CREST
CENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

WORKING PAPER SERIES

INVESTIGATING THE IMPACT OF INTEGRATION AGREEMENTS ON LABOR MARKET OUTCOMES FOR WELFARE RECIPIENTS: A RANDOMIZED CONTROLLED TRIAL

Gerard J. van den Berg, Sarah Bernhard, Gesine Stephan, Arne Uhlendorff

N°12 /November 2024

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: communication.crest@ensae.fr
https://crest.science/

Investigating the Impact of Integration Agreements on Labor Market Outcomes for Welfare Recipients: A Randomized Controlled Trial

Gerard J. van den Berg (U Groningen, UMCG, IFAU), Sarah Bernhard (IAB),

Gesine Stephan (IAB, FAU), Arne Uhlendorff (CREST, IAB)

October 2024

Abstract

Integration agreements (IA) outline the efforts the jobseeker should undertake to find employment and specify the services that the caseworker would provide to assist them in their job search. The agreements include a declaration of legal consequences, and punitive benefit sanctions could be imposed based on this declaration. Recent evidence has shown that these IAs are effective for recipients of unemployment insurance (UI) benefits. Using a randomized controlled trial, this paper investigates whether IAs support the integration of welfare benefit recipients into the labor market. This integration is of utmost importance from a policy and societal point of view. Newly registered recipients of means-tested benefits were randomly assigned to one of three groups, receiving either a) a standard integration agreement with the accompanying declaration of legal consequences at the beginning of the welfare spell, or b) an integration agreement without such a declaration, or c) no integration agreement within the first six months of the benefit receipt. Findings indicate that, on average, group assignment has no effect on the transition out of welfare or entry into employment. Based on a Random Forest analysis to capture heterogeneity, we find no effect by the degree of labor market prospects either.

Keywords: Social assistance, unemployment, active labor market policy, field experiment.

JEL-Codes: J68, J64, I38

Acknowledgements: We are grateful for valuable comments received at economics research seminars at Hohenheim and Maastricht and at conferences and workshops of the BGPE 2023, EALE 2024, VfS 2024 and VfS Social Policy 2024, especially from our discussant Jonas Jessen. We thank the German Federal Ministry of Labor and Social Affairs (BMAS) – in particular Dr. Reinhard Penz – for initiating and enabling the project, the KPI11 department at the headquarters of the Federal Employment Agency – especially Daniel Elferich – for the support in implementing the trial, and the participating job centers for taking part in the project. We thank the project group of the IAB (ProIAB) for training the participating employees in the job centers, Ulrike Büschel for coordinating the training, and the Data and IT Management (DIM) department of the IAB for their support.

1. Introduction

In many countries, unemployed welfare recipients are at a high risk of staying out of work for long periods and ultimately for losing the connection to the regular labor market (see e.g. Crépon and Van den Berg, 2016). This leads to persistence in economic inequality and as such it is not only a concern at the level of the affected individuals but also a major societal concern and a policy challenge. One way to address this challenge is to evaluate policy measures that have been shown to work well in other settings.

This paper considers the Integration Agreement (IA) as such a policy measure. IAs are signed contracts establishing the rights and obligations between unemployed persons and their caseworkers at the agency that takes care of benefits and job search counseling. The contracts aim to increase transparency and accountability for both parties involved, with the ultimate aim to improve the labor market outcomes of the unemployed. Their degree of formality varies across countries (Knotz 2018), with Germany exhibiting a particularly strong emphasis on formal practices (Van den Berg, Hofmann et al., 2024).

In the absence of empirical evidence, effects of IAs could be labeled as ambiguous. They may clarify obligations and induce job search efforts, but they can also diminish the trust that unemployed persons have in the system, or can just be incomprehensible to recipients. For the German UI system, a randomized controlled trial (RCT) investigated the effects of the timing of IAs on the labor market prospects of unemployed men (Van den Berg, Hofmann et al. 2024). The results show a significantly positive effect of early IAs on the average probability to leave unemployment for a job, and this positive effect is mostly driven by individuals with adverse prospects. Based on these findings, the Federal Employment Agency reformed its recommendations for handling IAs, providing caseworkers in the unemployment insurance system with more flexibility regarding the timing of these agreements, effectively enabling them to omit IAs for workers with good prospects.[1]

Welfare recipients differ as a group from UI recipients[2] and they face a different institutional context. In particular, in the German welfare system, IAs are arguably more intrusive than in the UI system, as they are more closely tied to potentially severe sanctions for non-compliance with job search duties. This paper investigates whether labor market outcomes of unemployed welfare recipients are affected

[1] In Australia, Gerards and Welters (2022), using propensity score matching techniques, find that exposure to requirements has negative effects on employment and wages and suggest that exposure decreases job search motivation and increases stress. Schiprowski et al. (2024) provide descriptive evidence that the time spent on job search by UI recipients in Germany increased after caseworker meetings where an IA was concluded.

[2] In Germany, the shares of individuals without any vocational qualifications and/or of foreign origin are much higher among welfare recipients and recipients of other means-tested social benefits than in UI. Also, unemployed welfare recipients have usually experienced less stable employment paths and their unemployment spells tend to be much longer.

if an IA is: a) not concluded directly upon entering the welfare system, or b) not immediately accompanied by potential legal consequences. To explore this, we conducted a randomized controlled trial in which we vary the timing and content of IAs in the German welfare system. This design potentially allows us to disentangle the behavioral aspect of an IA – commonly signing a contract about rights and duties – from the monitoring aspect due to the threat of a sanction in case of non-compliance.

In the light of the higher strictness of the IA regime (as compared to that in UI), it is interesting to note that a few recent studies raise questions about the presumption that IAs in the welfare system enhance transparency and accountability in the relationship between recipients and caseworkers. Bernhard and Senghaas (2021), Senghaas et al. (2020), and Senghaas and Bernhard (2021) analyze standardized survey data along with qualitative interviews and group discussions in job centers (these act as employment agencies for welfare recipients). They conclude that IA texts are overloaded by the number of objectives, such as monitoring, service provision, and ensuring transparency. The transparency of the written form of the IA was harmed by the requirement that the agreements are legally binding, which necessitates legal jargon.[3] Van den Berg, Kesternich et al. (2024) examine how self-reported job search effort among unemployed welfare recipients in Germany varies with idiosyncratic local variation in the stringency with which IAs are enforced. They find that recipients with negative reciprocity as a personality trait reduce their search effort in response to an increased stringency. Ulmestig et al. (2020) carry out a qualitative study of IAs for social assistance applicants in a Swedish municipality. In their setting, the legal security is weaker, but the authors also identify distinct problems with that, as the regime tends to reproduce pre-existing inequality.

As a preview, the baseline result of this paper is that we do find not effects of IAs on labor market outcomes of welfare recipients. Further, the effects of early IAs among welfare recipients are significantly lower than the estimated effects for UI benefit recipients - Van den Berg, Hofmann et al. (2024) estimate an increase in the average probability of entering employment within one year by over 4 percentage points, among UI benefit recipients. The findings in the current paper have been taken into account in a recent major reform of the German welfare system.[4]

In the following, Section 2 will give an overview about the German welfare system for unemployed works. Section 3 describes the experimental design, while Section 4 gives an overview about the data

[3] In a group discussion, caseworkers noted that the multi-page IA document contains numerous standardized text elements to ensure that any potential reduction in welfare due to breaches of duty withstands legal scrutiny. Caseworkers also indicated that the pressure to conclude IAs could hinder the development of trust in the caseworker-client relationship. Generally, they believed that IAs were least useful for those facing language barriers and/or severe placement challenges.

[4] As we observe inaccurate assignments into the project and incomplete compliance, results must be interpreted as intention-to-treat effects.

and the sample used. Section 5 discusses compliance issues. Section 6 presents our results, while Section 7 draws some conclusions.

2. Institutional background

In Germany, persons in need who are capable of working can apply for welfare benefits—also known as unemployment benefits II (UB II) until the end of 2022—paid from the federal budget. This includes individuals without entitlement to unemployment benefits (UB), individuals whose maximum UB duration has expired, or individuals whose UB is insufficient to cover their needs. These welfare benefits are means-tested and administered by local job centers. They cover ``Standardized Needs'' (SN), such as expenses for food, clothing, personal hygiene, household goods, and other personal needs, as well as the costs of accommodation and heating, along with various additional needs. During the year 2017, when our RCT took place, the standardized needs amount was €409 per month for singles. For individuals living in multi-person households, the amounts per person were lower.

To ensure the cooperation of unemployed welfare recipients in their job search, job centers can impose sanctions. During the time of our intervention, for individuals aged 25 and older, each sanction lasted for approximately three months. If UB II recipients did not attend an appointment with the job center, their SN could be reduced by 10%. If they failed to meet an obligation outlined in their IA, their SN could be reduced by 30% for the first sanction and by 60% for a second sanction within a year. Additional sanctions within the same year could result in a complete cut of UB II benefits. This regime is arguably much more stringent than in the UI system,

Our main interest lies in IAs. IAs are contracts between UB II recipients and their caseworkers. At the time of our intervention, they were legally prescribed by the German Social Code II. A caseworker could impose an IA through a unilateral administrative act if no agreement was reached and the UB II recipient refused to sign it. The IA outlined the rights and obligations of the UB II recipient, as well as the legal consequences of non-compliance. Rights could include, for example, access to the online job board of the Federal Employment Agency or participation in a training program. Obligations could also include, for instance, the required number of job applications per month or participation in a training program. The declaration of consequences for neglecting duties in the IA served as the legal basis for any potential later sanctions.

The primary goal of introducing IAs was to ensure transparency and commitment while guiding recipients through the placement process. However, there has since been some criticism regarding the implementation of IAs within this process. For example, the German Federal Court of Auditors published a report claiming that during the year 2017, around 20 percent of IAs with UB II recipients were missing (Bundesrechnungshof, 2019). The report outlined that deficits have existed for many years and that previous supervisory and quality assurance measures did not improve the situation. It concludes that each year, millions of IAs create a significant administrative burden, which is likely not always matched by corresponding benefits.

3. Experimental design

The development of a concept for an RCT on IAs in the German welfare system was carried out in close cooperation with the headquarters of the Federal Employment Agency (FEA) and the German Federal Ministry of Labor and Social Affairs (BMAS).[5] This concept was subsequently formally approved by the BMAS. Importantly, written permission was obtained to deviate from legal regulations in the Social Code II within the project's framework. In particular, as explained below, it was ruled that sanctions could not be given to those RCT participants in treatment arms that preclude such sanctions. We view this as a unique achievement in itself. It effectively means that part of the law was set aside in order to enhance the transparency and clarity of the RCT.

In the analyses, only the individuals who met the following conditions are included. They should be registered as unemployed on the day of assignment (or be unemployed and sick or in an active labor market program (ALMP)), have not received welfare for at least six months immediately prior to the assignment, be between 25 and 61 years old, and be registered at one of the participating job centers.

The RCT has three treatment arms denoted by V1, V2 and V3. Each of these provides a different approach to handling the IA and the associated "Legal Consequences Declaration" (LCD) in the first six months after assignment (see Table 1). In V1, the *usual IA* is supposed to be signed at the first meeting with the caseworker. In V2, an *IA without a LCD* was supposed to be signed at the first meeting. In V3, *no IA* was to be established during the first six months of benefit receipt.

[Table 1 about here]

In this context, a temporary waiver of the IA or the instruction on legal consequences implies a waiver of the associated sanctions for lack of personal effort, refusal to take up employment or training, and refusal to participate in measures. According to the experimental protocol, in groups V2 and V3, so-called "breaches of duty" were not to be sanctioned during the first six months of benefit receipt. However, if jobseekers did not keep appointments at the job center (known as "missed appointments"), their unemployment benefit II could still be reduced during this period. Proposals for placement and assignments to ALMPs were also to be made without instructions on legal conse-

[5] This followed a request from the BMAS to determine whether the findings on IAs in UI in van den Berg, Hofmann et al. (2024) can be extrapolated to the welfare system.

quences in groups V2 and V3. Otherwise, access to placement proposals, support services, and measures, as well as the frequency of contacts, should not differ from the counseling and placement processes involving a standard IA to avoid disadvantaging anyone. During the first six months of the field period, the job centers in these variants were also not to replace the IA with a unilaterally issued administrative act if jobseekers refused to sign it.

The project was conducted in seven job centers located in both East and West Germany, as well as in urban and rural regions. Random assignment occurred between July 2017 and September 2018, with the exact starting date at the discretion of the job centers. Assigned UB II recipients should not have received UB II during the previous six months, should be unemployed and capable of work, aged between 25 and 61, and not disabled.[6]

In each job center, UB II recipients who met the criteria described above were randomly assigned to one of three different treatment groups for a period of one year. The random assignment was conducted using an app. Caseworkers had access to the identification number and last name of the UB II recipients and needed to press an assignment button. The system then generated a random number to determine the assigned treatment status. Subsequently, the assignment result was displayed by the app, and the information was stored in a database that contained the time of randomization, the outcome, and anonymized identifiers of both the client and the caseworker. Each assigned individual could only be assigned once to prevent manipulation by the caseworkers. They were instructed to save the results in predefined fields of their general placement software.

Caseworkers in these job centers received project-related training from IAB staff. For implementation and documentation in the main placement software used by caseworkers, central text modules were provided by the headquarters of the Federal Employment Agency, and language guidelines were developed for possible inquiries from beneficiaries.[7]

4. Data, sample, outcome variables, and methods

Our data originate from various register data sources providing information up to the end of December 2020. First, the Integrated Employment Biographies (IEB V16.00.01-202012) of the IAB provide information on periods of employment, benefit receipt, unemployment, and participation in

[6] The inclusion of refugees in the project was left to the discretion of the participating job centers; however, all of them decided against it.

[7] The project was overseen by a project board consisting of participating researchers, job centers, and representatives from the FEA headquarters and the BMAS. After extensive legal counseling and ethical reflections, it was decided that informing participants and obtaining informed consent was not required for this particular project. The reasons for this decision were that no risks for participating individuals were anticipated. Furthermore, informing participants and asking for informed consent might have induced behavioral changes and selective dropouts, which would have affected the generalizability of the results.

labor market programs. Second, we use data from the Jobseeker History (ASU-EEI V06.13.00-202104) on the timing of IAs. Third, the Basic Benefits History (LHG V10.01.00-202104) contains information on sanctions and household structure. Fourth, we incorporate information on the reasons for sanctions from the sanction history (ALLEGRO data 2022). Information about the results of the random assignment from the electronic assignment tool covers entries from July 2017 to May 2018. For the period from June 2018 to September 2018, assignments were reconstructed from the general placement tool used by caseworkers.[8]

In total, 8,217 random assignments took place. In addition to the assignment conditions listed in Section 3, we require that the assignment should not have been subsequently canceled, the assignment date should fall within the assignment period, and a link to the IEB data should be possible. After applying these criteria, 4,520 individuals remain. The most frequent reasons for dropout are inaccurate assignments, particularly that individuals were not new benefit recipients or were not registered as unemployed on the assignment date.

Our focus will be on the moments at which relevant events (e.g., entering employment) occur after the assignment day, and we are primarily interested in individuals receiving UB II who were not employed on the day of assignment. It should be noted that if the application for unemployment benefit II was submitted to the job center simultaneously or after the random assignment, a (retroactive) payment need not have been made in every case after the documents were checked. It should also be noted that some individuals were employed on the day of random assignment because UB II can also be used to supplement earned income that does not meet needs. We exclude those individuals who did not receive UB II or who were employed on the day of random assignment, which leaves us with a final sample of 2,659 individuals.

Table A.1 in the Online Appendix presents a comparison of the characteristics of the three groups, including gender, nationality, age groups, educational degree, household type, various indicators of labor market history, the job center, and the month of assignment. The groups are not significantly different with respect to these characteristics—as expected due to random assignment.

Our main outcome variables are transitions out of welfare receipt and transitions into employment as an employee (so while paying social security contributions, including subsidized employment). We are also interested in the imposition of a benefit sanction due to a failure to report. The latter type of sanctions (unlike those due to neglect of duty) can occur in all three treatment groups. In each case we consider events within 720 days following random assignment.

[8] Accidentally, the data from the electronic tool for this time period were deleted in the course of server migrations. Additional analysis shows that data origin does not correlate with the group assignment.

First, we present Kaplan-Meier survival functions for the outcome variables. Second, we estimate the unconditional probabilities of transitions having been made into the relevant new states, within 180, 360, and 720 days following random assignment. In the case of sanctions this poses a methodological problem as sanctions can only be observed before leaving welfare. The duration until a sanction is right-censored by the duration until leaving welfare. These two durations constitute potentially dependent competing risks and the type of dependence may be influenced by the assigned treatment arm, obstructing identification of the effect of the treatment on sanctions (see Abbring and Van den Berg, 2005, and Van den Berg, Hofmann et al., 2024, for detailed discussions). This can be remedied by making a Conditional Independence Assumption (CIA). However, this is not straightforward to justify, as caseworkers may fine-tune the timing of a sanction based on worker characteristics that are unobserved in the data. A similar problem arises if we consider the timing of other types of sanctions and the actual timing of IAs in Section 5, because they are also right-censored by exits out of welfare. These important caveats should be borne in mind when assessing the corresponding results below.

Where we obtain statistically significant point estimates, we additionally apply the Romano-Wolf multiple hypothesis correction (Romano and Wolf 2005, 2016) using the Stata ado-file *rwolf* (Clarke et al. 2021), performing 250 bootstrap replications. This procedure accounts for the actual dependence structure among the test statistics through resampling, resulting in enhanced power compared to previous multiple-testing approaches. For this analysis, we consider all estimates using the same specification and sample as a family of tested hypotheses. When relevant, findings from this correction are discussed in the text.

5. Compliance

Before comparing the labor market outcomes across the three treatment arms, this section examines the extent to which participants in the field experiment actually received an IA immediately—as intended for treatment arms V1 and V2 —or only later in the placement process – as intended for treatment arm V3. In the case of incomplete compliance with the experimental protocol, results should be interpreted as intention-to-treat effects.[9]

Figure 1 presents Kaplan-Meier survival probabilities for individuals who received an initial IA (with or without a legal consequence statement) by a certain point in time. We do not censor this data when individuals leave the welfare system, but the results remain largely unchanged if we treat individuals as censored upon leaving. However, the caveat in Section 4 about the necessity of a CIA for causal

[9] Some sort of incomplete compliance is an inherent feature of most field experiments on active labor market policies. For instance, in an experiment comparing private and public provision of counseling to job seekers, Behaghel et al. (2014) observed a 40 percent compliance rate for both treatments. In an experiment on the private provision of counseling services, Bennmarker et al. (2013) reported a compliance rate of 28 percent.

inference also applies here. Therefore the results should be viewed with some caution and are only tentatively indicative of causal effects of the treatment on the timing of the IA. With this in mind, the results indicate that, on average, participants in group V3 receive an IA significantly less often and later than those in the other two groups, V1 and V2, as intended by the experimental protocol. After 180 days, almost all individuals in groups V1 and V2 have concluded an IA. However, 51 percent of those in group V3 also concluded an IA within 180 days after random assignment. Figure 2 shows the survival rates for the second aspect of our design. Although not perfect, compliance rates were substantially higher regarding the inclusion of a LCD in the IA. Only around 19 percent of individuals in group V3 and approximately 25 percent in group V2 received a declaration within 180 days after assignment, while the vast majority of individuals in treatment arm V1 received a LCD.

[Figure 1 and Figure 2 about here]

Table A.2 in the Online Appendix illustrates the extent to which non-compliance is correlated with the covariates used in our estimates. The strongest impact is observed for the month of assignment: compliance rates were significantly better for individuals who entered the project up to March 2018. Therefore, we conduct a sensitivity analysis that restricts the analysis to assignments made up to that calendar month. Otherwise, there are relatively few significant correlations. Non-compliance with V3 (no IA) increased with the proportion of time individuals spent employed during the year prior to random assignment. This may be attributed to caseworkers believing that individuals who worked a larger share of the previous year have the highest potential for reintegration into the labor market and that all available means should be used to incentivize their job search. Furthermore, we also observe some differences across job centers.

The trial was designed so that no sanctions were imposed for the neglect of duties in groups V2 and V3 during the first six months after allocation; only failures to report could be sanctioned. Table A.3 in the Online Appendix shows that, up to the 180th day after assignment, two percent of individuals in group V3 were sanctioned accordingly. Two years after assignment, the corresponding proportion is approximately four percent. Differences across the three groups are consistently shown to be not statistically significant. Overall, benefit recipients in all three groups received benefit cuts due to violations of obligations relatively rarely. However, we observe some degree of non-compliance in groups V2 and V3.

As mentioned above, unemployed welfare recipients from all three groups were also expected to have the same access to services and ALMPs. Table A.3 in the Online Appendix also provides information on entries into active labor market programs. In group V3, approximately 27 percent of participants participated in an ALMP within 180 days after assignment, while 43 percent did so during the 720 day period.[10] We find no significant effect of group assignment, which is in line with the experimental protocol.

To conclude, the treatment arm affects the probability of receiving an IA with or without a legal consequence instruction, but the estimated effects reported in the results section below should be interpreted as intention-to-treat effects. Although compliance is incomplete, compliance rates are within the range of what has been observed in other randomized controlled trials on labor market policy instruments.

6. Results

Main findings

Figure 3 illustrates the transitions of welfare recipients out of UB II receipt within two years after random assignment (these exits do not have to be permanent). Panel I of Table 2 presents the results from linear probability models analyzing exits from welfare receipt within 180, 360, and 720 days, controlling for a large set of covariates. The results indicate that, in all three groups, approximately 70 percent of individuals were able to leave benefit receipt at least temporarily within two years after random assignment. However, exit rates do not differ significantly between the three groups; this holds true both without and with controlling for covariates.

Figure 4 shows the survival curves until taking up employment subject to social insurance during the two years after random assignment, while Panel II of Table 2 presents the respective estimates for exits until 180, 360, and 720 days after random assignment. In all three groups, the cumulative proportion at the end of the observation period is approximately 62 percent. Once again, the assignment had no significant effect on transitions into employment, and this result holds true when controlling for covariates.[11]

[Table 2, Figure 3, and Figure 4 about here]

Sanctions for failure to report could also vary as a result of group assignment, as they were not covered by the experimental protocol. It is conceivable that individuals in groups V2 and V3—who did not receive a LCD or any IA at all—took their reporting obligations less seriously than those who had

[10] Most often, individuals participated in short activation measures. The second largest program category used are further training programs. Very few individuals took part in other programs.

[11] Additionally, Table A.4 in the Appendix presents the covariate effects on exits from welfare benefit receipt and entries into employment. Both transitions occur less frequently with increasing age and among single parents with children. They are more commonly observed for individuals with a vocational or university degree and for those who worked a larger share of the year prior to random assignment. There are also differences between job centers; however, these are much more pronounced regarding exits from UB II receipt and are are predominantly insignificant in relation to entries into employment.

concluded a standard IA. It is also possible that placement officers were more inclined to sanction reporting failures if, in accordance with the guidelines, they did not impose any sanctions for reporting failures in the first six months after assignment to these two groups. Indeed, Figure 5 and Panel III in Table 2 show that by the 180th day after assignment, around 3 percent in group V1 and 5 percent in the other two groups had received a sanction for failure to report (a different scale is used here compared to the previous figures). After 720 days, the proportions increased to approximately 10, 11, and 12 percent. However, the overall difference between the transition rates of the groups is not significant, regardless of whether covariates are accounted for in the regressions.

[Figure 5 about here]

Heterogeneous effects

Even in the absence of average treatment effects, there may be heterogeneous effects for subgroups. When analyzing the effects of IAs in the unemployment insurance system, Van den Berg, Hofmann et al. (2024) stratified the sample based on the predicted median unemployment duration until re-employment, distinguishing between individuals with a predicted median duration of less than or more than six months. For the prediction model, they used all inflows into unemployment in the same regions during the year prior to the RCT. They found that IAs did not speed up re-employment for individuals with good labor market prospects, while early IAs had significantly positive effects for individuals with lower re-employment prospects.

It is therefore a relevant question whether similar effect heterogeneity arises in the welfare system. Consequently, we stratify our sample according to predicted labor market prospects. For this, we utilize a sample of inflows into welfare receipt from the seven job centers that participated in the RCT during the period from June 2016 to May 2017, the calendar year before our project started. To align with our analysis sample from the experiment, we consider only individuals who had not previously received welfare for at least six months, were aged 25 to 61, and were not employed when entering welfare. As has been observed in Van den Berg et al. (2023), standard parametric prediction models for long-term unemployment are dominated by Machine Learning predictors based on a Random Forest. Hence, we advance on the approach in Van den Berg, Hofmann et al. (2024) by making a Random Forest and conducting out-of-sample predictions regarding exiting the UB II regime (not necessarily permanently) within 360 days for our analysis sample. We include all covariates used as control variables in the main analysis (utilizing the calendar month of welfare entry or assignment, respectively, to control for potential seasonal effects). We then split the sample at the median of the prediction score obtaining two equally sized groups ordered by their estimated prospects of leaving the welfare system.[12]

Table 3 displays the main results. The constants of the models indicate that the mean share of transitions out of welfare benefit receipt and into employment during the first year after assignment is approximately 19 and 16 percentage points higher in the group with predicted better labor market prospects. However, we do not find significant effects of group assignment for individuals in either group. There is one exception: longer-term entries into employment seem to be lower in group V2 (IA without a LCD). However, once we correct for multiple testing, this effect also becomes insignificant.

[Table 3 about here]

Sensitivity analysis

First, as motivated by the non-compliance findings in Section 5, we repeat our main analysis for assignments to the program that occurred until March 2018. Figure A.1 in the Appendix displays survival curves until receipt of the first IA for this group: After 180 days, around 30 percent of individuals in group V3 (no IA) had received an IA (contrary to the experimental protocol), providing tentative evidence that non-compliance is considerably lower than in the full sample.

The corresponding impact estimates can be found in Table 4. Even for earlier assignments we do not find significant effects among the different treatment groups concerning entries into employment or exits from UB II. In the short run, we observe a single significant negative effect of an assignment to the group with the standard IA (V1) on the share of individuals who received a sanction due to a reporting failure. This aligns with the previously stated hypothesis that individuals in groups V2 and V3 may have taken their reporting obligations less seriously, resulting in more sanctions due to reporting failures compared to those who had concluded a standard IA. Note, however, that once we correct for multiple testing, the effect is no longer significant.

Table 5 presents heterogeneous effects based on estimated labor market prospects. The assignment did not impact exits from welfare benefit receipt. For the group predicted not to leave welfare within 360 days, we identify a significant negative effect of an assignment to V2 (IA without an LCD) on longer-term exits into employment (or, equivalently, a significant positive effect of assignment to V1 or V3). Furthermore, the previously mentioned effect of receiving a sanction due to a failure to report

[12] We apply the stata ado file *rforest* (Schonlau and Zou 2020) for a dichotomous outcome. Tuning the model resulted in choosing the following hyperparameters: 235 iterations, 15 variables randomly selected at each split, maximum tree depth of 15, and a minimum size of 11 observations per leaf node. As a threshold for the heterogeneity analysis we use the median of the predicted probability, which amounts to 0.44. For this threshold, the accuracy of the model accounts to 59 percent; the true negative rate is 60 percent, and the true positive rate is 58 percent.

also originates from this group. If we correct for multiple testing, these effects are no longer significant.[13]

[Table 4 and Table 5 about here]

7. Concluding remarks

IAs are intended to be a central element of the counseling and placement activities for unemployed welfare benefit recipients. They also serve as a legal basis for sanctioning neglect of job search-related duties. By employing a randomized controlled trial that varied the timing and content of IAs, we aim to uncover whether IAs genuinely enhance employment prospects and reduce dependency on welfare.

For individuals who have newly started receiving UB II, results from our randomized controlled trial indicate that assignment to the groups expected to receive an IA (with and without LCDs) had, on average, no impact on shortening the duration of benefit receipt or the time until welfare recipients found new employment. Due to incomplete compliance, our results must be interpreted as intention-to-treat effects. However, even when we restrict the analysis to earlier assignments, which are much more consistent with the experimental protocol, we do not find evidence that IAs increased the rate of exits from welfare benefit receipt or entries into employment.

We do not find strong evidence for an effect of signing an IA on welfare exit or employment entry for unemployed welfare recipients. This may be due to the complex bureaucracy surrounding the IA as a treatment, or simply to the lack of suitable job opportunities for most welfare recipients. Another potential reason might be due to the fact that the baseline survival probabilities are simply very high, such that a sizeable effect on the hazard rate only leads to a small effect on the survival probability. This phenomenon has been observed in the literature on sanction effects in welfare; see Van den Berg et al. (2004). However, investigating this would require semi-parametric modeling. Yet another potential reason could be a lack of power due to a limited number of observations. However, the precision of the estimates reported in this paper is comparable to the precision in the study on the

[13] An interesting issue is whether we could use the assignment as an instrument for obtaining an IA and a LCD. However, the exclusion restriction may be violated if individuals realize that they are not subject to the regime. In such cases, the assignment could directly affect the outcome if their behavior changes before an IA or LCD is finalized. If we assume the exclusion restriction holds, a simple non-parametric Wald estimator would assess differences in outcome variables between assignment groups default by examining differences in realized treatments. This approach would inflate point estimates and standard errors—but would not necessarily increase significance. For a general discussion about instrumental variables with duration outcomes in the context of social experiments see Abbring and van den Berg (2005).

effectiveness of IAs for male jobseekers receiving UI benefits, which reports significantly positive effects of being supposed to sign an IA in the first month of the UI spell on the probability of entering employment within one year (Van den Berg, Hofmann et al. 2024). Further, the reported average effect in Van den Berg, Hofmann et al. (2024) of signing an IA early in the UI spell on exit to employment within one year is outside of the confidence interval of the estimated effects for welfare benefit recipients. This does not exclude that we would detect an effect with a larger sample size, but we can reject the hypothesis that the ITT effect of signing an IA early in the spell is of the same magnitude as the corresponding effects estimated for UI benefit recipients.

In either case, the results imply that the success of IAs in UI does not translate into the welfare system. As a topic for further research it may be interesting to consider more specific subgroups of welfare recipients. However, it may be more promising to consider other ALMPs with a less strong monitoring flavor as IAs.

References

Abbring, Jaap; Van den Berg, Gerard J. (2005): Social experiments and instrumental variables with duration outcomes, Tinbergen Institute Discussion Paper TI 05-047/3.

Behaghel, Luc; Crépon, Bruno; Gurgand, Marc (2014): Private and public provision of counseling to job seekers: Evidence from a large controlled experiment, American Economic Journal: Applied Economics 6, 142-174.

Bennmarker, Helge; Grönqvist, Erik; Öckert, Björn (2013): Effects of contracting out employment services: Evidence from a randomized experiment, Journal of Public Economics 98, 68-84.

Bernhard, Sarah; Senghaas, Monika (2021): Eingliederungsvereinbarungen im Jobcenter schaffen Verbindlichkeit, aber die Mitwirkungspflichten dominieren (Serie "Befunde aus der IAB-Grundsicherungsforschung 2017 bis 2020"), IAB-Forum, 07.07.2021.

Bundesrechnungshof (2019): Bericht an den Haushaltsausschuss des Deutschen Bundestages nach § 88 Abs. 2 BHO – Gesetzlicher Reformbedarf bei Eingliederungsvereinbarungen in den Rechtskreisen SGB III und SGB III, (last accessed on 17.6.2022).

Clarke, Damian; Romano, Joseph P.; Wolf, Michael (2021): The Romano-Wolf multiple-testing correction in Stata. The Stata journal, 20(4).

Crépon, Bruno; van den Berg, Gerard J. (2016): Active labor market policies. Annual Review of Economics, 8(1), 521-546.

Gerards, Ruuds, Welters, Riccardo (2022): Does eliminating benefit eligibility requirements improve unemployed job search and labour market outcomes?. Applied Economics Letters 29, 955-958.

Knotz, Carlo Michael (2018): A Rising Workfare State? Unemployment Benefit Conditionality in 21 OECD Countries, 1980-2012, Journal of International and Comparative Social Policy, 34(2), 92-108.

Romano, Joseph P.; Wolf, Michael (2005). Exact and approximate stepdown methods for multiple-hypothesis-testing. Journal of the American Statistical Association, 100(469), 94-108.

Romano, Joseph P.; Wolf, Michael (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Statistics & Probability Letters, 113, 38-40.

Schonlau, Matthias; Zou, Rosie Y. (2020): The random forest algorithm for statistical learning. The Stata Journal 20, 3-29.

Schiprowski, Amelie; Schmidtke, Julia; Schmieder, Johannes; Trenkle, Simon (2024): The Effects of Unemployment Insurance Caseworkers on Job Search Effort, AEA Papers and Proceedings 114, 567-571.

Senghaas, Monika; Bernhard, Sarah; Freier, Carolin (2020): Eingliederungsvereinbarungen aus Sicht der Jobcenter: Pflichten der Arbeitsuchenden nehmen viel Raum ein, IAB-Kurzbericht 05/2020.

Senghaas, Monika; Bernhard, Sarah (2021): Arbeitsvermittlung im Spannungsfeld von Dienstleistung und Kontrolle – Eine multimethodische Studie zu Eingliederungsvereinbarungen in der Grundsicherung für Arbeitsuchende, Sozialer Fortschritt 70, 487-507.

Ulmestig, Rickard,; Denvall, Verner; Nordesjö, Kettil (2020): Claiming equality and "doing" inequality - Individual action plans for applicants of social assistance, Social Work & Society 18.

Van den Berg, Gerard J.; Kunaschk, Max; Lang, Julia; Stephan, Gesine; Uhlendorff, Arne (2023): Predicting re-employment: machine learning versus assessments by unemployed workers and by their caseworkers, IZA Discussion Paper.

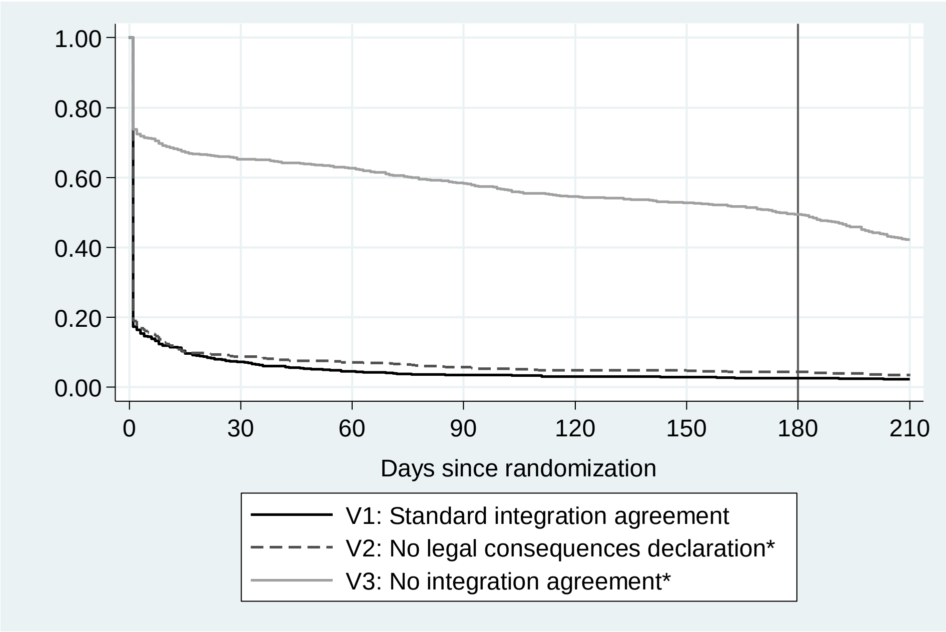
Van den Berg, Gerard J.; Kesternich, Iris; Müller, Gerrit; Siflinger, Bettina (2024): Reciprocity and the interaction between the unemployed and the caseworker, Journal of Economic Behavior and Organization 227, 106706.

Van den Berg, Gerard J.; Hofmann, Barbara; Stephan, Gesine; Uhlendorff, Arne (2024): Mandatory integration agreements for unemployed job seekers: a randomized controlled field experiment in Germany, International Economic Review, forthcoming.

Van den Berg, Gerard; Van der Klaauw, Bas; Van Ours, Jan (2004). Punitive sanctions and the transition rate from welfare to work. Journal of Labor Economics 22, 211-241.

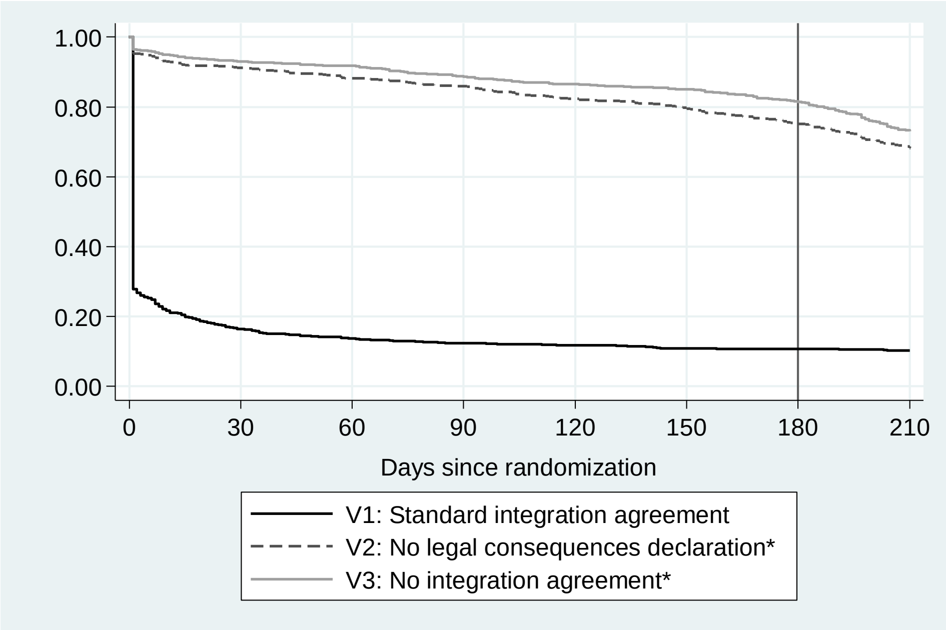
Figures

Figure 1 Duration until the conclusion of the first integration agreement



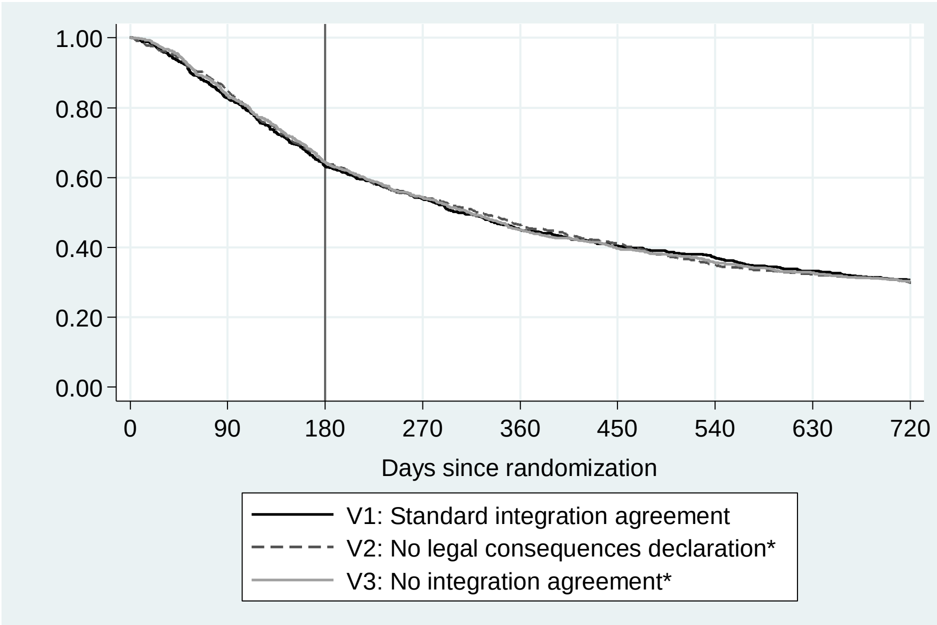
Notes: *) For a period of six months since assignment. Survival functions from Kaplan-Meier estimation. Log-rank test for equality of survival functions: Pr>chi2 = 0.00. 2,659 observations.

Figure 2 Duration until the first LCD



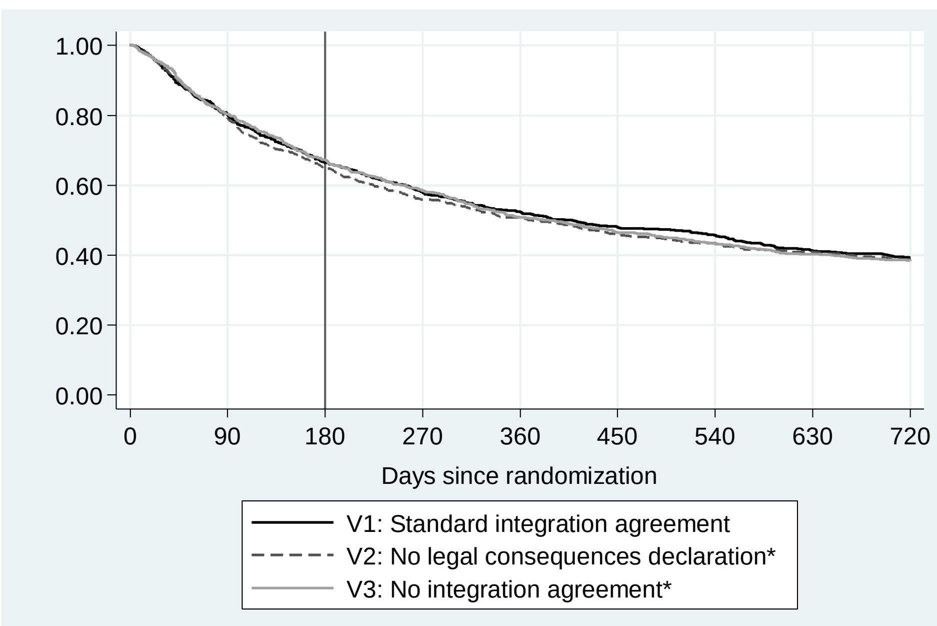
Notes: *) For a period of six months since assignment. Survival functions from Kaplan-Meier estimation. Log-rank test for equality of survival functions: Pr>chi2 = 0.00. 2,659 observations.

Figure 3 Duration until exiting UB II receipt



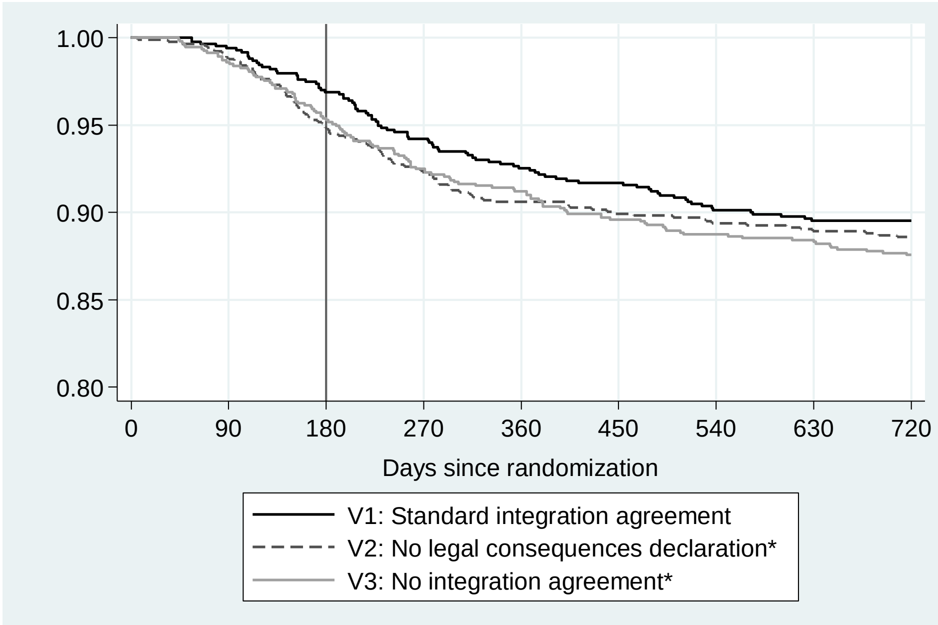
Notes: *) For a period of six months since assignment. Survival functions from Kaplan-Meier estimation (uncensored data). Log-rank test for equality of survival functions: : Pr>chi2 = 1.00. 2,659 observations.

Figure 4 Duration until entering employment



Notes: *) For a period of six months since assignment. Survival functions from Kaplan-Meier estimation (uncensored data). Log-rank test for equality of survival functions: Pr>chi2 = 0.90, 2,659 observations.

Figure 5 Duration until first sanction due to failure to report



Notes: *) For a period of six months since assignment. Survival functions from Kaplan-Meier estimation. Log-rank test for equality of survival functions: Pr>chi2 = 0.42. 2.659 observations.

Tables

Table 1 Treatment variants

	V1: Standard IA	**V2:** No legal consequences*	**V3:** No IA*
Contents of the integration agreement)			
Rights and obligations	x	x	-
Legal consequence information	x	-	-
Sanctions			
Neglect of duties from the integration agreement**	x	-	-
Failure to report	x	x	x

*) For a period of six months since assignment.
**) Lack of own efforts, refusal to take up work or training or participation in a labor market program.

Table 2 Intention-to-treat effects, controlling for covariates
Estimated coefficients and standards errors (in parenthesis)
Reference group: V3 (no IA)

	Until day…		
	180	360	720
I. Exit from UB II receipt			
V1: Standard IA	0.006	-0.007	-0.016
	(0.022)	(0.023)	(0.021)
V2: IA without LCD	-0.004	-0.017	-0.003
	(0.022)	(0.022)	(0.021)
Mean V3	0.357	0.550	0.698
II. Entry into employment			
V1: Standard IA	-0.000	-0.021	-0.020
	(0.022)	(0.023)	(0.022)
V2: IA without LCD	0.017	-0.006	-0.014
	(0.022)	(0.023)	(0.022)
Mean V3	0.328	0.492	0.616
III. Sanction of failure to report			
V1: Standard IA	-0.015	-0.013	-0.020
	(0.010)	(0.013)	(0.015)
V2: IA without LCD	0.005	0.007	-0.010
	(0.009)	(0.013)	(0.014)
Mean V3	0.047	0.088	0.124
Observations		2,659	

Notes: Linear probability model. For the list of included covariates, see Table A.1. IA = integration agreement, LCD = legal consequences declaration.

Table 3 Intention-to-treat effects by predicted labor market prospects, controlling for covariates

Estimated coefficients and standards errors (in parenthesis)

Reference group: V3 (no IA)

	Predicted to leave welfare within 360 days					
		No			Yes	
Until day...	180	360	720	180	360	720
I. Exit from UB II receipt						
V1: Standard IA	0.023	0.017	-0.004	-0.005	-0.035	-0.033
	(0.029)	(0.032)	(0.032)	(0.033)	(0.032)	(0.028)
V2: IA without LCD	-0.021	-0.036	-0.012	-0.004	-0.017	-0.005
	(0.030)	(0.033)	(0.032)	(0.032)	(0.031)	(0.027)
Mean V3	0.267	0.458	0.623	0.453	0.649	0.778
II. Entry into employment						
V1: Standard IA	0.018	-0.023	-0.029	-0.015	-0.019	-0.019
	(0.029)	(0.032)	(0.032)	(0.034)	(0.034)	(0.031)
V2: IA without LCD	0.003	-0.025	-0.067**	0.011	-0.009	0.016
	(0.029)	(0.032)	(0.032)	(0.032)	(0.032)	(0.030)
Mean V3	0.248	0.414	0.553	0.413	0.576	0.684
III. Sanction of failure to report						
V1: Standard IA	-0.022	-0.024	-0.026	-0.007	-0.001	-0.009
	(0.015)	(0.020)	(0.023)	(0.012)	(0.016)	(0.019)
V2: IA without LCD	0.017	0.018	0.007	-0.003	0.003	-0.017
	(0.015)	(0.021)	(0.023)	(0.011)	(0.015)	(0.018)
Mean V3	0.060	0.114	0.153	0.033	0.060	0.093
Observations		1,330			1,329	

) α = 0.05, *) α = 0.01.

Notes: Linear probability model. For list of covariates, see Table A.1. IA = integration agreement, LCD = legal consequences declaration.

Table 4 Intention-to-treat effects, controlling for covariates, for individuals who were assigned until March 2018

Estimated coefficients and standards errors (in parenthesis)

Reference group: V3 (no IA)

	Until day…		
	180	360	720
I. Exit from UB II receipt			
V1: Standard IA	0.013	-0.009	-0.030
	(0.028)	(0.029)	(0.027)
V2: IA without LCD	-0.019	-0.024	-0.010
	(0.028)	(0.028)	(0.026)
Mean V3	0.354	0.547	0.707
II. Entry into employment			
V1: Standard IA	0.016	-0.001	-0.014
	(0.028)	(0.030)	(0.029)
V2: IA without LCD	0.002	-0.009	-0.017
	(0.027)	(0.029)	(0.028)
Mean V3	0.319	0.477	0.613
III. Sanction of failure to report			
V1: Standard IA	-0.028**	-0.021	-0.018
	(0.012)	(0.017)	(0.019)
V2: IA without LCD	-0.007	-0.006	-0.016
	(0.012)	(0.016)	(0.018)
Mean V3	0.056	0.095	0.127
Observations	1,657		

Notes: Linear probability model. For list of covariates, see Table 2. IA = integration agreement, LCD = legal consequences declaration.

Table 5 Intention-to-treat effects by predicted labor market prospects, controlling for covariates, for individuals who were assigned until March 2018

Estimated coefficients and standards errors (in parenthesis)

Reference group: V3 (no IA)

	Predicted to leave welfare within 360 days					
		No			**Yes**	
Until day...	180	360	720	180	360	720
I. Exit from UB II receipt						
V1: Standard IA	0.019	-0.012	-0.035	0.025	-0.010	-0.031
	(0.038)	(0.042)	(0.040)	(0.043)	(0.041)	(0.036)
V2: IA without LCD	-0.056	-0.067	-0.043	-0.001	-0.007	0.006
	(0.039)	(0.042)	(0.041)	(0.039)	(0.038)	(0.033)
Mean V3	0.270	0.470	0.646	0.436	0.622	0.766
II. Entry into employment						
V1: Standard IA	0.026	-0.006	-0.029	0.016	0.004	-0.015
	(0.037)	(0.041)	(0.042)	(0.043)	(0.043)	(0.040)
V2: IA without LCD	-0.025	-0.044	-0.096**	0.011	-0.002	0.026
	(0.038)	(0.042)	(0.043)	(0.040)	(0.039)	(0.036)
Mean V3	0.232	0.386	0.540	0.405	0.567	0.684
III. Sanction of failure to report						
V1: Standard IA	-0.046**	-0.045*	-0.036	-0.009	0.000	-0.002
	(0.020)	(0.027)	(0.030)	(0.015)	(0.020)	(0.024)
V2: IA without LCD	-0.000	0.009	0.004	-0.006	-0.011	-0.024
	(0.020)	(0.028)	(0.030)	(0.014)	(0.019)	(0.022)
Mean V3	0.074	0.126	0.158	0.038	0.065	0.096
Observations		805			852	

Notes: Linear probability model. For list of covariates, see Table 2. IA = integration agreement, LCD = legal consequences declaration.

Appendix

Table A.1 Means of selected variables for the three treatment arms and *p*-values from F-tests for equal means (dichotomous variables) and χ^2-tests on equal distributions (categorial variables)

	V1: Default	V2: No legal consequeces	V3: No IA	*p*-value
Gender and nationality				
Female	0.33	0.34	0.35	0.70
Non-German nationality	0.34	0.37	0.36	0.43
Age group				
Age group 25-29	0.27	0.27	0.27	0.61
Age group 30-39	0.36	0.34	0.34	
Age group 40-49	0.20	0.22	0.20	
Age group 50-61	0.17	0.17	0.20	
Vocational education				
No vocational degree	0.36	0.36	0.35	0.54
Vocational degree	0.50	0.50	0.53	
University degree	0.14	0.12	0.11	
Information missing	0.01	0.02	0.02	
Household type				
Single	0.70	0.71	0.71	0.61
Partners without children	0.09	0.08	0.10	
Single parent with children	0.08	0.07	0.07	
Partners with children	0.13	0.13	0.11	
Information missing	0.01	0.01	0.01	
Position last job				
Helper activities	0.34	0.35	0.37	0.24
Professional activities	0.45	0.46	0.48	
Complex specialist activities	0.06	0.05	0.05	
Highly complex activities	0.06	0.06	0.04	
Information missing	0.09	0.08	0.07	
During the last year before random assignment: Share of year in				
Employment	0.14	0.16	0.15	0.35
UB receipt	0.21	0.20	0.20	0.46
UB II receipt	0.13	0.15	0.15	0.09
During the five years preceding the last year before random assignment: Years in				
Employment	1.28	1.28	1.29	1.00
UB receipt	0.30	0.27	0.28	0.55
UB II receipt	1.23	1.21	1.24	0.95
Jobcenter				
1	0.30	0.31	0.30	0.72
2	0.22	0.19	0.22	
3	0.11	0.10	0.10	
4	0.15	0.16	0.14	
5	0.05	0.04	0.05	
6	0.04	0.04	0.05	
7	0.14	0.16	0.14	

Table A.1 continued

	V1: Default	V2: No legal consequeces	V3: No IA	*p*-value
Month of assignment				
Up to September 2017	0.08	0.07	0.09	0.35
October 2017	0.07	0.10	0.09	
November 2017	0.09	0.09	0.08	
December 2017	0.06	0.06	0.08	
January 2018	0.11	0.11	0.11	
February 2018	0.09	0.10	0.08	
March 2018	0.10	0.11	0.08	
April 2018	0.13	0.09	0.11	
May 2018	0.08	0.08	0.09	
June 2018	0.08	0.07	0.07	
July 2018	0.05	0.04	0.06	
August 2018 and later	0.05	0.06	0.06	
Observations	832	894	933	

Table A.2 Determinants of non-compliance (within 180 days of random assignment)
Estimated coefficients and standards errors (in parenthesis)
Dependent variable: Non-compliance with prescribed treatment arm

	V1: Default	V2: No LCD	V3: No IA
Gender and nationality			
Female	-0.017	0.005	0.030
	(0.026)	(0.034)	(0.032)
Non-German nationality	-0.019	-0.018	-0.026
	(0.026)	(0.034)	(0.033)
Age group (reference: up to 29)			
Age group 30-39	0.001	0.019	0.035
	(0.030)	(0.041)	(0.037)
Age group 40-49	-0.037	0.050	-0.010
	(0.035)	(0.046)	(0.045)
Age group 50-61	-0.060*	-0.005	-0.025
	(0.036)	(0.049)	(0.045)
Vocational education (reference: no vocational degree)			
Vocational degree	0.022	-0.013	-0.011
	(0.027)	(0.036)	(0.033)
University degree	0.091**	-0.032	-0.003
	(0.038)	(0.054)	(0.053)
Information missing	0.067	-0.079	0.319***
	(0.112)	(0.127)	(0.119)
Household type (reference: single)			
Partners without children	0.058	0.006	-0.017
	(0.042)	(0.055)	(0.050)
Single parent with children	-0.020	-0.112*	-0.033
	(0.046)	(0.063)	(0.061)
Partners with children	0.044	-0.014	0.000
	(0.035)	(0.048)	(0.048)
Information missing	0.327**	-0.175	-0.051
	(0.145)	(0.161)	(0.134)
Position last job (reference: helper activities)			
Professional activities	0.014	-0.023	0.019
	(0.026)	(0.035)	(0.031)
Complex specialist activities	-0.019	-0.027	0.048
	(0.051)	(0.070)	(0.071)
Highly complex activities	-0.007	0.021	0.067
	(0.052)	(0.069)	(0.078)
Information missing	-0.053	-0.069	0.097
	(0.046)	(0.064)	(0.064)

Continued next page

Table A.2 continued

	V1: Default	V2: No LCD	V3: No IA
During the last year before random assignment: Share of year in			
Employment	0.089*	-0.037	0.186***
	(0.051)	(0.064)	(0.061)
UB receipt	-0.092**	-0.109*	0.002
	(0.044)	(0.060)	(0.058)
UB II receipt	0.016	0.138*	0.113
	(0.062)	(0.074)	(0.070)
During the five years before the last year before random assignment: Years in			
Employment	0.017	-0.001	0.010
	(0.011)	(0.015)	(0.015)
UB receipt	-0.023	-0.029	0.018
	(0.028)	(0.036)	(0.035)
UB II receipt	-0.010	0.008	-0.001
	(0.008)	(0.012)	(0.011)
Jobcenter (reference: 1)			
2	-0.022	0.055	0.019
	(0.032)	(0.046)	(0.042)
3	0.030	0.299***	0.121**
	(0.040)	(0.055)	(0.053)
4	-0.051	-0.050	0.005
	(0.036)	(0.047)	(0.048)
5	0.270***	0.015	0.041
	(0.057)	(0.082)	(0.070)
6	-0.040	0.017	0.041
	(0.057)	(0.079)	(0.071)
7	-0.007	0.151***	0.073
	(0.038)	(0.048)	(0.048)
Month of assignment (reference: up to September 2017)			
October 2017	-0.152***	0.086	0.077
	(0.057)	(0.073)	(0.068)
November 2017	-0.137**	0.142*	0.236***
	(0.054)	(0.075)	(0.070)
December 2017	-0.130**	0.085	0.269***
	(0.060)	(0.081)	(0.071)
January 2018	-0.080	0.272***	0.051
	(0.052)	(0.071)	(0.064)
February 2018	-0.051	0.193***	0.094
	(0.054)	(0.072)	(0.069)
March 2018	-0.029	0.183**	0.364***
	(0.053)	(0.072)	(0.070)
April 2018	-0.055	0.164**	0.688***
	(0.049)	(0.074)	(0.066)

Continued next page

Table A.2 continued

	V1: Default	V2: No LCD	V3: No IA
May 2018	-0.072	0.201***	0.725***
	(0.056)	(0.077)	(0.068)
June 2018	-0.105*	0.216***	0.609***
	(0.055)	(0.079)	(0.073)
July 2018	-0.023	0.297***	0.677***
	(0.063)	(0.092)	(0.079)
August 2018 and later	-0.109*	0.258***	0.625***
	(0.066)	(0.085)	(0.080)
Constant	0.188***	0.091	0.053
	(0.056)	(0.076)	(0.073)
R-squared	0.090	0.102	0.309
Observations	832	894	933

*) α = 0.10, **) α = 0.05, ***) α = 0.01.
Notes: Linear probability model. For list of covariates, see Table 2. IA = integration agreement, LCD = legal consequences declaration.

Table A.3 Intention-to-treat effects on entries into active labor market programs, controlling for covariates

Estimated coefficients and standards errors (in parenthesis)

Reference group: V3 (no IA)

	Until day…		
	180	360	720
I. Sanction of duty neglect			
V1: Standard IA	0.007	0.011	0.010
	(0.006)	(0.009)	(0.010)
V2: IA without LCD	-0.001	0.005	0.002
	(0.006)	(0.009)	(0.010)
Mean V3	0.016	0.030	0.044
II. Entry into ALMP			
V1: Standard IA	0.005	-0.012	0.002
	(0.020)	(0.022)	(0.023)
V2: IA without LCD	-0.016	-0.028	-0.027
	(0.020)	(0.021)	(0.023)
Mean V3	0.272	0.346	0.427
Observations		2,659	

Notes: Linear probability model. For the list of included covariates, see Table ~~2~~A.1. IA = integration agreement, LCD = legal consequences declaration, ALMP = active labor market programs.

Table A.4 Covariate effects on exits from UB II receipt and entries into employment
Estimated coefficients and standards errors (in parenthesis)

	Exit from UB II receipt until day…			Entry into employment until day…		
	180	360	720	180	360	720
Gender and nationality						
Female	-0.014	-0.026	-0.011	-0.027	-0.035	-0.017
	(0.020)	(0.021)	(0.019)	(0.020)	(0.021)	(0.021)
Non-German nationality	-0.001	-0.007	0.006	-0.007	-0.006	0.017
	(0.021)	(0.022)	(0.020)	(0.021)	(0.022)	(0.021)
Age group (reference: up to 29)						
Age group 30-39	-0.102***	-0.074***	-0.036	-0.095***	-0.094***	-0.055**
	(0.024)	(0.025)	(0.023)	(0.024)	(0.025)	(0.024)
Age group 40-49	-0.122***	-0.151***	-0.116***	-0.100***	-0.116***	-0.090***
	(0.028)	(0.029)	(0.027)	(0.028)	(0.030)	(0.029)
Age group 50-61	-0.190***	-0.231***	-0.222***	-0.171***	-0.225***	-0.223***
	(0.029)	(0.030)	(0.028)	(0.029)	(0.031)	(0.029)
Vocational education (reference: no vocational degree)						
Vocational degree	0.070***	0.037*	0.045**	0.065***	0.060***	0.054**
	(0.021)	(0.022)	(0.020)	(0.021)	(0.023)	(0.022)
University degree	0.140***	0.152***	0.150***	0.111***	0.158***	0.162***
	(0.032)	(0.033)	(0.031)	(0.032)	(0.034)	(0.033)
Information missing	0.102	0.166**	0.079	-0.036	-0.004	0.082
	(0.080)	(0.082)	(0.076)	(0.080)	(0.084)	(0.081)
Household type (reference: single)						
Partners without children	0.076**	0.089***	0.052*	-0.031	-0.006	-0.018
	(0.033)	(0.033)	(0.031)	(0.032)	(0.034)	(0.033)
Single parent with children	-0.090**	-0.114***	-0.115***	-0.087**	-0.054	-0.018
	(0.038)	(0.039)	(0.036)	(0.038)	(0.040)	(0.039)
Partners with children	0.003	-0.008	-0.007	0.068**	0.039	0.042
	(0.029)	(0.030)	(0.028)	(0.029)	(0.031)	(0.030)
Information missing	0.023	0.135	0.044	0.115	0.080	0.036
	(0.097)	(0.099)	(0.092)	(0.096)	(0.101)	(0.098)
Position last job (reference: helper activities)						
Professional activities	0.012	0.036*	0.015	-0.000	-0.009	-0.030
	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)	(0.021)
Complex specialist activities	-0.001	0.035	0.009	-0.029	-0.050	-0.095**
	(0.043)	(0.044)	(0.041)	(0.043)	(0.045)	(0.044)
Highly complex activities	0.050	0.048	0.037	-0.026	-0.054	-0.044
	(0.044)	(0.045)	(0.042)	(0.044)	(0.046)	(0.045)
Information missing	-0.087**	-0.062	-0.000	0.054	0.051	0.161***
	(0.039)	(0.040)	(0.037)	(0.039)	(0.041)	(0.039)

Continued next page

Table A.4 continued

	Exit from UB II receipt until day…			Entry into employment until day…		
	180	360	720	180	360	720
During the last year before random assignment: Share of year in						
Employment	0.202***	0.270***	0.185***	0.258***	0.312***	0.346***
	(0.040)	(0.041)	(0.038)	(0.039)	(0.042)	(0.040)
UB receipt	0.006	0.049	0.058*	0.068*	0.107***	0.144***
	(0.036)	(0.037)	(0.035)	(0.036)	(0.038)	(0.037)
UB II receipt	-0.135***	-0.188***	-0.160***	0.000	0.022	-0.033
	(0.046)	(0.047)	(0.044)	(0.046)	(0.048)	(0.047)
During the five years before the last year before random assignment: Years in						
Employment	-0.013	-0.013	-0.009	-0.003	-0.006	-0.009
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
UB receipt	-0.011	-0.024	-0.008	-0.001	-0.008	-0.022
	(0.022)	(0.023)	(0.021)	(0.022)	(0.023)	(0.022)
UB II receipt	-0.031***	-0.030***	-0.027***	-0.031***	-0.032***	-0.032***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Jobcenter (reference: 1)						
2	-0.044	-0.053*	-0.006	-0.016	-0.033	0.001
	(0.027)	(0.028)	(0.026)	(0.027)	(0.028)	(0.027)
3	0.104***	0.036	0.057*	0.086***	0.030	0.070**
	(0.033)	(0.034)	(0.032)	(0.033)	(0.035)	(0.034)
4	0.094***	0.088***	0.137***	0.041	0.042	0.047
	(0.029)	(0.030)	(0.028)	(0.029)	(0.031)	(0.030)
5	0.112**	0.160***	0.154***	0.072	0.063	0.012
	(0.047)	(0.048)	(0.045)	(0.046)	(0.049)	(0.047)
6	0.077*	0.127***	0.101**	0.041	0.063	0.081*
	(0.046)	(0.048)	(0.044)	(0.046)	(0.049)	(0.047)
7	0.026	0.036	0.066**	0.028	0.050	0.057*
	(0.030)	(0.031)	(0.029)	(0.030)	(0.032)	(0.031)
Month of assignment (reference: up to September 2017)						
October 2017	-0.053	-0.044	-0.052	-0.019	-0.010	0.029
	(0.045)	(0.046)	(0.043)	(0.044)	(0.047)	(0.045)
November 2017	-0.072	-0.050	0.005	-0.049	-0.082*	0.001
	(0.044)	(0.046)	(0.042)	(0.044)	(0.047)	(0.045)
December 2017	-0.043	-0.054	-0.050	-0.024	-0.012	0.046
	(0.047)	(0.049)	(0.045)	(0.047)	(0.050)	(0.048)
January 2018	0.021	0.027	-0.020	-0.022	-0.047	0.016
	(0.042)	(0.043)	(0.040)	(0.042)	(0.044)	(0.043)
February 2018	0.014	-0.045	-0.025	0.043	0.007	0.100**
	(0.044)	(0.045)	(0.042)	(0.044)	(0.046)	(0.045)
March 2018	0.002	-0.019	-0.023	0.011	-0.055	-0.008
	(0.044)	(0.045)	(0.042)	(0.044)	(0.046)	(0.044)
April 2018	0.052	0.040	0.023	0.086**	0.046	0.102**
	(0.042)	(0.043)	(0.040)	(0.042)	(0.044)	(0.043)

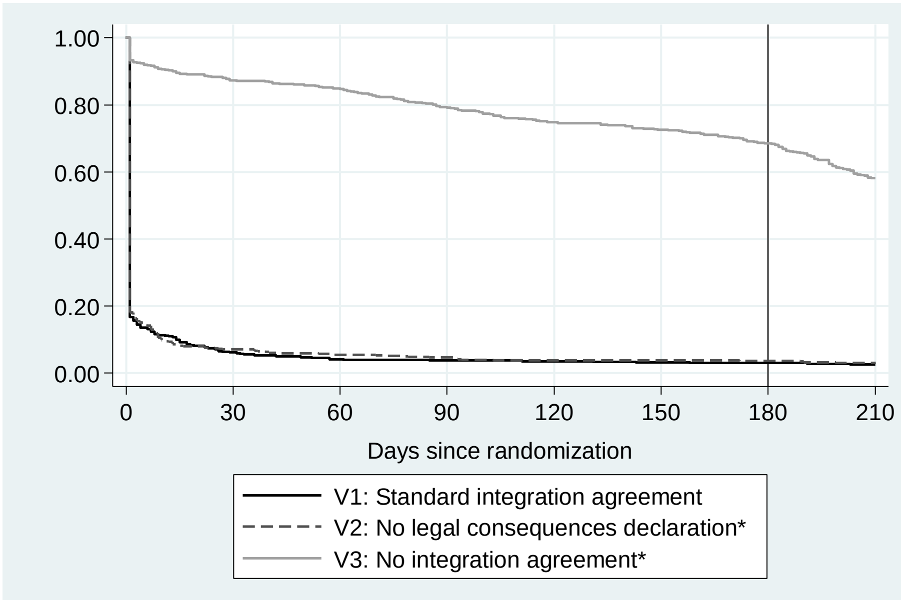
Continued next page

Table A.4 continued

	Exit from UB II receipt until day…			Entry into employment until day…		
	180	360	720	180	360	720
May 2018	-0.043	-0.029	-0.056	-0.037	-0.048	-0.022
	(0.045)	(0.046)	(0.043)	(0.045)	(0.047)	(0.046)
June 2018	-0.015	-0.013	-0.002	-0.003	-0.016	0.019
	(0.046)	(0.048)	(0.044)	(0.046)	(0.049)	(0.047)
July 2018	-0.062	-0.025	-0.076	-0.049	-0.041	0.019
	(0.052)	(0.054)	(0.050)	(0.052)	(0.055)	(0.053)
August 2018 and later	-0.028	-0.079	-0.062	0.009	0.013	0.051
	(0.052)	(0.053)	(0.049)	(0.015)	(0.020)	(0.024)
Constant	0.434***	0.638***	0.741***	0.020	0.033*	0.057***
	(0.047)	(0.048)	(0.045)	(0.014)	(0.019)	(0.021)
R-squared	0.088	0.106	0.093	0.068	0.075	0.093
Observations		2,659			2,659	

*) α = 0.10, **) α = 0.05, ***) α = 0.01.
Notes: Linear probability model. Treatment effects: See Table 3.

Figure A.1 Duration until the conclusion of the first integration agreement for individuals who were assigned until March 2018



Notes: *) For a period of six months since assignment. Survival functions from Kaplan-Meier estimation. Log-rank test for equality of survival functions: Pr>chi2 = 0.00. 2,659 observations.

CRESTCENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

The Center for Research in Economics and Statistics (CREST) is a leading French scientific institution for advanced research on quantitative methods applied to the social sciences.

CREST is a joint interdisciplinary unit of research and faculty members of CNRS, ENSAE Paris, ENSAI and the Economics Department of Ecole Polytechnique. Its activities are located physically in the ENSAE Paris building on the Palaiseau campus of Institut Polytechnique de Paris and secondarily on the Ker-Lann campus of ENSAI Rennes.