.063)	0.019	(0.081)
UI-ratio prev. 5 years	0.107	(0.135)	0.068	(0.139)
Public benefit ratio prev. 5 years	0.213	(0.210)	0.114	(0.168)
Labor market history				
Danish Trade Union Assocition	0.559	(0.497)	0.599	(0.490)
Academics Association	0.202	(0.401)	0.090	(0.287)
Non-organized UI-funds	0.186	(0.389)	0.309	(0.462)
Any employment in year t-1	0.376	(0.485)	0.981	(0.138)
Number of employers in year t-1	0.494	(0.749)	1.260	(0.589)
Wage in year t-1 (10,000 DKK)	0.768	(1.122)	3.560	(1.988)
Working hours in year t-1 (divided by 10)	1.128	(1.559)	3.388	(0.628)
Prev. industry				
Other	0.012	(0.107)	0.122	(0.327)
Business services	0.095	(0.293)	0.114	(0.317)
Trade and transportation	0.218	(0.413)	0.193	(0.395)
Manufacturing	0.000	(0.000)	0.147	(0.354)
Information, communication	0.024	(0.153)	0.049	(0.216)
Employment not known	0.626	(0.484)	0.000	(0.000)
Culture, leisure, other service	0.026	(0.158)	0.032	(0.177)
Public administration, education, health	0.000	(0.000)	0.344	(0.475)
Obs	1636		4863	
Bandwidth	127		90	

Notes: Summary statistics for individuals falling within the optimal bandwidth for the main specification (unemployed in week 26 as the outcome). Demographics, level of education, labor market history, transfers and previous industry are based on Danish registers, whereas field of education is based on survey behind profiling Tool.

5.2 First Stage

Since we estimate a fuzzy RDD, where the treatment assignment T_i could differ from the actual treatment D_i, both must be sufficiently correlated in order to avoid issues due to weak instruments (Staiger and Stock, 1997). As shown graphically in Figure 2, the assigned treatment is strongly correlated with the actual treatment, which shows that we are able to correctly reconstruct the relevant input factors of the profiling model for the large majority of individuals. For instance, for the old sample, being assigned to the treatment increases the likelihood of actually receiving the information treatment by about 90 percentage points, while the first stage effect is slightly lower (+73 percentage points) for the young sample.[11]

Figure 2: First stage of regression discontinuity design



(a) Young (b) Old

Note: The sample (within the optimal bandwidth) has been sliced in 30 equally spaced bins on either side of the cutoff. Each dot represents sample averages by bins. The solid line represents lines fitted to the underlying data, estimated with a linear polynomial, triangular kernel and using the optimal bandwidth.

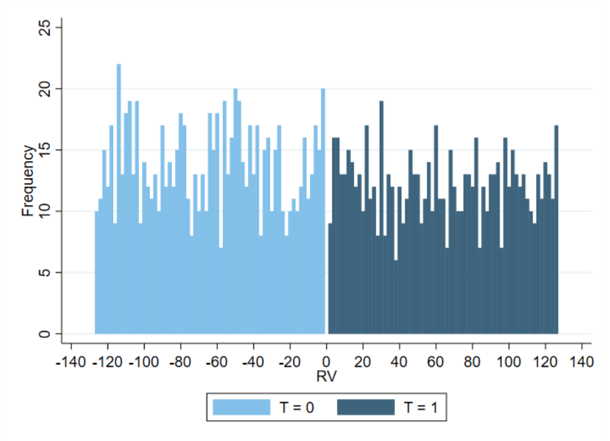
5.3 Validity of Empirical Strategy

Next, we empirically test the validity of the underlying identification assumption. First we consider the continuity of the density function for the running variable around the cutoff (see McCrary, 2008). Figure 3 shows histograms for the distribution around the cutoff in both the young and old sample. In both histograms the distribution in the neighborhood around the cutoff shows few differences compared to the overall distribution. We note however that formal

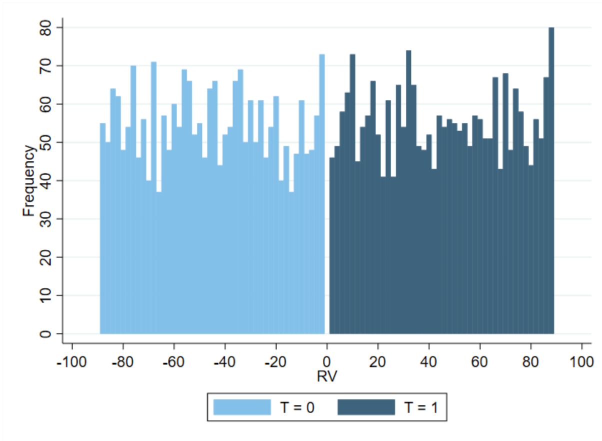
[11]The corresponding RD regression relies on a local linear polynomial with triangular kernel weights and the optimal bandwidth from the main specification using an indicator for whether the individual is still unemployed in week 26 as outcome variable.

statistical test do in fact show evidence of a discontinuity for the old sample at the 10 percent level ($p = 0.079$ for the old sample, while $p = 0.667$ for the young).

Figure 3: Density around cutoff



(a) Young



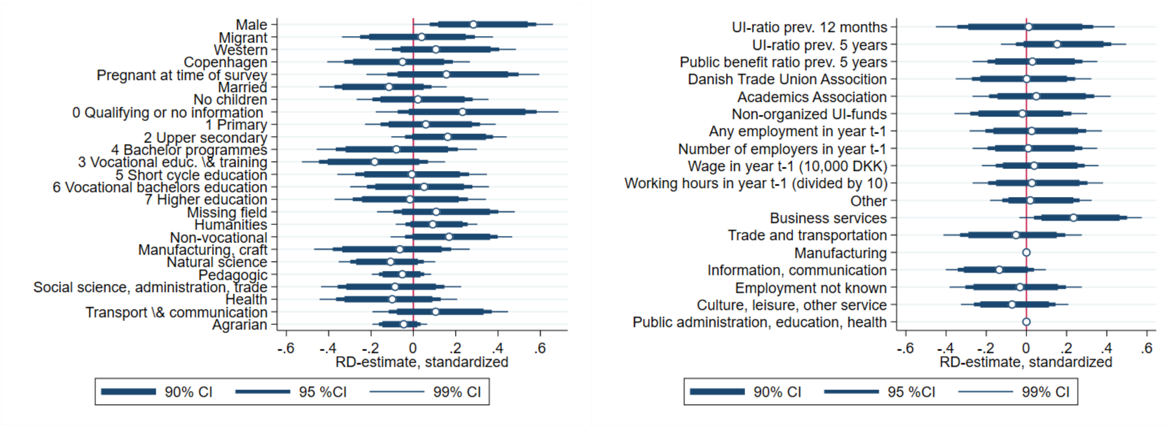
(b) Old

Note: Depicted is the number of observations in two-week bins within the optimal bandwidth (using unemployed in week 26 as outcome). Test for continuous density around the cutoff: $P_{young} = 0.667$ ($N = 1,636$); $P_{old} = 0.079$ ($N = 4,863$). Density test is based on specification including first order polynomial.

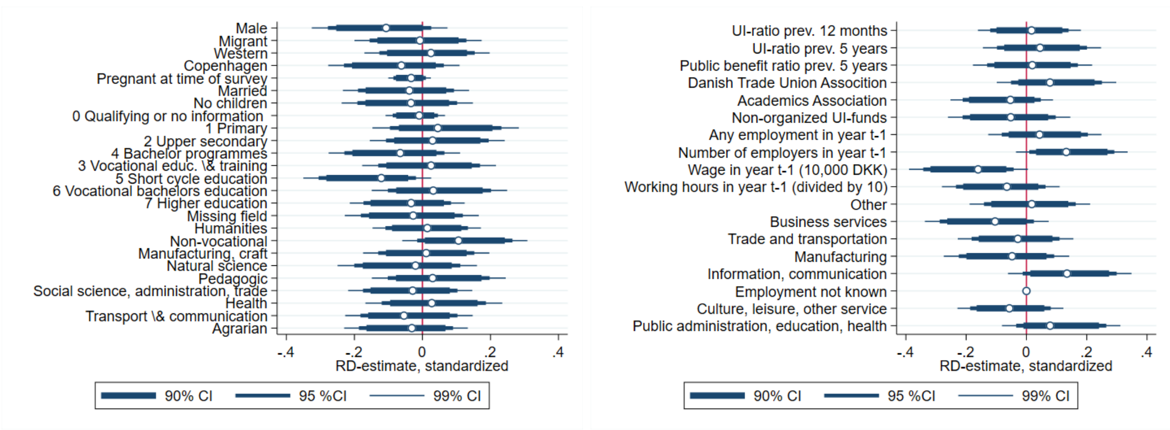
As a second validity check, we run (ITT) RD regressions using pre-determined covariates as the outcome of interest. This allows us to test whether there exists discontinuities with respect to other characteristics that should not be affected by the information treatment. Figure 4 summarizes the RD-estimates and robust 90, 95 and 99% confidence intervals for the pre-determined covariates. In Figure 4 the covariates have been standardized so that the estimated discontinuities are measured in standard deviations. The non-standardized estimates are shown in appendix table A3 and A4. All regressions are specified as depicted in Equation 1 using triangular kernel weights and the optimal bandwidth calculated for the main outcome variable (being unemployed in week 26).

Looking across Figure 4, we note that while estimates are small and insignificant for most baseline covariates, there are in fact several large estimates in both samples, some of which are also highly significant. In the young sample, individuals just above the cutoff are 14 percentage points more likely to be male and this difference is significant even at the 1 percent level. At the 5 percent level, individuals just above the cutoff are also significantly more likely to have worked in business services within the last year. In the old sample, individuals just above the cutoff are significantly more likely to hold a short-cycle education degree, had significantly more employers over the past year and significantly lower wages over the past year (all these coefficients are significant at the 5 percent level).

Figure 4: Covariate balance



(a) Young



(b) Old

Note: RD-estimates from an RD-regression using the predetermined covariate as the outcome of interest. The RD regressions use a linear polynomial, a triangular kernel and the optimal bandwidth from the main specification (unemployed in week 26 as outcome).

Given the strong theoretical reasons that the identifying assumptions should be satisfied in the RD setting (see Section 4.4), these discontinuities are very surprising. In principle, they could of course reflect random variation (type I errors), however, the results of the validity checks presented above also raise severe concerns about the validity of our RDD design.

In the rest of the empirical analysis, we present RDD estimates after controlling for all covariates that show significant discontinuities in Figure 4.[12] If the discontinuities in Figure 4

[12]The problematic covariates are included linearly and we use the bandwidth optimal for covariate inclusion.

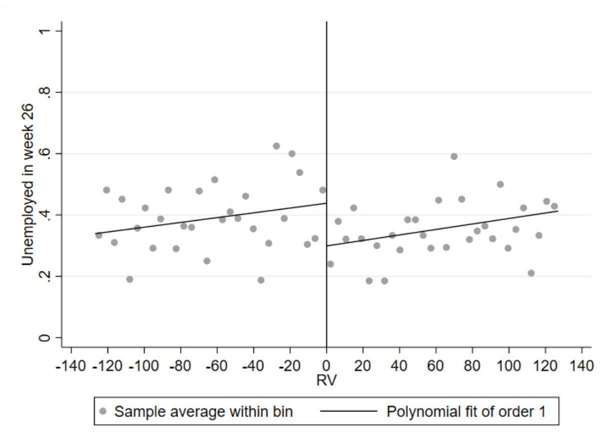
represent the only systematic differences between individuals that fall on either side of the RDD cutoff, the use of these control variables allow us to estimate causal effects. In fact it turns out that the inclusion of these covariates does not affect the estimated effects presented later (see table 2). However, the findings in Figure 4, obviously still raise severe concerns about other, unobserved differences that we can not address. This major concern needs to be kept in mind when interpreting the estimated effects over the next sections.

5.4 Baseline Results

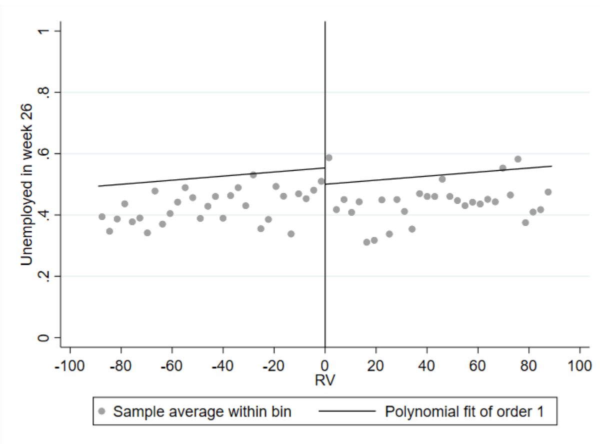
Figure 5 shows RDD plots for the probability of being registered unemployed in week 26 after receiving the information treatment. Each dot represents the average for individuals within bins of the running variable, while the solid lines represent an RD-regression fitted to the underlying data. When considering the young sample (see Figure 5a), we can see a marked discontinuity at the age-cutoff. Treated individuals just above the cutoff are about 14 percentage points less likely to still be unemployed compared to the non-treated just below the cutoff. The corresponding RD estimates from our preferred specification are presented in column (2) of Table 5 and show that the ITT is statistically significant at the 1%-level. Relative to the sample average, the treatment is estimated to reduce the likelihood of being unemployed by about 34.5%. For the old sample (see column (4) of Table 2), the magnitude of the effect is substantially smaller, but we still find a reduced likelihood of being unemployed of about six percentage points (-14% relative to the sample average), which is statistically significant at the 10% level. Finally, looking at columns (1) and (3) of Table 2 we see that our inclusion of control variables do not actually affect our results.

Appendix figure A3 shows that after controlling for the problematic covariates, we see no discontinuities in any of the other baseline covariates.

Figure 5: RD estimates: unemployed in week 26 after survey



(a) Young



(b) Old

Note: The sample (within the optimal bandwidth) has been sliced in 30 equally spaced bins on either side of the cutoff. Each dot represents sample averages by bins. The solid line represents lines fitted to the underlying data. These are is estimated while controlling for sample-specific covariates and with a linear polynomial, triangular kernel and using the optimal bandwidth (using unemployed in week 26 as the outcome). The sample-specific covariates are gender and business services for the young sample, while for the old sample it is short cycle tertiary education, number of employers and previous wages.

In Table 5, we also present the estimates of the Fuzzy RD approach. It can be seen that the FRD (LATE) for the young sample is slightly larger (-17 percentage points) than the ITT, while it is slightly smaller (-4 percentage points) for the old sample.[13] Overall, the results indicate that the information treatment affects job seekers' behavior in a way that reduces the likelihood of remaining registered as unemployed. Hence, one could speculate that the profiling tool may serve its purpose as it seems to reduce the specific risk that was predicted by the underlying statistical model based on historical data. In this context, it seems to be more effective for young job seekers (below age 30) than for older job seekers (above age 50). There could be various explanations for these differential effects with respect to age. For instance, they may reflect the higher responsiveness of young people to digital information or their greater labor market flexibility

In the appendix, we test the robustness of our results to various different specifications, i.e. including/excluding controls, using an alternative definition of the running variables, considering a higher order polynomial, different kernel weights and clustered standard errors. As shown in Table A6 and Table A5, the main estimates are very robust across these different specifications.

[13]It should be noted that the fuzzy RDD estimates are obtained with a different (optimal) bandwidth than the ITTs, which implies that the LATE is not necessarily equivalent to the ratio of the ITT and the first stage estimates.

Table 2: Unemployed in week 26

	Young		Old	
	(1)	(2)	(3)	(4)
ITT	-0.138***	-0.139***	-0.057*	-0.053*
	(0.052)	(0.052)	(0.032)	(0.032)
FS	0.729***	0.724***	0.902***	0.902***
	(0.047)	(0.047)	(0.017)	(0.017)
FRD	-0.170*	-0.170*	-0.040	-0.038
	(0.100)	(0.100)	(0.045)	(0.045)
Controls	NO	YES	NO	YES
Polynomial	1	1	1	1
Kernel	Triangular	Triangular	Triangular	Triangular
Clustering	NO	NO	NO	NO
Mean dep. var	0.400	0.400	0.420	0.420
Effective obs (ITT)	1636	1636	4863	4863
Bandwidth (ITT)	127	127	90	90
Robust 95% CI (ITT)	[-.271 ; -.029]	[-.272 ; -.031]	[-.137 ; .013]	[-.134 ; .017]
Effective obs (FS)	859	859	3163	3110
Bandwidth (FS)	67	67	59	58
Robust 95% CI (FS)	[.623 ; .822]	[.618 ; .817]	[.862 ; .934]	[.862 ; .935]
Effective obs (FRD)	859	859	3163	3110
Bandwidth (FRD)	67	67	59	58
Robust 95% CI (FRD)	[-.388 ; .031]	[-.389 ; .032]	[-.138 ; .053]	[-.136 ; .055]

Notes: RD-coefficients with 'unemployed in week 26' as the outcomes of interest. All coefficients are estimated with a linear polynomial, triangular kernel and the optimal bandwidth. Note that we allow the optimal bandwidth to be different for the ITT and IV specification. Hence, the FS and ITT coefficients are estimated using different bandwidths, and therefore a simple division of ITT by FS does not give the FRD (LATE) estimate. Controls refer to gender and business services for the young sample, while for the old sample it is short cycle tertiary education, number of employers and previous wages.

5.5 Employment Prospects

Next, we want to investigate whether the reduced likelihood of being unemployed translates into improved employment prospects. Therefore, Table 3 shows the effect on three additional outcome variables related to paid employment six months after the initial survey: (i) an indicator for whether the individual has any paid employment, (ii) the average weekly working hours and (iii) the total labor earnings. All three variables are measured in the corresponding calendar month (six months after the survey). Interestingly, for both the young and the old sample, the estimated negative effect on unemployment is *not* accompanied by any significant positive effects on employment outcomes. The information treatment has no statistically significant effect on any of the outcome variables related to paid employment and the size of the estimated effects is substantially smaller than the results presented for unemployment in Table 2.[14]

[14] Again, the overall pattern is very robust with respect to the inclusion, respectively exclusion of control variables as shown in Table A7 in the Appendix A.

Table 3: Employment outcomes in month 6

	Young			Old		
	Employed	Hours	Wage	Employed	Hours	Wage
ITT	0.026	1.889	496.308	-0.005	-0.899	345.977
	(0.063)	(2.136)	(1790.931)	(0.033)	(1.196)	(1002.626)
FS	0.716***	0.719***	0.721***	0.910***	0.911***	0.910***
	(0.050)	(0.049)	(0.048)	(0.014)	(0.013)	(0.014)
FRD	0.012	1.940	-670.464	-0.007	-0.532	152.878
	(0.109)	(3.684)	(3223.569)	(0.037)	(1.215)	(1233.823)
Controls	YES	YES	YES	YES	YES	YES
Polynomial	1	1	1	1	1	1
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Clustering	NO	NO	NO	NO	NO	NO
Mean dep. var	0.504	15.096	11679.866	0.557	16.657	14776.383
Effective obs (ITT)	1225	1253	1360	4648	3914	5403
Bandwidth (ITT)	97	98	107	86	73	100
Robust 95% CI (ITT)	[-.116 ; .179]	[-2.904 ; 7.141]	[-3737.042 ; 4726.755]	[-.091 ; .059]	[-4.036 ; 1.412]	[-2226.22 ; 2444.04]
Effective obs (FS)	798	806	833	4463	4577	4347
Bandwidth (FS)	61	63	64	82	85	80
Robust 95% CI (FS)	[.606 ; .819]	[.61 ; .819]	[.613 ; .818]	[.878 ; .937]	[.879 ; .937]	[.877 ; .937]
Effective obs (FRD)	798	806	833	4463	4577	4347
Bandwidth (FRD)	61	63	64	82	85	80
Robust 95% CI (FRD)	[-.223 ; .246]	[-5.761 ; 9.925]	[-7547.005 ; 6101.063]	[-.093 ; .067]	[-3.5 ; 1.793]	[-2700.757 ; 2677.259]

Notes: RD-coefficients are estimated with a linear polynomial, triangular kernel and the optimal bandwidth. Note that we allow the optimal bandwidth to be different for the ITT and IV specification. Hence, the FS and ITT coefficients are estimated using different bandwidths, and therefore a simple division of ITT by FS does not give the FRD estimate. Controls refer to gender and business services for the young sample, while for the old sample it is short cycle tertiary education, number of employers and previous wages.

One interpretation of these findings is that being flagged as high risk by the profiling tool causes individuals to leave UI for either self-support or other types of public benefits, rather than for employment. This would be consistent with some previous findings from related settings.[15] At the same time, however, we note that the confidence intervals around the estimated effects on employment are large. Accordingly, we are not able to rule out that being flagged as high risk in the profiling tool in fact also has substantial effects on employment.

5.6 Treatment Effects over Time

Next, we extend our baseline analysis to investigate how the treatment effects evolve over time. Therefore, we estimate ITTs on the likelihood of being unemployed, respectively employed for every week relative to the moment when the job seeker received the information treatment.

The results are presented in Figure 6. For the young sample, the negative effect of the information treatment on the likelihood of being unemployed gradually increases between month two and six after filling out the survey and vanishes afterwards. At the end of the observation period (one year after receiving the treatment), there is no difference between treated and non-treated individuals. While individuals marginally flagged as high risk thus exit unemployment

[15]For instance, Card et al. (2007) document a huge spike in exits from UI around the exhaustion of unemployment benefits without being mirrored in employment outcomes.

faster, we see that unemployment exits eventually catch up among individuals who marginally do not get flagged as high risk.

Again, the overall pattern looks very different when considering the likelihood of being employed in a given month. As shown in Panel A.2 of Figure 6, there are no significant employment effects throughout the first year of the unemployment spell. Although there are some sizable (but insignificant) positive effects on the employment probability within the first months of the unemployment spell,[16] they are substantially smaller than the negative effects on the likelihood of still being unemployed.[17]

For the old sample, the reduction of the unemployment probability throughout the first six months is again substantially smaller than for the young sample, while the employment effects are very similar. Hence, the discrepancy between the effects on exits from UI and starting paid employment is particularly pronounced for job seekers who receive the information regarding their personal risk of becoming long-term unemployed at an early stage of their working life.

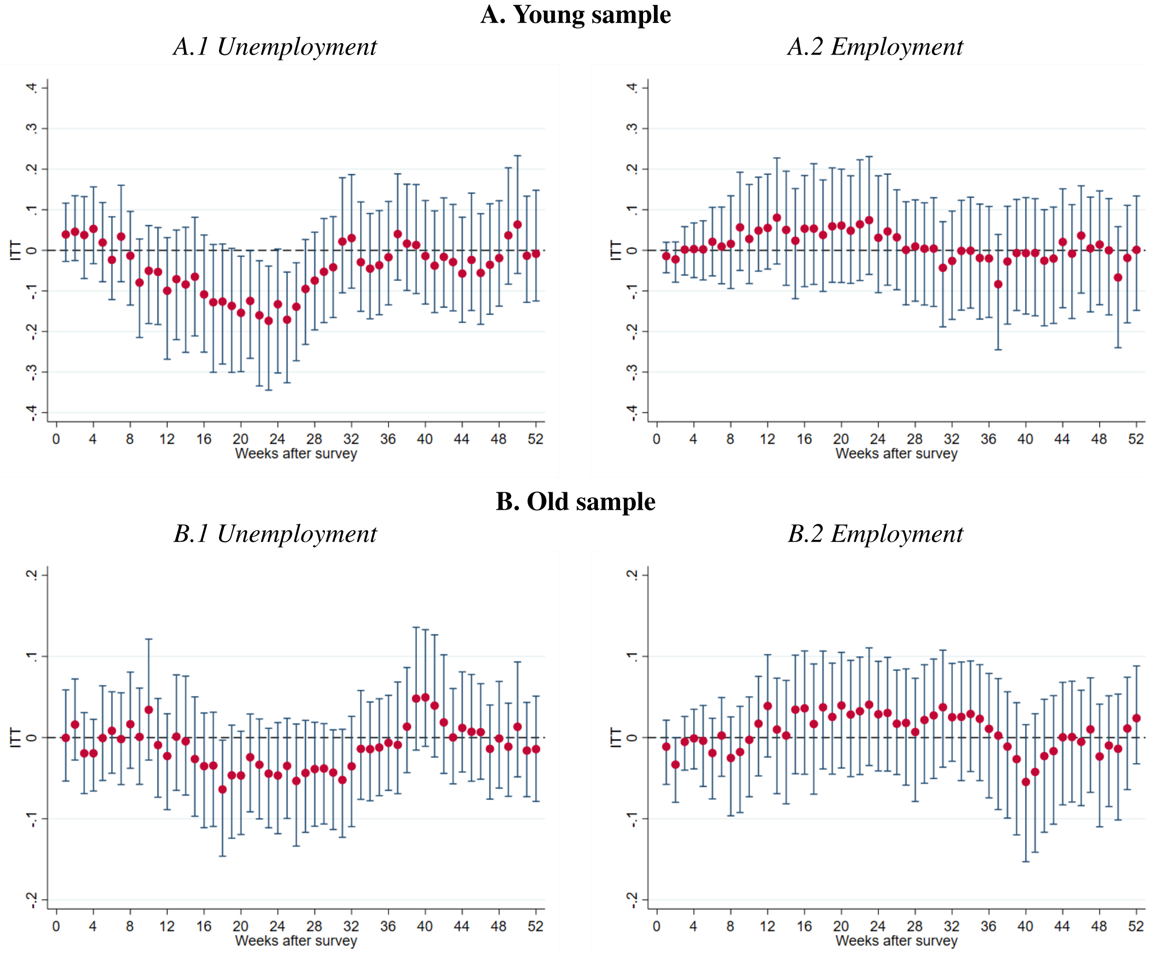
One could speculate that the differential effects with respect to age might be related to differences in how job seekers cope with job loss and the role of reemployment expectations. For instance, existing evidence indicates that late career job losses are associated with particular financial and psychological burdens (see e.g. Stevens, 1997; Theodossiou, 1998; Robb et al., 2003; Siegel et al., 2004) and that subjective employment expectations are significant predictors of depressions among older workers (Mandal et al., 2011). This may explain why older workers react to the information provision to a lesser extent than those receiving the treatment at age 28 or 29.

Moreover, information frictions and learning regarding the job seekers' labor market prospects may also contribute to our results (Farber and Gibbons, 1996). For instance, younger workers with less labor market experience might be more likely to update their expectations about their future employment prospects in reaction to the information treatment. Informing job seekers about an increased risk of becoming long-term unemployed most likely reduces the expected returns from continuing job search, which could explain why treated individuals more often leave UI without starting paid employment.

[16] For instance, we see an increase in the employment probability of around 9 percentage points 14 weeks after the individual received the treatment.

[17] Figure A4 in Appendix A shows the effects on the number of monthly working hours and labor earnings. The results indicate that, if at all, any intensive margin effects are rather small.

Figure 6: Unemployment and employment over time



Note: ITT-estimates and robust 95 % confidence intervals. These are estimated while controlling for problematic covariates and with a linear polynomial, triangular kernel and using the optimal bandwidth. Controls refer to gender and business services for the young sample, while for the old sample it is short cycle tertiary education, number of employers and previous wages. The figures show the effect of profiling on the probability to be unemployed (1) or employed (2) in a given week relative to the survey. We define the individual as being unemployed if she receives UI-benefits in a given week. We define the individual as being employed if she i) does not receive any public transfers in a given week and ii) does earn a wage income in the month corresponding to that week.

6 Conclusion

In this paper, we provide a first attempt to combine statistical profiling with an econometric policy evaluation of an information treatment for unemployed workers. Specifically, we are interested in the causal effect of providing personalized information about job seekers' risk of becoming long-term unemployed. The latter is obtained using a tree-based Machine Learning procedure, which allows us to exploit age discontinuities in the underlying decision for an RDD.

Our results indicate that the information treatment may be effective in promoting exits from unemployment for a significant number of UI benefit recipients. Receiving information about being at high risk of long-term unemployment is estimated to reduce the likelihood of remaining unemployed after 6 months by 14 percentage points in a sample of young workers and by 5 percentage points in sample of old workers. These effects vanish over time as the individuals not receiving the information catch up. Moreover, despite the significant reduction in unemployment, we find no evidence that the information treatment improves the job seekers' labor market performance in terms of employment probabilities, working hours and cumulated earnings.

The validity of our estimates rest on the usual RDD assumption that unemployed individuals immediately above and below the relevant age cutoffs are not systematically different. Conducting standard tests of these assumptions, however, we find clear indications that these assumptions are violated. In particular, we see important discontinuities in several predetermined covariates at the age cutoffs. Although all results are robust to controlling for these predetermined covariates, this finding raises severe concerns about the possibility of additional unobserved differences.

Overall, the results of our analysis appear promising for the possibility of using profiling tools to help move workers out of unemployment faster. At the same time, the potential violations of the identifying assumptions precludes us from making any firm conclusions. Additional evaluation of profiling tools are necessary, preferably using an explicit randomized controlled trial.

References

Almond, D. and Doyle, J. J. (2011), 'After midnight: A regression discontinuity design in length of postpartum hospital stays', *American Economic Journal: Economic Policy* **3**(3), 1–34.

Barreca, A. I., Guldi, M., Lindo, J. M. and Waddell, G. R. (2011), 'Saving babies? revisiting the effect of very low birth weight classification', *The Quarterly Journal of Economics* **126**(4), 2117–2123.

Behncke, S., Frölich, M. and Lechner, M. (2009), 'Targeting labour market programmes—results from a randomized experiment', *Swiss Journal of Economics and Statistics* **145**(3), 221–268.

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 (Statistics in Society)* **173**(1), 67–92.

Belot, M., Kircher, P. and Muller, P. (2018), 'Providing advice to jobseekers at low cost: An experimental study on online advice', *The Review of Economic Studies* **86**(4), 1411–1447.

Bjerk, D. (2007), 'Racial profiling, statistical discrimination, and the effect of a colorblind policy on the crime rate', *Journal of Public Economic Theory* **9**(3), 521–545.

Black, D. A., Galdo, J. and Smith, J. A. (2007), 'Evaluating the worker profiling and reemployment services system using a regression discontinuity approach', *American Economic Review* **97**(2), 104–107.

Burgess Jr, J. F., Christiansen, C. L., Michalak, S. E. and Morris, C. N. (2000), 'Medical profiling: improving standards and risk adjustments using hierarchical models', *Journal of Health Economics* **19**(3), 291–309.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.

Card, D., Chetty, R. and Weber, A. (2007), 'The spike at benefit exhaustion: Leaving the unemployment system or starting a new job?', *American Economic Review* **97**(2), 113–118.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2017), 'A practical introduction to regression discontinuity designs', *Cambridge Elements: Quantitative and Computational Methods for Social Science-Cambridge University Press I* .

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2019), *A practical introduction to regression discontinuity designs: Foundations*, Cambridge University Press.

Desiere, S. and Langenbucher, K. (2018), 'Profiling tools for early identification of job seekers who need extra support', *OECD Policy Brief on Activation Policies* (dec), 1–4.

Farber, H. S. and Gibbons, R. (1996), 'Learning and wage dynamics', *The Quarterly Journal of Economics* **111**(4), 1007–1047.

Frölich, M. (2008), 'Statistical treatment choice: an application to active labor market programs', *Journal of the American Statistical Association* **103**(482), 547–558.

Horton, J. J. (2017), 'The effects of algorithmic labor market recommendations: Evidence from a field experiment', *Journal of Labor Economics* **35**(2), 345–385.

Jacobson, L. S., LaLonde, R. J. and Sullivan, D. G. (1993), 'Earnings losses of displaced workers', *The American Economic Review* pp. 685–709.

Kassenboehmer, S. C. and Schatz, S. G. (2017), 'Re-employment expectations and realisations: Prediction errors and behavioural responses', *Labour Economics* **44**, 161–176.

Lechner, M. and Smith, J. (2007), 'What is the value added by caseworkers?', *Labour economics* **14**(2), 135–151.

Mandal, B., Ayyagari, P. and Gallo, W. T. (2011), 'Job loss and depression: The role of subjective expectations', *Social Science & Medicine* **72**(4), 576–583.

McCrary, J. (2008), 'Manipulation of the running variable in the regression discontinuity design: A density test', *Journal of econometrics* **142**(2), 698–714.

Mueller, A. I., Spinnewijn, J. and Topa, G. (2021), 'Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias', *American Economic Review* **111**(1), 324–63.

Neal, D. (1995), 'Industry-specific human capital: Evidence from displaced workers', *Journal of labor Economics* pp. 653–677.

Persico, N. and Todd, P. E. (2005), 'Passenger profiling, imperfect screening, and airport security', *American Economic Review* **95**(2), 127–131.

Phillips, P. J. and Pohl, G. (2012), 'Economic profiling of the lone wolf terrorist: can economics provide behavioral investigative advice?', *Journal of Applied Security Research* **7**(2), 151–177.

Robb, C., Haley, W., Becker, M., Polivka, L. and Chwa, H.-J. (2003), 'Attitudes towards mental health care in younger and older adults: Similarities and differences', *Aging & Mental Health* **7**(2), 142–152.

Santtila, P., Ritvanen, A. and Mokros, A. (2004), 'Predicting burglar characteristics from crime scene behaviour', *International Journal of Police Science & Management* **6**(3), 136–154.

Schiprowski, A. (2020), 'The role of caseworkers in unemployment insurance: Evidence from unplanned absences', *forthcoming: Journal of Labor Economics* .

Schmieder, J. F. and Trenkle, S. (2020), 'Disincentive effects of unemployment benefits and the role of caseworkers', *Journal of Public Economics* **182**, 104096.

Siegel, M. J., Bradley, E. H., Gallo, W. T. and Kasl, S. V. (2004), 'The effect of spousal mental and physical health on husbands' and wives' depressive symptoms, among older adults: longitudinal evidence from the health and retirement survey', *Journal of Aging and Health* **16**(3), 398–425.

Spinnewijn, J. (2015), 'Unemployed but optimistic: Optimal insurance design with biased beliefs', *Journal of the European Economic Association* **13**(1), 130–167.

Staghøj, J., Svarer, M. and Rosholm, M. (2010), 'Choosing the best training programme: is there a case for statistical treatment rules?', *Oxford Bulletin of Economics and Statistics* **72**(2), 172–201.

Staiger, D. and Stock, J. H. (1997), 'Instrumental variables regression with weak instruments', *Econometrica* **65**(3), 557–586.

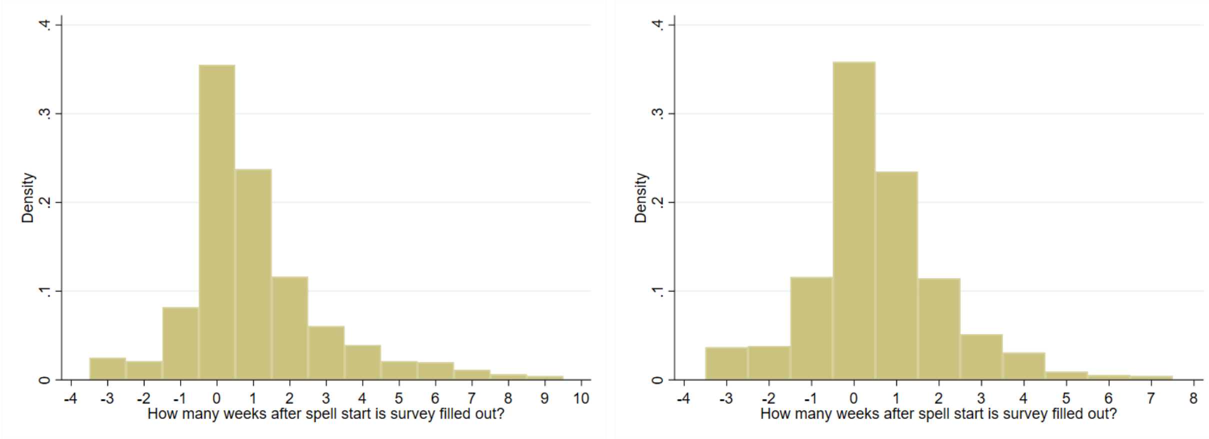
Stevens, A. H. (1997), 'Persistent effects of job displacement: The importance of multiple job losses', *Journal of Labor Economics* **15**(1, Part 1), 165–188.

Sullivan, D. and von Wachter, T. (2009), 'Job displacement and mortality: An analysis using administrative data', *The Quarterly Journal of Economics* **124**(3), 1265–1306.

Theodossiou, I. (1998), 'The effects of low-pay and unemployment on psychological well-being: a logistic regression approach', *Journal of Health Economics* **17**(1), 85–104.

A Additional Tables and Figures

Figure A1: Survey relative to spell start



(a) Young

(b) Old

The plot shows the number of weeks between the spell start and take-up of the survey, winsorized at the 1st and 99st percentile. A positive (negative) number means that the survey is taken after (before) the spell start. The spell start is found by going back in time from the survey week until the first week in which the individual does not receive UI-benefits in 7 consecutive weeks.

Figure A2: Message to job seekers (Danish)

"Tak for dine svar. Som supplement hertil er der indhentet oplysninger om din alder, tidligere forsørgelses- og lønhistorik mv. Disse oplysninger kan give et godt udgangspunkt for en indsats tilpasset din vej mod job.

Du skal tale med din sagsbehandler om din vej videre. Derfor anbefaler vi, at du printer dine svar ud og tager dem med til din første samtale. På den måde, har du og din sagsbehandler bedst mulig chance for at iværksætte de indsatser, der kan hjælpe dig hurtigst muligt tilbage på arbejdsmarkedet. Du kan genfinde dine svar på "Min side" på Jobnet de næste tre måneder."

(a) Low risk

"Tak for dine svar. Som supplement hertil er der indhentet oplysninger om din alder, tidligere forsørgelses- og lønhistorik mv.

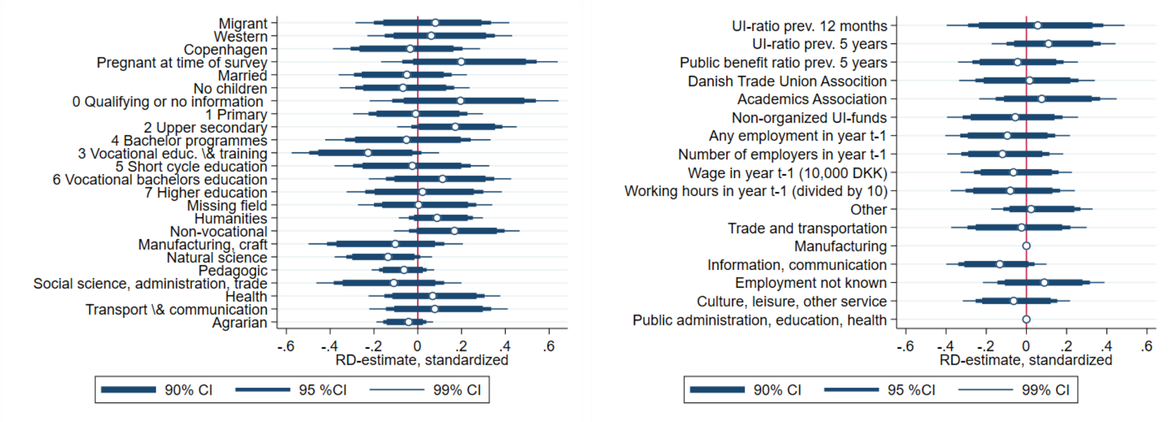
Der er flere af dine oplysninger, der peger i retning af, at du kan opleve udfordringer med hurtigt at vende tilbage i beskæftigelse. Analyser viser, at personer der har samme karakteristika som dig og som tidligere har oplevet ledighed, i højere grad end andre dagpengemodtagere, har haft udfordringer med at opnå beskæftigelse. Derfor er det ekstra vigtigt for dig og din sagsbehandler at I drøfter, hvordan din indsats skal tilrettelægges, således at du kommer hurtigt tilbage på arbejdsmarkedet.

Du skal tale med din sagsbehandler om din vej videre. Derfor anbefaler vi, at du printer dine svar ud og tager dem med til din første samtale. På den måde, har du og din sagsbehandler bedst mulig chance for at iværksætte de indsatser, der kan hjælpe dig hurtigst muligt tilbage på arbejdsmarkedet. Du kan genfinde dine svar på "Min side" på Jobnet de næste tre måneder."

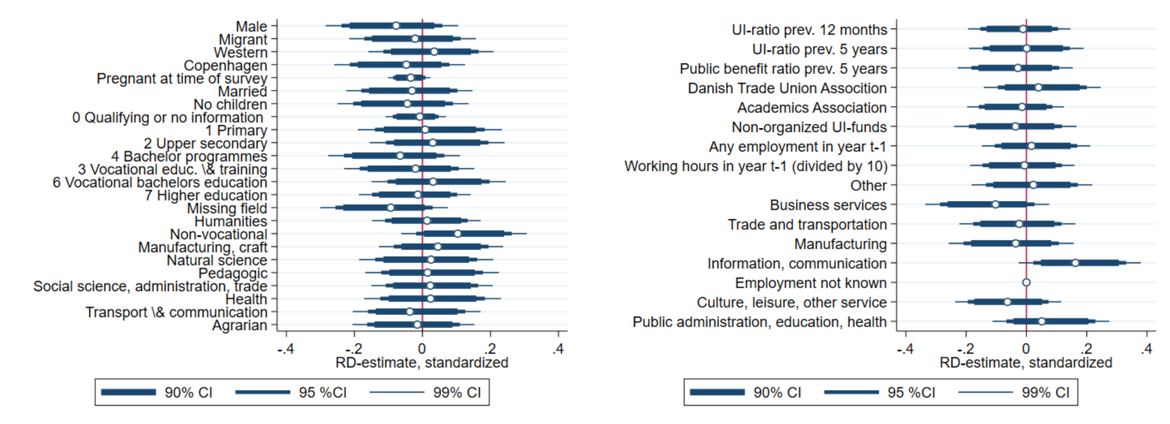
(b) High risk

Note: Panel a and b is the actual message send to low and high risk job seekers. The yellow box marks the additional paragraph added to jobseekers flagged as high risk.

Figure A3: Covariate balance - with controls



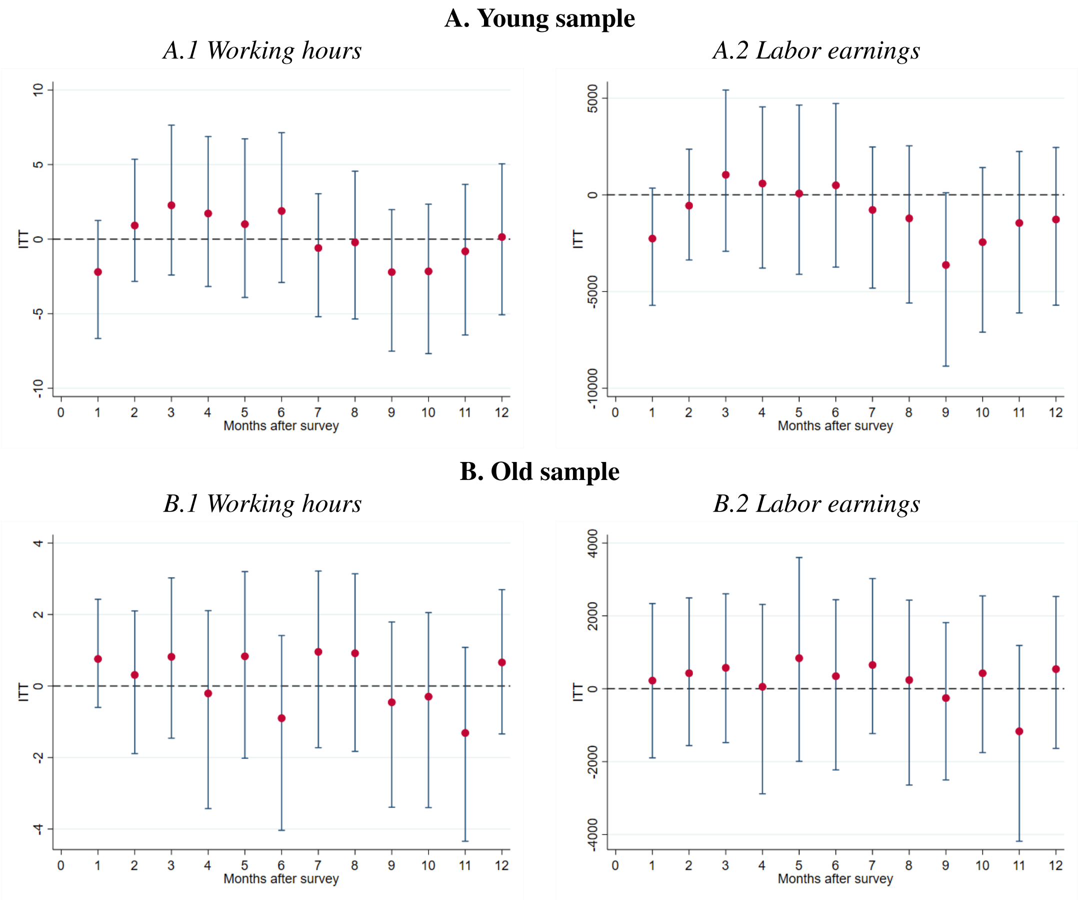
(a) Young



(b) Old

Note: RD-estimates from an RD-regression using the predetermined covariate as the outcome of interest and controlling for the problematic covariates. For the young sample these are gender and business services, while for the old sample these are short cycle tertiary education, number of employers and previous wages. The RD regressions use a linear polynomial, a triangular kernel and the optimal bandwidth from the main specification (unemployed in week 26 as outcome).

Figure A4: Working Hours and Labor Earnings over Time



Note: ITT-estimates and robust 95 % confidence intervals. The figures show the effect of profiling on working hours and wages in a given month relative to the survey (zeros are included - i.e. not conditional on being employed).

Table A1: Nodes in the original representation of the Decision Tree

	Input factors								
Node	Wage inc. in last 12 months	Education[a]	Industry[b]	UIB-ratio in last[c] 12 months	UIB-ratio in last[c] 5 years	PB-ratio[d] in last 5 years	Origin[e]	Age in years	Pred. risk in %
1	No	—	—	—	—	—	—	>50	72.2
2	No	—	—	—	—	—	—	45-50	67.4
3	No	—	—	—	—	—	—	29-44	63.8
4	No	Cat. 1	—	—	—	—	—	<29	68.6
5	No	Cat. 2	—	—	—	—	—	<29	59.4
6	—	—	Cat. 1	—	>0.254	—	—	>55	68.2
7	—	—	Cat. 1	—	<0.001	—	—	>55	71.2
8	—	—	Cat. 2	>0.289	—	—	—	>57	63.5
9	—	—	Cat. 2	<0.003	—	—	—	>57	74.1
10	—	—	Cat. 2	<0.005	—	—	—	54-57	62.1
11	—	—	Cat. 2	<0.019	—	>0.247	Cat. 1/2	28-53	61.7
12	—	—	Cat. 2	—	—	—	Cat. 2/3	28-53	63.5
13	—	—	Cat. 2	—	—	—	Cat. 3/4/5	28-53	62.1

Source: Danish Employment Agency (STAR)
Note: Depicted are the 13 nodes with the highest predicted risk of being unemployed for more than six months
[a] *Education categories:* 1 - humanities, religion, aesthetic or missing education; 2 - social science, business, office training non-vocational training, pedagogical training, fishery, agriculture or food, scientific education
[b] *Industry categories:* 1 - public administration, health, teaching, unknown activity, manufacturing, mining and quarrying, utilities, agriculture, forestry or fishery; 2 - trade, logistics, business services, culture, leisure or other services, real estate, information and communication, financial or insurance services.
[c] Unemployment-benefit-ratio defined as the ...
[d] Public-benefit-ratio defined as the ...
[e] *Origin categories:* 1 - Danish, 2 - descendent (Western), 3 - immigrant (Western), 4 - descendent (non-Western), 5 - immigrant (non-Western).

Table A2: Subgroups in the representation of the decision tree that isolates age cutoffs

Subgroup	Employee	Industry	Education field	UI-ratio 5	UI-ratio 12	Public ratio 5	Origin	Nodes	Age induced treatment	Obs (+/- 3 years)
1	No	0	1	-	-	-	-	1, 2, 3, 4	-	-
2	No	0	2	-	-	-	-	1, 2, 3, 5	-	-
3	No	0	3	-	-	-	-	0, 1, 2, 3	Age >29	905
4	Yes	1	-	ratio <0.001	-	-	-	0, 7	Age >56	2456
5	Yes	1	-	0.001 <ratio <0.254	-	-	-	0	-	-
6	Yes	1	-	ratio >0.254	-	-	-	0, 6	Age >56	910
7	Yes	2	-	-	ratio <0.003	ratio <0.246	1	0, 9, 10	Age >54	2559
8	Yes	2	-	-	ratio <0.003	ratio >0.246	1	0, 9, 10, 11	Age >28	552
9	Yes	2	-	-	ratio <0.003	-	2	0, 9, 10, 13	Age <28 or Age >54	1072 ; 144
10	Yes	2	-	-	ratio <0.003	-	3	9, 10, 12, 13	-	-
11	Yes	2	-	-	0.003 <ratio <0.004	ratio <0.246	1	0, 10	54 <Age <58	36 ; 32
12	Yes	2	-	-	0.003 <ratio <0.004	ratio >0.246	1	0, 10, 11	28 <Age <58	10 ; 13
13	Yes	2	-	-	0.003 <ratio <0.004	-	2	0, 10, 13	Age <28 or 54 <Age <57	19 ; 0 ; 1
14	Yes	2	-	-	0.003 <ratio <0.004	-	3	0, 10, 12, 13	Age <58	4
15	Yes	2	-	-	0.004 <ratio <0.018	ratio <0.246	1	0	-	-
16	Yes	2	-	-	0.004 <ratio <0.018	ratio >0.246	1	0, 11	28 <Age <54	33 ; 27
17	Yes	2	-	-	0.004 <ratio <0.018	-	2	0, 13	Age <28	60
18	Yes	2	-	-	0.004 <ratio <0.018	-	3	0, 12, 13	Age <54	14
19	Yes	2	-	-	0.018 <ratio <0.288	ratio <0.246	1	0	-	-
20	Yes	2	-	-	0.018 <ratio <0.288	ratio >0.246	1	0	-	-
21	Yes	2	-	-	0.018 <ratio <0.288	-	2	0, 13	Age <28	192
22	Yes	2	-	-	0.018 <ratio <0.288	-	3	0, 12, 13	Age <54	76
23	Yes	2	-	-	ratio >0.288	ratio <0.246	1	0	Age >58	88
24	Yes	2	-	-	ratio >0.288	ratio >0.246	1	0, 8	Age >58	247
25	Yes	2	-	-	ratio >0.288	-	2	0, 8, 13	Age <28 or Age >54	40 ; 12
26	Yes	2	-	-	ratio >0.288	-	3	0, 8, 12, 13	Age <54 or Age >58	19 ; 15
27	Yes	3	-	-	-	-	-	0	-	-

Notes: The 27 subgroups resulting from a partitioning of the full sample using all input variables and associated cutoffs except age. The 'Nodes' column shows what risk nodes each subgroup consists of. Here, 1-13 refer to the original 13 risk nodes classified as high risk (see figure A1) and 0 refers to low risk. The last column shows the number of observations in a 3 year window around each cutoff.

[a] *Education categories:* 1 - humanities, religion, aesthetic or missing education; 2 - social work, office, non-commercial, or pedagogical training, fishery, agriculture or food, scientific education 3 - Manufacturing and crafts, health, transportation and communications

[b] *Industry categories:* 1 - public administration, health, teaching, employed in unknown activity, manufacturing, mining and quarrying, utilities, agriculture, forestry or fishery; 2 - trade, logistics, business services, culture, leisure or other services, real estate, information and communication, financial or insurance services. 3 - Construction

[c] *UI-ratio* defined as the fraction of days spent on UI-benefits in the previous 12 months or 5 years.

[d] *Origin categories:* 1 - Danish, 2 - descendent (Western country), 3 - immigrant (Western country), 4 - descendent (non-Western country), 5 - immigrant (non-Western country)

[e] *Public ratio* defined as the fraction of days spent on public transfers in the previous 5 years.

Table A3: Covariate balance
Young sample

	Mean	SD	RD-estimate	Robust 95% CI
Demographics				
Male	0.382	0.486	0.138	[0.038 : 0.283]
Migrant	0.134	0.341	0.013	[-0.083 : 0.095]
Western	0.933	0.251	0.026	[-0.024 : 0.097]
Copenhagen	0.274	0.446	-0.023	[-0.146 : 0.084]
Pregnant at time of survey	0.081	0.273	0.038	[-0.030 : 0.121]
Married	0.161	0.368	-0.048	[-0.157 : 0.036]
No children	0.699	0.459	0.010	[-0.094 : 0.136]
Level of Education				
0 Qualifying or no information	0.013	0.113	0.020	[-0.006 : 0.050]
1 Primary	0.094	0.292	0.018	[-0.047 : 0.097]
2 Upper secondary	0.079	0.270	0.047	[-0.011 : 0.109]
4 Bachelor programmes	0.075	0.263	-0.018	[-0.084 : 0.048]
3 Vocational educ. & training	0.251	0.434	-0.081	[-0.197 : 0.030]
5 Short cycle education	0.092	0.290	-0.002	[-0.080 : 0.077]
6 Vocational bachelors education	0.256	0.437	0.022	[-0.094 : 0.119]
7 Higher education	0.140	0.347	-0.006	[-0.099 : 0.089]
Field of Education				
Missing	0.256	0.436	0.048	[-0.042 : 0.178]
Humanities	0.016	0.125	0.017	[-0.007 : 0.048]
Non-vocational	0.011	0.104	0.019	[-0.004 : 0.044]
Manufacturing, craft	0.197	0.398	-0.025	[-0.148 : 0.069]
Natural science	0.007	0.085	-0.010	[-0.029 : 0.005]
Pedagogic	0.007	0.082	-0.005	[-0.015 : 0.005]
Social science, administration, trade	0.055	0.228	-0.022	[-0.092 : 0.038]
Health	0.376	0.485	-0.047	[-0.172 : 0.061]
Transport & communication	0.073	0.261	0.028	[-0.031 : 0.099]
Agrarian	0.002	0.043	-0.003	[-0.009 : 0.002]
Transfers				
UI-ratio prev. 12 months	0.018	0.063	0.001	[-0.020 : 0.020]
UI-ratio prev. 5 years	0.107	0.135	0.022	[-0.008 : 0.061]
Public benefit ratio prev. 5 years	0.213	0.210	0.007	[-0.044 : 0.063]
Labor Market History				
Academics Association	0.559	0.497	0.001	[-0.135 : 0.121]
Danish Trade Union Assocition	0.202	0.401	0.019	[-0.071 : 0.129]
Non-organized UI-funds	0.186	0.389	-0.008	[-0.114 : 0.091]
Any employment in year t-1	0.376	0.485	0.013	[-0.101 : 0.147]
Number of employers in year t-1	0.494	0.749	0.007	[-0.154 : 0.222]
Wage in year t-1 (10,000 DKK)	0.768	1.122	0.051	[-0.201 : 0.380]
Working hours in year t-1 (divided by 10)	1.128	1.559	0.045	[-0.312 : 0.498]
Previous Industry				
Other	0.012	0.107	0.003	[-0.017 : 0.037]
Business services	0.095	0.293	0.074	[0.012 : 0.159]
Trade and transportation	0.218	0.413	-0.022	[-0.139 : 0.081]
Manufacturing	0.000	0.000	0.000	[0.000 : 0.000]
Information, communication	0.024	0.153	-0.025	[-0.064 : 0.007]
Employment not known	0.626	0.484	-0.016	[-0.151 : 0.098]
Culture, leisure, other service	0.026	0.158	-0.014	[-0.051 : 0.028]
Public administration, education, health	0.000	0.000	0.000	[0.000 : 0.000]

Notes: RD-estimates from an RD-regression using the predetermined covariate as the outcome of interest, a linear polynomial, a triangular kernel and the optimal bandwidth from the main specification (unemployed in week 26 as outcome).

Table A4: Covariate balance
Old sample

	Mean	SD	RD-estimate	Robust 95% CI
Demographics				
Male	0.399	0.490	-0.052	[-0.136 : 0.013]
Migrant	0.067	0.250	-0.002	[-0.041 : 0.034]
Western	0.975	0.156	0.004	[-0.021 : 0.025]
Copenhagen	0.221	0.415	-0.026	[-0.098 : 0.027]
Pregnant at time of survey	0.001	0.025	-0.001	[-0.004 : 0.000]
Married	0.810	0.393	-0.016	[-0.079 : 0.039]
No children	0.638	0.481	-0.017	[-0.095 : 0.051]
Level of Education				
0 Qualifying or no information	0.001	0.025	-0.000	[-0.002 : 0.001]
1 Primary	0.185	0.389	0.017	[-0.035 : 0.085]
2 Upper secondary	0.058	0.234	0.007	[-0.026 : 0.047]
4 Bachelor programmes	0.015	0.120	-0.008	[-0.030 : 0.008]
3 Vocational educ. & training	0.424	0.494	0.012	[-0.064 : 0.083]
5 Short cycle education	0.065	0.246	-0.030	[-0.077 : -0.005]
6 Vocational bachelors education	0.184	0.388	0.012	[-0.039 : 0.078]
7 Higher education	0.068	0.252	-0.009	[-0.049 : 0.024]
Field of education				
Missing	0.540	0.498	-0.014	[-0.091 : 0.059]
Humanities	0.029	0.168	0.003	[-0.021 : 0.026]
Non-vocational	0.028	0.165	0.018	[-0.003 : 0.044]
Manufacturing, craft	0.039	0.194	0.002	[-0.025 : 0.030]
Natural science	0.039	0.193	-0.004	[-0.041 : 0.022]
Pedagogic	0.114	0.318	0.010	[-0.032 : 0.063]
Social science, administration, trade	0.151	0.358	-0.011	[-0.066 : 0.039]
Health	0.040	0.196	0.005	[-0.023 : 0.036]
Transport & communication	0.010	0.099	-0.006	[-0.019 : 0.011]
Agrarian	0.010	0.101	-0.003	[-0.018 : 0.009]
Transfers				
UI-ratio prev. 12 months	0.019	0.081	0.001	[-0.010 : 0.011]
UI-ratio prev. 5 years	0.068	0.139	0.006	[-0.013 : 0.027]
Public benefit ratio prev. 5 years	0.114	0.168	0.003	[-0.022 : 0.028]
Labor Market History				
Academics Association	0.599	0.490	0.039	[-0.025 : 0.124]
Danish Trade Union Assocition	0.090	0.287	-0.017	[-0.066 : 0.015]
Non-organized UI-funds	0.309	0.462	-0.024	[-0.099 : 0.045]
Any employment in year t-1	0.981	0.138	0.006	[-0.011 : 0.027]
Number of employers in year t-1	1.260	0.589	0.080	[0.006 : 0.176]
Wage in year t-1 (10,000 DKK)	3.560	1.988	-0.313	[-0.668 : -0.083]
Working hours in year t-1 (divided by 10)	3.388	0.628	-0.042	[-0.152 : 0.041]
Previous Industry				
Other	0.122	0.327	0.006	[-0.045 : 0.052]
Business services	0.114	0.317	-0.032	[-0.090 : 0.008]
Trade and transportation	0.193	0.395	-0.012	[-0.074 : 0.045]
Manufacturing	0.147	0.354	-0.017	[-0.078 : 0.032]
Information, communication	0.049	0.216	0.028	[-0.003 : 0.063]
Employment not known	0.000	0.000	0.000	[0.000 : 0.000]
Culture, leisure, other service	0.032	0.177	-0.010	[-0.034 : 0.015]
Public administration, education, health	0.344	0.475	0.037	[-0.016 : 0.126]

Notes: RD-estimates from an RD-regression using the predetermined covariate as the outcome of interest, a linear polynomial, a triangular kernel and the optimal bandwidth from the main specification (unemployed in week 26 as outcome).

Table A5: Robustness, young sample.
Dep. var: Unemployed in week 26

	(1)	(2)	(3)	(4)	(5)	(6)
ITT	-0.138***	-0.139***	-0.141***	-0.127	-0.155**	-0.138***
	(0.052)	(0.052)	(0.053)	(0.081)	(0.064)	(0.052)
FS	0.729***	0.724***	0.716***	0.723***	0.749***	0.729***
	(0.047)	(0.047)	(0.045)	(0.054)	(0.044)	(0.050)
FRD	-0.170*	-0.170*	-0.187*	-0.172	-0.172*	-0.170*
	(0.100)	(0.100)	(0.099)	(0.114)	(0.090)	(0.090)
Controls	NO	YES	NO	NO	NO	NO
RV	rv_birth	rv_birth	rv_survey	rv_birth	rv_birth	rv_birth
Polynomial	1	1	1	2	1	1
Kernel	Triangular	Triangular	Triangular	Triangular	Uniform	Triangular
Clustering	NO	NO	NO	NO	NO	YES
Mean dep. var	0.400	0.400	0.400	0.400	0.400	0.400
Effective obs (ITT)	1636	1636	1577	1489	914	1646
Bandwidth (ITT)	127	127	123	116	71	129
Robust 95% CI (ITT)	[-.271 ; -.029]	[-.272 ; -.031]	[-.278 ; -.031]	[-.291 ; .068]	[-.31 ; -.014]	[-.272 ; -.031]
Effective obs (FS)	859	859	929	1447	824	859
Bandwidth (FS)	67	67	72	113	64	67
Robust 95% CI (FS)	[.623 ; .822]	[.618 ; .817]	[.615 ; .801]	[.609 ; .835]	[.643 ; .831]	[.617 ; .828]
Effective obs (FRD)	859	859	929	1447	824	859
Bandwidth (FRD)	67	67	72	113	64	67
Robust 95% CI (FRD)	[-.388 ; .031]	[-.389 ; .032]	[-.404 ; .007]	[-.412 ; .066]	[-.378 ; .008]	[-.368 ; .011]

Notes: Each RD-coefficient are estimated with a bandwidth optimal for the particular robustness check. Controls refer to the problematic covariates: For the young sample these are gender and business services, while for the old sample these are short cycle tertiary education, number of employers and previous wages.

Table A6: Robustness, old sample.
Dep. var: Unemployed in week 26

	(1)	(2)	(3)	(4)	(5)	(6)
ITT	-0.138***	-0.139***	-0.141***	-0.127	-0.155**	-0.138***
	(0.052)	(0.052)	(0.053)	(0.081)	(0.064)	(0.052)
FS	0.729***	0.724***	0.716***	0.723***	0.749***	0.729***
	(0.047)	(0.047)	(0.045)	(0.054)	(0.044)	(0.050)
FRD	-0.170*	-0.170*	-0.187*	-0.172	-0.172*	-0.170*
	(0.100)	(0.100)	(0.099)	(0.114)	(0.090)	(0.090)
Controls	NO	YES	NO	NO	NO	NO
RV	rv_birth	rv_birth	rv_survey	rv_birth	rv_birth	rv_birth
Polynomial	1	1	1	2	1	1
Kernel	Triangular	Triangular	Triangular	Triangular	Uniform	Triangular
Clustering	NO	NO	NO	NO	NO	YES
Mean dep. var	0.400	0.400	0.400	0.400	0.400	0.400
Effective obs (ITT)	1636	1636	1577	1489	914	1646
Bandwidth (ITT)	127	127	123	116	71	129
Robust 95% CI (ITT)	[-.271 ; -.029]	[-.272 ; -.031]	[-.278 ; -.031]	[-.291 ; .068]	[-.31 ; -.014]	[-.272 ; -.031]
Effective obs (FS)	859	859	929	1447	824	859
Bandwidth (FS)	67	67	72	113	64	67
Robust 95% CI (FS)	[.623 ; .822]	[.618 ; .817]	[.615 ; .801]	[.609 ; .835]	[.643 ; .831]	[.617 ; .828]
Effective obs (FRD)	859	859	929	1447	824	859
Bandwidth (FRD)	67	67	72	113	64	67
Robust 95% CI (FRD)	[-.388 ; .031]	[-.389 ; .032]	[-.404 ; .007]	[-.412 ; .066]	[-.378 ; .008]	[-.368 ; .011]

Notes: Each RD-coefficient are estimated with a bandwidth optimal for the particular robustness check. Controls refer to the problematic covariates: For the young sample these are gender and business services, while for the old sample these are short cycle tertiary education, number of employers and previous wages.

Table A7: Employment outcomes in month 6 - no controls

	Young			Old		
	Employed	Hours	Wage	Employed	Hours	Wage
ITT	0.040	2.219	792.660	0.004	-0.867	41.465
	(0.063)	(2.149)	(1797.744)	(0.033)	(1.210)	(1048.813)
FS	0.721***	0.723***	0.726***	0.910***	0.911***	0.909***
	(0.050)	(0.049)	(0.048)	(0.014)	(0.013)	(0.014)
FRD	0.036	2.577	-133.127	0.002	-0.438	-308.246
	(0.110)	(3.691)	(3224.875)	(0.037)	(1.227)	(1281.968)
Controls	NO	NO	NO	NO	NO	NO
Polynomial	1	1	1	1	1	1
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Clustering	NO	NO	NO	NO	NO	NO
Mean dep. var	0.504	15.096	11679.866	0.557	16.657	14776.383
Effective obs (ITT)	1225	1241	1376	4712	3914	5181
Bandwidth (ITT)	96	98	107	87	72	95
Robust 95% CI (ITT)	[-.102 ; .195]	[-2.555 ; 7.537]	[-3411.89 ; 5084.663]	[-.081 ; .069]	[-4.052 ; 1.474]	[-2672.141 ; 2190.658]
Effective obs (FS)	798	806	833	4403	4577	4204
Bandwidth (FS)	61	63	64	82	84	78
Robust 95% CI (FS)	[.611 ; .824]	[.615 ; .823]	[.618 ; .823]	[.878 ; .937]	[.879 ; .937]	[.876 ; .937]
Effective obs (FRD)	798	806	833	4403	4577	4204
Bandwidth (FRD)	61	63	64	82	84	78
Robust 95% CI (FRD)	[-.197 ; .273]	[-5.088 ; 10.611]	[-6965.786 ; 6668.715]	[-.084 ; .077]	[-3.425 ; 1.916]	[-3285.766 ; 2270.521]

Notes: RD-coefficients are estimated with a linear polynomial, triangular kernel and the optimal bandwidth. Note that we allow the optimal bandwidth to be different for the ITT and IV specification. Hence, the FS and ITT coefficients are estimated using different bandwidths, and therefore a simple division of ITT by FS does not give the FRD estimate.

B Technical Appendix

B.1 Input variables

Field of education: This variable is taken directly from the survey.

Origin: Linking to the population register, we obtain information on jobseeker origin

Age: We reconstruct the respondent's *age* based on the population register (CPR). Here we used the formula originally used by the programmers behind the profiling tool:
$Age = int((Calculationdate - Birthday)/365.25)$.

Industry in the previous 12 months: Industry in the previous 12 months is defined in terms of firm-specific accumulated wages. Using the monthly earned income register collected by the Danish tax authorities (EINDKOMST), we find accumulated wages from each of the firms that the individual was employed during the previous 12 months (relative to the calculation date). We use the firm in which the individual had the highest accumulated wages. Linking to the firm register (CVR), we can find the industry in which this firm was operating.

Employee in the previous 12 months: We define the individual as an employee if she appears in the monthly earned income register in the previous 12 months (EINDKOMST).

UI benefit ratio in the previous 12 months and 5 years: The UI-benefit ratio is defined as the fraction of days spent on UI-benefits over the previous 12 months or 5 years. Linking to registers with detailed (daily) information on UI-benefit payments[18] (FACT Y01A02, Y37A02, Y34A02), we can calculate the UI-benefit ratio.

Public benefit ratio in the previous 5 years: The public-benefit ratio is defined as the fraction of days spent on public transfers UI-benefits over the previous 12 months or 5 years. Linking to registers with detailed (daily) information on public subsidies and transfers[19] (FACT Y01A02 Y04A02, Y05A02, Y07A02, Y08A02, Y09A02, Y10A02, Y11A02, Y12A02, Y34A02, Y35A02, Y36A02, Y38A02, Y37A02, Y39A02), we can calculate the public-benefit ratio.

[18] UI-benefits ('dagpenge') as well as temporary subsidies ('særlig uddannelsesydelse', 'arbejdsmarkedsydelse', 'kontantydelse')

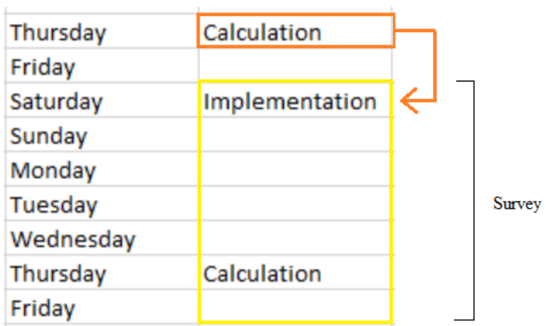
[19] In Danish 'kontanthjælp', 'sygedagpenge', 'revalidering', 'uddannelseshjælp', 'integrationsydelse', 'forrevalidering', 'ressourceforløb', 'jobafklaring', 'førtidpension', 'fleksydelse', 'dagpenge', 'særlig uddannelsesydelse', 'arbejdsmarkedsydelse', 'kontantydelse'

B.2 Imputation of Calculation Dates

For the reconstruction of input variables historically used in the Profiling Tool, it is crucial to know the exact date at which the variables were calculated. Unfortunately, we do not have this information, however, with knowledge about the operation of the Profiling Tool we can impute these dates. From discussions with operators, we know the *typical* data flows behind the Profiling Tool [20].

- STAR *calculates* the input variables and sends the data to an external operator, VISMA. This typically happens every a Thursday.
- Once a week, the BI-unit in VISMA *reads* the latest data sent from STAR. This typically happens on a Friday.
- At 2AM the morning after the data has been read, the new data is *implemented* in the statistical model. This typically happens on a Saturday.

This sequence of events is illustrated in the figure below. If STAR calculates the input variables on a Thursday, and VISMA implements the new data on Jobnet the following Saturday, it implies that all individuals filling out the survey between Saturday and Friday (yellow) will be profiled according to variables calculated on the same Thursday (orange). Hence, although the survey is filled out on a daily basis, the input variables used in the Profiling Tool are calculated on a weekly basis.



We do not have information about the individual calculation dates, but we can back them out from other pieces of information. First of all, we have information about the exact date at which the individual has filled out the survey. Second of all, we have information about the dates at which data was sent from STAR to VISMA. Assuming that the data was sent on the same day as the input variables were calculated, we get a *grid of calculation dates*. Third of all, we have access to all dates at which the data files were read in the BI-unit in VISMA. Assuming that the new data was implemented the day after it was read in the BI-unit, we therefore also have a *grid of implementation dates*. We can use this information to deduce the calculation date for each individual by a two-step mapping.

[20]The process is not automated, and hence a server breakdown, holidays or illness could delay the process.

1. We map from the individual survey date to one particular implementation date in the grid. Here, we choose the *last* implementation date that happens before or coincides with the individual survey date.

2. Second, we map from the implementation date to one particular calculation date in the grid. In some weeks, STAR has sent more than one data file to VISMA. The data-file that is implemented is the last one sent to VISMA the day before implementation. We therefore choose the *last* calculation date that happens before the implementation. Thus, we get a *calculation date* for each individual in the sample.

B.3 Running Variable

The running variable in the analysis takes the survey day as given and ask *when the individual should have been born in order to be treated.* For the calculation of this distance measure it is important to know how age was calculated in the Profiling Tool. Namely, age was historically calculated according to the formula

$$Age = int[(Calculation\_day - Birthday)/365.25]$$

There are two things to note here. First, the formula introduces a small error, since a year only has 365.25 days on average. Second, many $calculation\_days$'s will result in the same age due to to the use of the integer-function. The latter is particularly important, since we want to find the day at which the individual should have been born such that the individual's age equals the cut-off. This translates into solving for *Birthday_cut* in the formula

$$Cut\_off = int[(Calculation\_day - Birthday\_cut)/365.25]$$

Due to the use of the integer function, this equation will have multiple solutions. We go about this by going back, day by day, until the first day at which the equation holds. We define this day as the day at which the individual should have been born in order to be assigned to treatment. We then calculate the distance between the actual birthday and this day,

$$Dist = Birthday - Birthday\_cut$$

Finally, we create weekly bins of this distance.

$$RV_i = \begin{cases} floor((Dist + 6)/7) & \text{if Dist < -6} \\ 0 & \text{if -6} \geq \text{Dist} \geq 0 \\ ceil(Dist/7) & \text{if Dist > 0} \end{cases}$$