SHOULD WELFARE ADMINISTRATION BE CENTRALIZED OR DECENTRALIZED? EVIDENCE FROM A POLICY EXPERIMENT

Bernhard Boockmann*,§,†
Stephan L. Thomsen¶,$
Thomas Walter$
Christian Göbel$
Martin Huber#

This version: June 13, 2011

Abstract

The 2005 reform of the German welfare system introduced two competing organizational models for welfare administration. In most districts, a centralized organization was established where local welfare agencies are bound to central directives. At the same time, 69 districts were allowed to opt for a decentralized organization. We evaluate the relative success of both types in terms of integrating welfare recipients into employment. Compared to centralized organization, decentralized organization has a negative effect on employment chances of males. For women, no significant effect is found. These findings are robust to aspects of internal organization common to both types of agencies.

JEL-Codes: I38, J64, C31

Key Words: Welfare System, Organization, Decentralization, Labor Market Integration

Acknowledgements: This study is based on results from the evaluation mandated by the "experimentation clause" of Chapter 6 of Book II of the German Social Code (*Sozialgesetzbuch Zweites Buch, SGB II*) and was commissioned by the German Parliament. All opinions expressed are our own. We wish to thank the Institute for Employment Research (IAB), Nürnberg, for providing the administrative data for this study. We gratefully acknowledge the helpful comments of Martin Brussig, Bernd Fitzenberger, Gerd Heyer, Matthias Knuth, Michael Lechner, and Conny Wunsch. We wish to thank Markus Clauss and Martina Hartig for helping us to obtain the database. Moritz Hennig, Stefan Langer, Verena Niepel and Hans Verbeek provided excellent research assistance. Stephan L. Thomsen thanks the *Stifterverband für die deutsche Wissenschaft* for financial support.

* Corresponding author. Institute for Applied Economic Research (IAW), Ob dem Himmelreich 1, D-72074 Tübingen, Tel. (Fax): +49-7071-989620 (-989699), Email: bernhard.boockmann@iaw.edu
§ University of Tübingen; † IZA, Bonn; ¶ Otto-von-Guericke University, Magdeburg; $ Centre for European Economic Research (ZEW), Mannheim; # University of St. Gallen

1 Introduction

Studies from economics, management, and organization theory suggest that the form of organization of an institution, particularly the centralization or decentralization of responsibilities, may have far-reaching implications for their outcomes (see, for example, Besley and Coate, 2003; Richardson et al., 2002; Hutchcroft, 2001). Also, in the case of welfare administration, different organizational systems are likely to result in different incentives and strategies and can influence the success of integrating unemployed welfare recipients into employment. Given that public welfare spending accounts for a significant portion of total government expenditure and given that labor market integration of welfare recipients is the principal task of the public welfare administration for the unemployed, the improvement of organizational effectiveness is a question of foremost economic importance.

One key component in the organization of welfare administration is the degree of local autonomy. In a decentralized setting, local authorities are responsible for the activation of welfare recipients and act independently from central directives and guidelines. Conversely, in a centralized structure, welfare administration is organized by a countrywide government agency that issues directives on how the activation of welfare recipients should be implemented at the local level. Theoretical arguments in favor of a decentralized organization are based on the idea that local authorities are better informed about the characteristics of the local labor market. They are assumed to have detailed knowledge about the specific regional attributes relevant for a successful activation process, and, therefore, they are effective in providing services that are tailored to local conditions. Centralized organizations, on the other hand, are often considered to have an advantage in bundling resources, collecting information from various sources, and imposing best-practice strategies for its local offices (see e.g. Finn, 2000).

The degree of local autonomy of welfare administration varies considerably across countries. In the Netherlands, local authorities form the basis of the public welfare system. In the UK, by contrast, public welfare administration is part of the central government structure. In other countries, welfare reform has changed the degree of centralization of welfare administration. The 1996 U.S. welfare reform, for instance, devolved greater program authority from the federal level to the states, and the Canadian reform that same year gave greater discretion to the provinces (Blank, 2002).

Even though there is an increasing evaluation literature concerning the effectiveness of active labor market programs (ALMP) and certain elements of welfare reform (most of them from the United States, Germany, or other European countries), evidence of the effects of the welfare system organization is scarce.[1] One reason for this is that centralization or decentralization applies to countries as a whole, which makes it difficult to disentangle the effects of a particular organizational setting from other aspects of the welfare system or its reform. So far, conclusions are derived from case studies

[1] For a review of U.S. welfare reforms and the related empirical literature, we refer to Blank (2002), Moffitt (2002), and Grogger and Karoly (2005). Bloom and Michalopoulos (2001) synthesize the results of 29 studies investigating the effects of various US welfare-to-work programs. German welfare-to-work programs that were introduced after 2005 have been analyzed by Hohmeyer and Wolff (2007), Wolff and Jozwiak (2007), Bernhard et al. (2008), Boockmann et al. (2009), Aldashev et al. (2010), Huber et al. (2010), and Thomsen and Walter (2010a). Surveys on welfare reforms in Europe are provided by Torfing (1999), Kildal (2001), and Halvorsen and Jensen (2004) for the Nordic countries, Finn (2000), Beaudry (2002), and Dostal (2008) for the UK, and Finn (2000) and Knijn and van Wel (2001) for the Netherlands. See also Martin and Grubb (2001) and Kluve (2010) for comprehensive overviews.

only (see, for example, Lindsay and McQuaid, 2008; Tergeist and Grubb, 2006). To the best of our knowledge, this is therefore the first study to provide a quantitative assessment of the relative performance of a centralized and a decentralized organization of welfare administration.

We exploit the 2005 reform of the German welfare system that introduced two competing types of organization – a centralized and a decentralized one – in an otherwise homogenous institutional framework. Both approaches were pursued in parallel for a fixed period of time, after which, the more successful model should be determined.[2] In most of the 439 German districts, a centralized organization was established, in which the welfare agencies are subject to the directives and guidelines of the Federal Employment Agency. However, a total of 69 districts were allowed to opt out in favor of a decentralized organization that is legally and organizationally independent from central directives and guidelines. All other components of public welfare and labor market policy – such as benefit entitlements, the tax-benefit system in general, and labor market institutions such as minimum wages and employment protection – apply equally to the centralized and decentralized systems of welfare administration.

Based on a unique data set that is compiled from surveys of welfare administration, Federal Employment Agency (FEA) register data, comprehensive surveys of welfare recipients, and extensive regional information, we evaluate the relative performance of the two organizational systems in terms of successful integration of welfare recipients into the labor market. For our purpose, successful integration means that an unemployed welfare recipient takes up employment without receiving public welfare transfers any longer.[3] To estimate the effect, we apply a propensity score matching estimator that controls for all selective influences between both types of organization. We use specifications with different sets of covariates included in the propensity score. In all cases, the results indicate a very good matching quality; in addition, the estimation results of treatment effects are robust with respect to different specifications.

The estimated effects show that decentralized welfare agencies are less successful than centralized welfare agencies in placing male welfare recipients in employment; for female welfare recipients, the point estimates are also negative, but mostly not statistically significant. Hence, gender differences in the effects can be established. We also estimate the effects for persons living as singles and persons in non-single households separately. Since the German welfare system explicitly targets the household as a whole, only for single households there is coincidence between the relevant unit for welfare and employment. The results tend to be more pronounced for singles. Finally, given the relative success of centralized agencies, we investigate whether the success hinges on centralization itself or is due to internal organization features by exploiting data on the organizational strategies applied in the welfare agencies. We find that the significant negative effect of decentralized welfare agencies on employment for men is largely robust to the inclusion of further information on organizational strategies. The better performance of centralized compared to decentralized welfare agencies therefore re-

[2] This setting was introduced in the so-called experimentation clause in Chapter 6 of Book II of the German Social Code (Sozialgesetzbuch Zweites Buch, SGB II). A description of the experimentation clause with details of implementation, context and policy results is provided by Deutscher Bundestag (2008).

[3] This definition does not preclude that the employer receives public employment subsidies for hiring. Since German hiring subsidies were found to give rise to huge deadweight effects (Boockmann et al., forthcoming), we feel justified to neglect the distinction between subsidized and non-subsidized hiring.

lates to inherent differences between the two types of organization, and not to the adoption of particular forms of internal organization.

2 The German Welfare Reform of 2005

Before 2005, the same organization of welfare administration applied to all 439 districts (in German, *Kreise* and *kreisfreie Städte*) in Germany. There were two different types of welfare benefits: Unemployment and social assistance, which were administered by two different authorities. The centrally organized FEA, represented by the local employment offices, was in charge of unemployment assistance, a means-tested benefit for long-term unemployed individuals whose claims to unemployment insurance benefits had expired. In contrast, local authorities were responsible for social assistance, a benefit for individuals who were not eligible for unemployment assistance or unemployment insurance benefits. This organization of the welfare system, with its two distinct administrative bodies, was often judged as overly fragmented (Tergeist and Grubb, 2006; Eichhorst et al., 2010) and resulted in disincentives with respect to integration into the labor market.

In January 2005, the welfare system reform merged unemployment and social assistance into a single benefit, the so-called Unemployment Benefit II (UBII), to remove the shortcomings mentioned above. In contrast to unemployment assistance, and similar to the former social assistance, UBII is not conditional on former earnings. To be eligible for UBII, persons must be aged between 15 and 64 and must be able to work for at least 15 hours per week. Means-testing takes into account the wealth and income of all individuals living in the household. Individuals who are employed but have insufficient household income are also eligible for the benefit. Recipients of UBII are obliged to actively look for work and to participate in the welfare-to-work programs that are assigned to them. An important part of the reform was the reorganization of the welfare agencies: After the reform, for each district, all welfare services (benefit payments, counseling, labor market activation, etc.) were provided by one welfare agency, as opposed to the previous division of tasks and responsibilities between two administrative entities.

However, there was no political consensus on where the new welfare agencies should be established: Within the system of the centralized FEA or decentralized, at the level of local authorities. Ultimately, the legislator mandated a policy experiment and the evaluation of the relative performance of the two competing models. In the majority of the 439 German districts, local employment offices and local authorities formed a joint venture that is subject to the central controlling standards of the FEA (centralized agencies; in German: *Arbeitsgemeinschaft (ARGE)*). Within the joint venture, the FEA is in charge of the administration of benefits, job placement, and the application of the main instruments of ALMP. In particular, guidelines for these aspects and technical standards as computer software of the FEA are binding for centralized districts. Local authorities are in charge of administering payments for housing costs and for additional needs. Moreover, they provide counseling in specific contexts such as lone parent families, home care for elderly or disabled relatives, or alcohol and drug addiction.[4]

[4] A variant of this model arose where the local employment office and local authorities could not agree on forming a joint venture. In 19 out of 439 cases, both institutions continued to work separately in the district. Howev-

Out of the 439 German districts, 69 were allowed to opt for a decentralized organization of welfare administration (decentralized agencies, in German: *zugelassener kommunaler Träger (zkT)*). Under this system, local authorities autonomously operate the entire activation process including counseling, benefits disbursement, job placement, and the allocation of benefit recipients to ALMP. In particular, local welfare agencies are legally and organizationally independent from central directives and guidelines in the decentralized system.

Table 1: Organizational Features of Decentralized and Centralized Welfare Agencies

	Decentralized Agencies	**Centralized Agencies**
Number of Entities	69	370
Legal Form	Part of local administration	Part of FEA, but is a separate legal entity
Organizational Affiliation	Local authorities	Joint venture between local employment office of the FEA and local authorities
Main Source of Financing	Federal government	Federal government
Centralized Standards of FEA	Not binding, although legal restrictions exist	Binding for job placement, provision of ALMP, monitoring of efforts
Software	Specific solutions for each local authority	Standard system of FEA

Remarks: The numbers of decentralized and centralized welfare agencies presented here refer to October 2006 and are based on the 439 German districts at this time.

Table 1 summarizes the key characteristics of centralized and decentralized welfare agencies. Decentralized welfare agencies can adopt their own organization and integration strategies, and are not bound to central directives. On the other side, centralized agencies are subject to central directives and codes of best-practice.[5] In both the centralized and the decentralized systems, the largest share of welfare payment is financed by the federal government; only a small fraction of overall expenditure – identical in all districts – is taken from local tax budgets. This is different from reforms in other countries, where budgets have been shifted to local authorities as part of the decentralization process. We are therefore able to investigate the effect of decentralized organization independently of budgetary matters.

To evaluate the relative performance of both regimes, it is important to understand selection of districts into the two types. The number of decentralized districts (69) is equal to the number of deputies in the *Bundesrat*, the second chamber of the German Parliament. Each federal state could have between three and six decentralized districts, corresponding to its number of deputies in the *Bundesrat.*

er, because tasks are shared in a similar way as in the case of the centralized system, we do not differentiate between these two types in the empirical analysis.

[5] As will be shown in section 6, the organizational independence of decentralized agencies leads to considerable variance in the implemented integration strategies.

Within each state, districts could apply to opt out of the centralized system. In cases of excess demand, the state government made a selection from the applying districts. In several federal states, the maximum number of districts that could opt for decentralized organization was not exhausted. The vacant places could then be filled by the districts not selected from other states in the first round. With respect to the regional distribution of applications, it appears that the selection process was strongly influenced by political affiliations of the state governments. In two states, Lower Saxony and Hesse, where the conservative governments were strongly in favor of the decentralized system, 13 districts were allowed to opt out, even though these states only had 6 and 5 seats in the *Bundesrat*, respectively. In contrast, hardly any districts were proposed from Mecklenburg-Western Pomerania or Rhineland-Palatinate, both of which were run at that time by social democrats. Hence, the rules for selection resulted in a concentration of decentralized agencies in certain states (WZB et al., 2008).

3 Description of the Data

3.1 The Estimation Sample

In order to investigate whether centralized or decentralized welfare agencies are more successful in integrating welfare recipients into employment, we use a unique data set that was collected for this research question.[6] The data are confined to a pre-selection of 154 districts, a subset of all 439 German districts. Of the 154 districts, 51 are decentralized. The remaining 103 districts have a centralized organization. They were selected to obtain regional units with characteristics similar to the 51 decentralized agencies. Although evidence suggests that the adoption of a decentralized system was driven by the political affiliation of the state governments (WZB et al., 2008), some association could remain between local labor market characteristics and the opt-out from centralized welfare administration. Therefore, the distribution of regional characteristics is accounted for in the sampling procedure.

The selection of comparable districts is explained in detail in Arntz et al. (2006).[7] Based on a comprehensive description of the regional labor market situation until 2004 (before the reform took place), the authors chose variables that are relevant to the transition of the long-term unemployed into the labor market. In a second step, the authors used the reduced set of relevant regional variables and applied the distance matching suggested by Zhao (2004) to identify comparable districts in a regional matching procedure. Appendix 1 illustrates the regional location of the districts in our sample. Appendix 2 shows that the matching of regions equalizes the means of the relevant regional variables. The table reveals that equality of means cannot be rejected for the majority of the variables; the only exceptions are some variables that depend on the degree of urbanization of the district such as, for example, the share of commuters. Here, the mean in centralized districts is slightly higher than it is in decentralized districts. We will account for the possible effects of these differences in the estimation below.

[6] Parts of this data set are publicly available as a scientific use file at the Federal Employment Agency. See Oertel et al. (2009) for details on data access.
[7] The study by Arntz et al. (2006) was conducted to prepare the evaluation of the welfare reform.

3.2 Available Information

To obtain data on the organizational structure of the welfare agencies, repeated interviews were conducted with the agencies' management and staff in the 154 sample units. These surveys have been used to build aggregate measures of the type of case management, the activation concept, the placement strategies, and the mix of ALMP. In addition, a wide range of regional variables (e.g., unemployment rates, welfare ratios, GDP, population density, share of foreigners, etc.) were collected on district-level for several months before and after the 2005 reform.

The individual-level data consist of a survey of welfare recipients who were registered at the 154 agencies. Between January and April 2007, 100 to 300 telephone interviews were conducted within each agency; the number of interviews depended on the size of the welfare agency. In total, nearly 20,300 individuals were interviewed who were drawn from the stock of UBII recipients in October 2006. This sampling scheme could impose a difficulty for the estimation of the relative effects of decentralized and centralized welfare agencies since the sample was not drawn in January 2005 (when the reform was introduced) but in 2006, i.e. more than one year after the implementation of the reform.[8] The reason for this delay is that the disruptions caused by the reform created considerable problems for the quality of administrative data during several months after the introduction of the reform. This particularly applied to decentralized welfare agencies, which continued to use their local computer systems. In principle, an interface for data collection was provided by the FEA, allowing these welfare agencies to interact directly with the FEA's mainframe computers. In practice, however, the adoption of the interface was incomplete until the second half of 2006. Centralized agencies, on the other hand, had issues with a newly introduced software system. For these reasons, the quality of the data during the early periods after the reform is insufficient for empirical analysis. Therefore, we rely on data from 2006 and 2007. Because a large share of UBII recipients depend on welfare benefits for an extended period of time, the stock sample covers those individuals for whom the organization of welfare administration matters the most.

The survey data include individual characteristics (gender, age, marital and parental status, education, health and disability status, migration background, etc.), information on members of the household (number and age of household members and respondent's relation to them), and details concerning the labor market status and labor market history (current labor market state, former spells of insured and minor employment, former spells of unemployment, receipt of welfare benefits, participation in activation programs). Moreover, information is available about basic skills (e.g. reading, writing, math, and computer skills), further qualifications (e.g. driver's license), job search activities, and the concessions that respondents would be willing to make in order to obtain a new job. The survey data were linked with administrative data from the FEA at the individual level. The administrative data include daily information about periods of employment and unemployment, job seeking, participation in ALMP, and benefit receipt. This information allows for the construction of comprehensive

[8] The composition of welfare recipients in the districts could, to some extent, itself be an outcome of decentralized or centralized organization at this point of time. If, for example, the centralized system were faster in integrating welfare recipients with good employment prospects in the early periods after the reform, the stock of welfare recipients in 2006 may contain fewer welfare recipients with favorable characteristics than in decentralized districts. Potential compositional differences of welfare recipients are considered in the estimation, by taking selection at the individual level into account.

labor market histories of the sampled individuals. An overview of the available information is given in Table 2.

Table 2: Overview on Characteristics Included in the Analysis

Basic socio-demographic information	age (18 to 24, 25 to 34, 35 to 44, and 45 to 57 years), schooling (secondary general school, intermediate secondary school, university entrance diploma, other), migration background, household size (1 person, 2 persons, 3 or more persons), no. of children (no children, 1 child, 2 or more children)
Obstacles to employment	disability, care obligation
Labor market and employment history	status before welfare receipt ((minor) employment), no. of half-months unemployed in 2004, no. of half-months unemployed in 2003, no. of half-months unemployed in 2002, no. of half-months unemployed in 2001, no. of half-months out of labor force from 2001 to 2004, mean duration out of labor force from 2003 to 2004 in half-months, no. of programs from 2003 to 2004, mean duration of programs from 2003 to 2004 in half-months
Current welfare spell	months in welfare before 10/2006, start after 10/2006 or missing
Regional information	unemployment ratio (binary), urban district, GDP per employed person (binary), population density (binary), labor market conditions (above average, on average, below average), East Germany
Further socio-demographic variables	at least one child aged below 3 in the household, lone parent status, vocational qualification (none, in-firm training, off-the-job training, university degree, other), self-assessment of overall state of health (good, satisfactory, poor), impairments to health (gastro-intestinal diseases, cardiovascular diseases, rheumatism and other articular trouble, sleep disorders, nervous disorders, allergies, back complaint, other complaints, no health problems), self-assessment of daily working capacity (less than 3 hours, 3 to 6 hours, 6 to 8 hours, 8 or more hours), Self-assessment of basic skills (reading and writing in mother tongue, mathematics, emails and internet), driver's license
Further information on the labor market history from 2001 to 2004	no. of half-months employed in 2004, no. of half-months employed in 2003, no. of half-months employed in 2002, no. of half-months employed in 2001, no. of half-months seeking for a job while employed in 2004, no. of half-months seeking for a job while employed in 2003, no. of half-months seeking for a job while employed in 2002, no. of half-months seeking for a job while employed in 2001, no. of half-months in a program in 2004, no. of half-months in a program in 2003, no. of half-months in a program in 2002, no. of half-months in a program in 2001, no. of employment spells in 2004, no. of employment spells in 2003, no. of employment spells in 2002, no. of employment spells in 2001, no. of unemployment spells in 2004, no. of unemployment spells in 2003, no. of unemployment spells in 2002, no. of unemployment spells in 2001, no. of spells of job seeking while employed in 2004, no. of spells of job seeking while employed in 2003, no. of spells of job seeking while employed in 2002, no. of spells of job seeking while employed in 2001, no. of programs in 2002, no. of programs in 2001, no. of spells out of labor force in 2004, no. of spells out of labor force in 2003, no. of spells out of labor force in 2002, no. of spells out of labor force in 2001

The information used for the outcome variable is also provided by the FEA and indicates for each month between January and December 2007 the employment status of individuals. We define employment without welfare receipt as the outcome of interest. In this case, gross labor earnings (plus any income from other sources such as capital earnings) exceed the income threshold which limits eligibility for welfare benefits.[9] Because our analysis focuses on integration into employment, we restrict the sample to individuals who were unemployed at the time they entered the welfare system and

[9] The administrative data only contain information regarding employment that is subject to social insurance contributions. Therefore, our outcome variable does not include spells of minor employment or self-employment. The outcome variable is measured as a binary dummy variable.

at the time of sampling. Furthermore, we restrict the data to persons aged between 18 and 57 years. Persons aged 58 or older are no longer required to actively search for employment but may remain on welfare benefits until they reach the official retirement age of 65. Individuals aged 15 to 17 years are subject to compulsory schooling and cannot be expected to take up employment. Due to these restrictions, we have 13,286 observations in the estimation sample (4,489 persons from districts with decentralized welfare organization and 8,797 from districts with centralized organization).

4 Estimation Approach

4.1 Estimation of the Average Treatment Effect on the Treated

To evaluate the relative performance of decentralized versus centralized organization on the individual level, the organization of the local welfare administration is the treatment variable of interest. Accordingly, we define two possible treatment states for individual *i*, with $D = 1$ for being registered at a decentralized welfare agency and $D = 0$ for being registered at a centralized welfare agency. The corresponding potential outcomes are denoted Y^1_i and Y^0_i. The individual treatment effect is the difference between the two potential outcomes. Since the individual cannot be in both states at the same time, we estimate the average effect of treatment on the treated (ATT), defined as

$$ATT = E(Y^1 - Y^0|D=1) = E(Y^1|D=1) - E(Y^0|D=1). \quad (1)$$

To estimate the ATT, we use a propensity score matching estimator (see Heckman et al., 1999 for an overview on the identification and estimation of the ATT with matching estimators). The counterfactual outcome of the treatment group is estimated based on information of persons living in centralized districts who are similar in all relevant characteristics to the treatment group. The method of matching puts strong requirements on the data. In order for the ATT to be identified, the Conditional Independence Assumption (CIA) must be fulfilled (Lechner, 2001), i.e. conditional on the set of relevant observable covariates *X*, the potential outcome Y^0 is independent of the organizational model. Hence, all relevant variables have to be observed. Instead of conditioning on *X* directly, Rosenbaum and Rubin (1983) showed that it is sufficient to match on the propensity score, *p(X)*, a scalar function of the set of characteristics. In addition, it has to be ensured that people who resemble the welfare recipients' characteristics distribution in districts with decentralized organization are available in districts with centralized welfare administration (common support condition). Finally, we need to invoke the Stable Unit Treatment Value Assumption (SUTVA, Rubin, 1986), which rules out cross-effects. This requires the regional labor markets to be sufficiently separated so that the success of one welfare agency with respect to job placement does not come at the cost of another welfare agency. This has been checked by IFO and IAW (2008), and the findings indicate robust evidence that SUTVA holds.

In the estimation, we use a kernel density matching estimator with bootstrapped standard errors and 250 replications.[10] The kernel matching estimator weights the control observations according to their "distance" (in terms of the propensity score) to the treated individuals by means of an Epanechnikov kernel function and a bandwidth of 0.06.[11] Persons residing in the same district are affect-

[10] According to Abadie and Imbens (2008), bootstrapped standard errors are unbiased for kernel matching (due to the smoothness of the objective function).
[11] We use the matching algorithm provided by Leuven and Sianesi (2003).

ed by common shocks. Although this does not affect the consistency of the estimates, it may affect efficiency (see, e.g., Moulton, 1986; 1990). We account for this by estimating clustered standard errors at the agency level (using the non-overlapping block bootstrap), i.e. by re-sampling persons on the agency level, but not on the individual level.

4.2 Plausibility of the CIA: Specification of the Propensity Score

To estimate the causal effect of decentralized welfare administration on individual labor market outcomes, we have to rule out selective participation in treatment. The most common types of selectivity in the evaluation of labor market policies are self-selection and selection by a caseworker into the program. In our case, these types of selectivity are very unlikely. From the point of view of a welfare recipient or the caseworker, the 2005 reform of welfare administration and organization is an exogenous event that cannot be easily influenced or avoided. The only way to select into treatment would be to move to another district. However, welfare recipients usually cannot afford to relocate and are not encouraged to move as long as they remain on welfare. Another possible self-selection concerns the inflow into welfare receipt. UBII recipients have to be able to work for at least 15 hours a week. In determining whether claimants meet this requirement, welfare agencies possess a considerable degree of leeway. If ability criteria differ systematically between centralized and decentralized welfare agencies, this may result in a different composition of welfare recipients with regard to characteristics such as illness or disability. In a similar manner, different distributions of welfare recipients between centralized and decentralized agency districts may result from the sampling scheme (see above).[12]

In order to solve these potential problems, we can include a comprehensive characterization of the individual situation in the propensity score. We have access to socio-demographic characteristics beyond the standard set of controls such as migration background, basic mathematics, literacy and computer skills, self-assessed working capacity (measured in hours per day), and obstacles to employment such as provision of long-term care of relatives. In addition, detailed information on the labor market history of each individual, including frequency and duration of employment, unemployment, job seeking activity, ALMP participation, and benefit receipt between 2001 and 2004 as well as information on recent labor market history is available. Direct measures of individual motivation and attitudes are not included, but it is likely that these characteristics are relatively persistent over time. For this reason, it is crucial that we are able to condition on individual employment histories in a detailed manner.[13] Lechner and Wunsch (2011) analyze the sensitivity of matching-based evaluations with respect to the availability of control variables. Their results indicate that besides variables that are commonly used in evaluation studies, such as pre-treatment outcomes, caseworker assessments, and labor market histories also information on the health status, short- and long-run employment histories, the timing of unemployment and job search behavior is required for the CIA to be fulfilled. Our data include all of these variables; they even exceed the set of information used by Lechner and Wunsch

[12] Different behavior of agencies, with respect to inflow and activation, can lead to selection at the individual level. To avoid composition bias, it is therefore necessary to control for individual characteristics. This applies to all studies where the sampling date is after the treatment, even in case of random treatment.
[13] This is also emphasized by Card and Sullivan (1988) and Heckman et al. (1998).

(2011). Thus, making use of our unusually rich data set, we are confident that we capture all relevant factors that affect both participation in treatment and our outcome variable of interest.

Since we have a large number of potential control variables at our disposal for the specification of the propensity score, and because including irrelevant covariates may introduce noise into the calculation of the propensity score, we choose different specifications in order to check the robustness of the estimated treatment effects. The first specification contains the most important individual characteristics (age, schooling, migration background, household size, number of children, obstacles to employment, and several indicators for labor market history) as well as information on the duration of the current welfare spell and limited regional information.[14] Based on the results of balancing tests, this parsimonious specification is our preferred specification. To these variables, we add further regional information in the second specification. The third specification contains the full set of covariates. Descriptive statistics for all variables included in the different propensity score specifications tested are provided in Appendix 3.

Since many evaluation studies have found that the effectiveness of labor market activation differs between genders (see e.g. Bergemann and van den Berg, 2008, for a survey on recent findings for Europe), all estimations are done separately for men and women. In addition, we split the sample according to household size and differentiate between individuals living on their own (singles) and individuals who live with at least one other person in the same household (non-singles). Activation by welfare agencies targets the household as a whole. Only for single households, this is the unit which may also be integrated in employment. Therefore, we look at single and multi-person households separately. The estimation results for the propensity score models in the preferred specification (baseline specification) are given in Appendix 4.[15]

4.3 Balancing Quality of the Matching Estimator

To assess the quality of matching, we apply four balancing tests. First, we compare the means of the variables included in the propensity score between treatment and the control group and test for differences in means, by applying *t*-tests. After successful matching, there should be no remaining differences in the distribution of the covariates. Second, we follow Rosenbaum and Rubin (1985), who suggest the use of the so-called standardized difference in percent. Differences in means of single covariates between the treatment and control group are compared before and after matching, standardized by the mean standard deviation across groups before matching. A third test relates to the explanatory power of the propensity score model after matching. Re-estimation of the same probit regression on the matched sample should result in an explained treatment variation of almost zero, as measured by the McFadden-R^2 (Sianesi, 2004). Fourth, we apply the quality indicator as suggested by Smith and Todd (2005):

[14] The duration of the welfare spell is measured as the number of months on welfare benefits before the sampling date. Due to the time span between sampling and interview date, not all individuals report a starting date of welfare receipt before the sampling date. Some left and re-entered the welfare system during fall and winter 2006/2007 and thus report a starting date after the sampling date. For these individuals, the duration variable is set to 0. An additional dummy variable takes these late starting dates into account.

[15] A complete set of estimation results of the propensity score models can be provided by the authors on request.

$$Z_k = \beta_0 + \beta_1\hat{P}(Z) + \beta_2\hat{P}(Z)^2 + \beta_3\hat{P}(Z)^3 + \beta_4\hat{P}(Z)^4 + \beta_5 D + \beta_6 D\hat{P}(Z) + \beta_7 D\hat{P}(Z)^2 + \beta_8 D\hat{P}(Z)^3 + \beta_9 D\hat{P}(Z)^4 + \eta \quad (2)$$

Each variable included in the propensity score is regressed on a higher-order polynomial of the propensity score, the treatment indicator, and the interaction between both. In the ideal case, coefficients β_5 to β_9 should be jointly zero, indicating that there is no further observable selection into the treatment conditional on the propensity score.

Table 3: Indicators for Matching Quality

	Men	**Women**
Before Matching		
McFadden-R^2	0.012	0.009
LR-Test	92.730	77.630
p-value	0.000	0.000
Mean standardized difference in percent	4.686	4.977
After Matching		
McFadden-R^2	0.000	0.001
LR-Test	2.740	9.050
p-value	1.000	0.999
Mean standardized difference in percent	0.829	1.440
Smith and Todd (2005) balancing test		
p-values > 0.05	25	19
p-values > 0.01	25	24

Remarks: McFadden-R^2 derives from a probit estimation of the propensity score on all covariates considered. The LR-statistic and the corresponding p-value derive from a likelihood-ratio test of the joint insignificance of all covariates. The mean standardized difference in percent has been calculated as an unweighted average of all covariates. The Smith-Todd test displays the number of covariates passing the test at the indicated significance level. There are 26 covariates included in the preferred specification.

As can be seen from the results of the balancing tests depicted in Table 3, matching quality is very satisfactory.[16] According to Table 3, the mean standardized difference in percent is strongly reduced after matching. The McFadden-R^2 estimates of the third test are almost zero after matching; thus, as intended, re-estimation of the propensity score specification on the matched sample does not result in any explanatory power of the included covariates. Almost all of the variables included in the propensity score model pass the test suggested by Smith and Todd (2005). In addition, Appendix 3 shows that the equality of means of the variables included in the propensity score specification between treatment and control group cannot be rejected in just about any case.

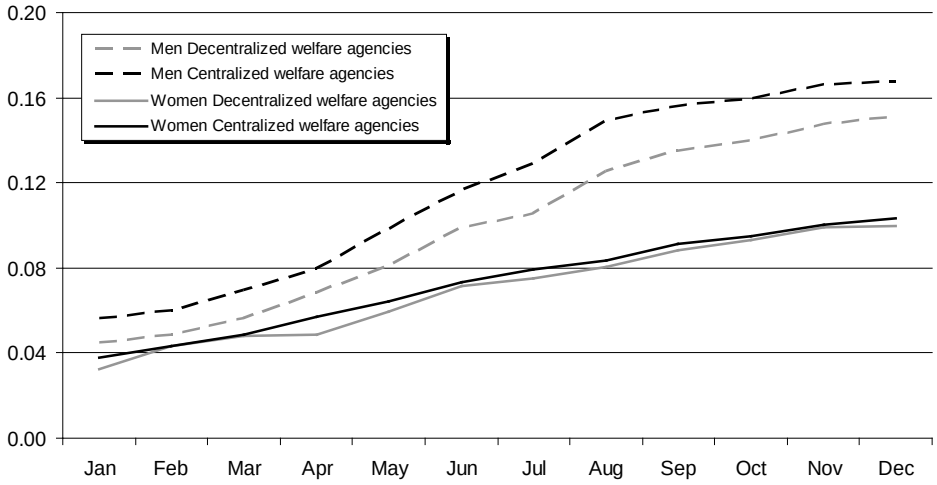
5 Empirical results

Before presenting the estimation results, we briefly describe the development of our outcome variable, employment without welfare receipt, where we distinguish between individuals who are registered at

[16] The corresponding results for the samples of singles and non-singles are given in Appendix 5 and Appendix 6.

centralized and decentralized welfare agencies (see Figure 1). For men, employment rates in districts with centralized welfare agencies are larger than they are in districts with decentralized organization. By December 2007, we observe a mean difference of about one and a half percentage points between decentralized and centralized welfare agencies (16.8% for centralized and 15.2% for decentralized welfare agencies). There is no difference between the two organizational models for women.

Figure 1: Means of the Outcome Variable "Employment *Without* Welfare Receipt"

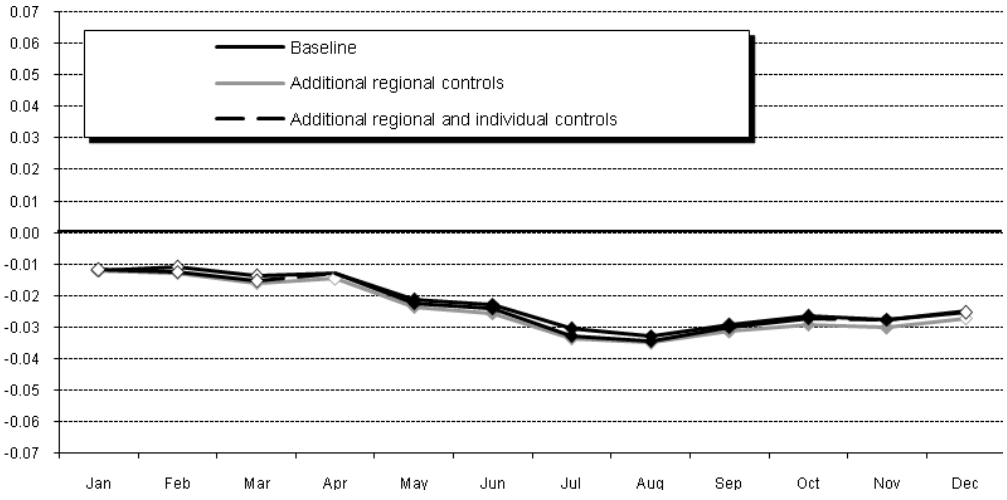


Note: Displayed results for 2007; sampling date: October 2006. All sampled persons are receiving welfare benefits at sampling date.

Our econometric analysis is consistent with these descriptive findings. As discussed in Section 4, we use three different specifications of the propensity score (with baseline denoting the preferred specification). The estimated treatment effects of decentralized welfare agencies are presented in Figures 2 and 3 for both men and women. Rather than showing treatment effects at a single observation date, we display their evolution over the course of 2007, the year after sampling.

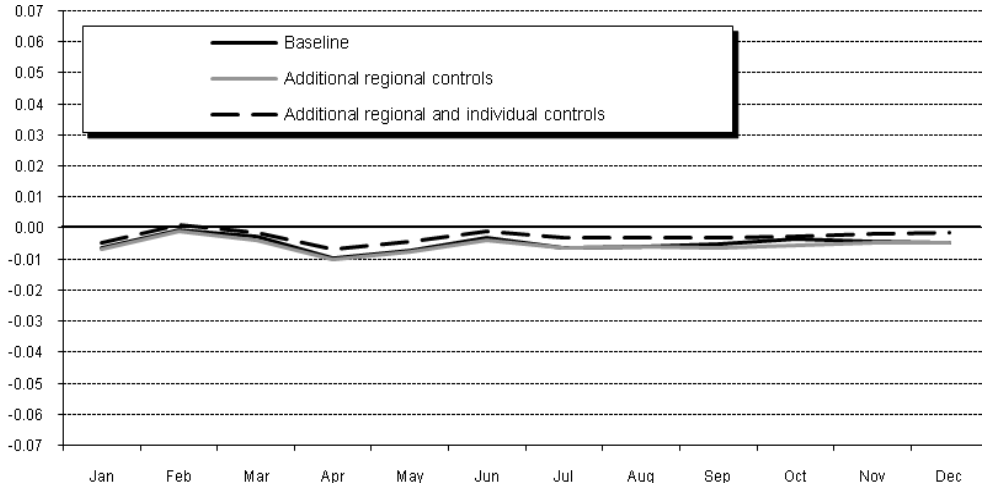
For men, we observe a negative treatment effect, i.e. decentralized welfare agencies are less successful than centralized agencies in placing welfare recipients in jobs that provide a sufficient living income. The absolute effect rises from one to over three percentage points from January to August 2007, and declines moderately thereafter. These magnitudes are slightly larger than the descriptive evidence presented in Figure1. The effects for May to November are significant at the 5% level, with *t*-statistics ranging from 1.96 to 2.91. With the exception of April, the effects for the other months are significant at the 10% level. The inclusion of further covariates leaves the estimated effects virtually unaffected (see Figure 2).

Figure 2: Estimated Treatment Effects on Employment, Men



Notes: ♦ indicates significance at the 5% level, ◊ significance at the 10% level; displayed results for 2007; sampling date: October 2006.

Figure 3: Estimated Treatment Effects on Employment, Women



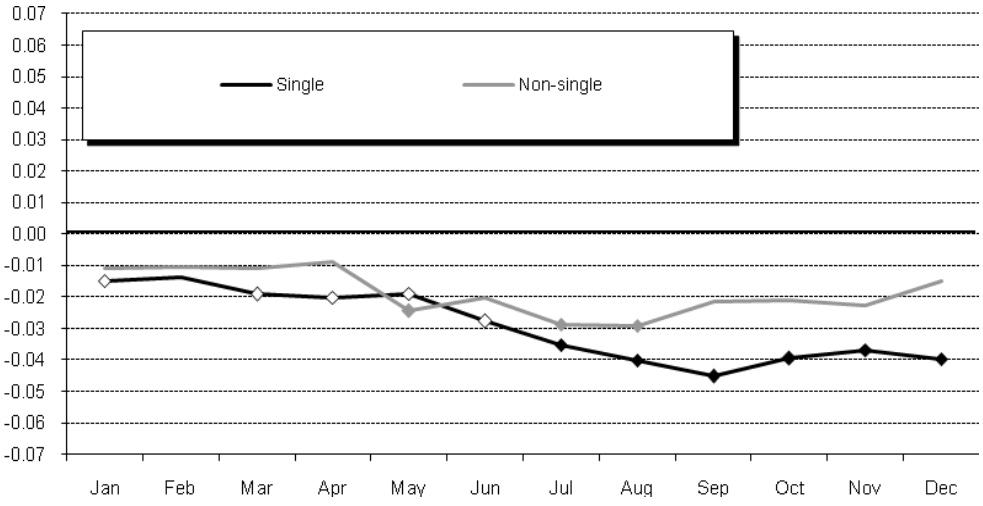
Notes: ♦ indicates significance at the 5% level, ◊ significance at the 10% level; displayed results for 2007; sampling date: October 2006.

Given the relatively small fraction of people taking up employment (Figure 1), the effects for men are substantial. The largest estimated effect of nearly -3.5 percentage points, estimated for August 2007, implies that decentralized agencies have an about 24% lower integration quota than centralized agencies. For women, we also find negative treatment effects, which are however smaller in magnitude and not statistically significant (see Figure 3). Again, the results are insensitive to the specification of the propensity score.

Gender differences are also present when we split the sample into single and non-single households (see Figures 4 and 5 providing estimated treatment effects based on the baseline specifica-

tion of the propensity score models). For single men, we estimate a substantially negative employment effect of decentralized welfare agencies. The negative effect amounts up to 4.5 percentage points in absolute terms. For single women, we observe a negative treatment effect, too. This effect, however, is only slightly significant at the beginning of our observation period. Thereafter, it is insignificant and of smaller magnitude than the effect found for single men. In case of non-single men, we estimate a negative treatment effect of decentralized welfare agencies which has an absolute value of up to 2.9 percentage points. This effect is of smaller magnitude than the effect found for single men, but it is larger than the effect for non-single women. For the latter subgroup, we cannot establish a significant treatment effect.

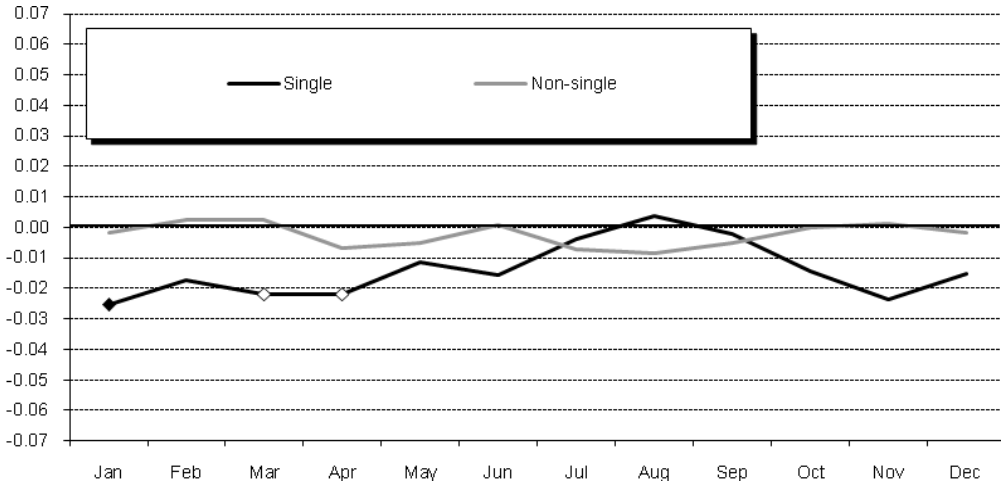
Figure 4: Treatment Effects on Employment, Singles and Non-singles, Men



Notes: ♦ indicates significance at the 5% level, ◊ significance at the 10% level; displayed results for 2007; sampling date: October 2006.

To discuss potential reasons for the gender differences in our results, we refer to the study of IAQ et al. (2009). This study combines analyses of survey and administrative data with case studies within welfare agencies. It shows that women are less intensively activated than men, irrespectively of the agency type. In particular, women are less frequently assigned to ALMP programs than men (see also Thomsen and Walter, 2010b, and Boockmann et al., 2011).

Figure 5: Treatment Effects on Employment, Singles and Non-singles, Women



Notes: ♦ indicates significance at the 5% level, ◊ significance at the 10% level; displayed results for 2007; sampling date: October 2006.

In addition, according to Chapter 10 of Book II of the German Social Code, parents of small children under the age of three years may not be activated at all. According to the results of IAQ et al. (2009), many more mothers than fathers make use of the option to withdraw from active job search. Case study evidence also suggests that activation efforts of welfare agencies further differ between genders for efficiency reasons (IAQ et al., 2009). Due to limited time resources of the caseworkers and the overall goal to realize as many transitions to employment as possible, activation is mainly targeted to the most easy-to-place individuals. In most cases, welfare agencies assume that men are the easy-to-place individuals. If women are much less intensively activated than men or even not activated at all, we would not expect any significant difference in the success of decentralized and centralized welfare agencies to integrate female welfare recipients into employment without welfare receipt. Differences can only be present for individuals who are subject to activation like men. Thus, the findings of IAQ et al. (2009) and other studies might explain why we observe gender differences in our results.

6 A First Glance at the Black Box of Welfare Administration

The significant treatment effect for men raises the question of why centralized organization performs better in placing welfare recipients into jobs. Is the relative success of centralized agencies due to their use of more successful organizational approaches and strategies that could also be adopted by decentralized agencies as well? All centralized welfare agencies are subject to central FEA guidelines, central controlling, and certain directives regarding the use of activation strategies. Nevertheless, welfare agencies have leeway in the way they internally organize their services for welfare recipients. The implementation of organizational approaches is not specific to either administrative model, and we observe variations within both agencies with different organizational features. In the following we

analyze the effect of the adopted approaches and strategies and check if they are able to explain the positive effect of centralized organization.

In order to do so, we exploit data on the organizational strategies applied in the welfare agencies. According to studies conducted to evaluate the implementation of Germany's 2005 welfare reform (IAW and ZEW, 2008; WZB et al., 2008), the following features are the most important elements in the internal organization of tasks and the cooperation with external partners:[17]

1) *Generalized case management* for all clients as opposed to case management by specialized staff for clients with multiple obstacles to employment,
2) *Integration of activation and placement* as opposed to the separation of these functions,
3) Use of *customer segmentation* procedures,
4) Establishment of an *employer service*, i.e. specialized staff maintaining contact with employers,
5) *Subcontracting of placement services* to private providers.

Table 4 provides a more detailed description of the organizational features, and outlines some arguments as to why they could affect the integration success of welfare recipients. Customer segmentation and particularly generalized case management tend to be used much more frequently by decentralized agencies, integration of activation and placement is slightly more common among centralized agencies, while the other two strategies are not related to agency type.

To check whether the effect of decentralized agencies can be attributed to one of these strategies, we require a multivariate framework. For this purpose, we use binary probit models. The probit estimations contain all covariates used in the preferred specification of the propensity score (see above). In addition, dummy variables for decentralized welfare agencies and for the organizational features are included. We then test whether a significant effect of decentralized agencies on employment without welfare receipt remains despite controlling for organization.

[17] The effects of further characteristics and strategies of the welfare agencies are considered in ZEW et al. (2008).

Table 4: Definition of Organizational Variables

Definition	Possible Impact on Integration	Frequency in Sample
	Generalized Case Management	
Case managers counsel all types of clients. There is no assignment of welfare recipients with multiple obstacles to employment to specialist caseworkers.	Better placement under specialized case management if clients with specific problems require specialized expertise. Generalized case management facilitates individual counseling as clients have fewer contact persons.	0.69 (decentralized agencies) 0.24 (centralized agencies)
	Integration of Activation and Placement	
Clients are counseled (activated) and placed into employment by the same staff members. There is no assignment of specialized staff to the two tasks.	Integration reduces the number of contact persons for each welfare recipient, and facilitates a holistic approach. In contrast, separation leads to gains from specialization but may create coordination problems at the interface of both tasks.	0.51 (decentralized) 0.59 (centralized)
	Customer Segmentation	
Classification of clients into different groups receiving different treatment during activation	Segmentation may increase employment rates among groups that are activated more intensely but reduces integration into employment in other groups.	0.84 (decentralized) 0.66 (centralized)
	Employer Service	
A team of agency staff members maintains a network with employers and serves as contact persons for them	Networking may result in better placement. However, internal coordination problems between the employer service and caseworkers may arise.	0.86 (decentralized) 0.83 (centralized)
	Subcontracting of Placement Services	
The welfare agency uses private employment services to place some of their clients into employment.	Specialization gains may occur. However, private agencies may work more or less effectively compared to the public employment service. Requires proper assignment of welfare recipients to service providers.	0.41 (decentralized) 0.40 (centralized)

Table 5 displays our estimation results for April, August and December, 2007, which account for potential clustering of error terms at agency level (see e.g. Moulton, 1986; 1990). The entries in the table are marginal effects of the dummy variables on the outcome variable, and their magnitudes and treatment effects from matching are, therefore, comparable. Since results did not differ much between randomly chosen individuals and single or non-single households, we rely on the overall samples of men and women.

Similar to the matching results, we find a negative effect of decentralization for men. However, the effect is slightly smaller with a magnitude of up to 2.5 percentage points. For women, the effect of

decentralized agencies is again insignificant. The organizational variables themselves are mostly insignificant, the effect of an employer service in the subsample of women being the only exception. The complete set of results shows that this effect is negative and significant at the 10% level in the last four months of the observation period. This could be interpreted as indication that the presence of a specific organizational unit of the agency specializing in employer contacts is a disadvantage. It might be that the installation of an employer service shifts away valuable resources from the activation process of welfare recipients, which, in turn, might have negative consequences for employment uptake. Unfortunately, there is no direct evidence with which these results can be validated.

Despite the slight decline in magnitude, the significant negative effect of decentralized welfare agencies on employment for men is largely robust to the inclusion of further organizational strategies. Therefore, we conclude that the effect of organization of welfare agencies is not due to the adoption of particular forms of internal organization. A more likely explanation of the difference in effects relates to the theoretical argumentation. The advantages of centralized organization in bundling resources, collecting information from various sources, and imposing best-practice strategies for the local offices tend to outperform the favorable properties of decentralized organization.

Table 5: Estimated Probit Effects of Organizational Features

	Men			Women		
	April	August	December	April	August	December
Decentralized welfare agency	-0.0111	-0.0259**	-0.0156	-0.0053	-0.0032	-0.0039
	(0.0070)	(0.0119)	(0.0145)	(0.0058)	(0.0074)	(0.0105)
Generalized case management	0.0037	0.0014	-0.0094	-0.0065	0.0000	0.0029
	(0.0074)	(0.0119)	(0.0141)	(0.0061)	(0.0073)	(0.0103)
Integration of activation and placement	-0.0011	-0.0057	-0.0035	0.0035	0.0088	-0.0010
	(0.0068)	(0.0100)	(0.0126)	(0.0055)	(0.0068)	(0.0086)
Customer segmentation	0.0015	-0.0087	0.0066	0.0005	-0.0040	0.0021
	(0.0066)	(0.0098)	(0.0117)	(0.0060)	(0.0074)	(0.0088)
Employer service	-0.0085	0.0035	-0.0185	-0.0131	-0.0129	-0.0245*
	(0.0079)	(0.0126)	(0.0173)	(0.0087)	(0.0109)	(0.0138)
Subcontracting of placement services	-0.0009	0.0010	-0.0004	-0.0055	-0.0027	-0.0077
	(0.0061)	(0.0102)	(0.0121)	(0.0060)	(0.0072)	(0.0090)
Pseudo R^2	0.069	0.067	0.069	0.081	0.078	0.072
Log-Likelihood	-1559.59	-2368.69	-2571.75	-1358.90	-1838.24	-2141.99

Remarks: Displayed are marginal effects and standard errors in brackets. *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable in each model and for each month is defined to be 1 if an individual is employed and does not receive welfare benefits anymore. The variable is 0 otherwise. The number of observations in the sample of men (women) is 6,217 (6,992). One centralized welfare agency had to be dropped from the analysis due to missing information on the organizational variables. The standard errors take into account clustering at the agency level. All models include as further regressors the covariates used in the preferred propensity score specification of the matching analysis. The detailed results are not displayed here but are available from the authors upon request. The displayed results refer to the year 2007.

7 Conclusions

The German welfare reform of 2005 introduced two competing organizational systems for the labor market activation of welfare recipients in an otherwise homogenous institutional setting: decentralized and centralized welfare agencies. In order to evaluate their relative performance, we have estimated their effect on the integration of welfare recipients into employment without welfare receipt, regarding regional differences as well as individual selection. The estimation is based on exceptionally rich data from various sources. We have combined a detailed survey of welfare recipients with administrative records from the Federal Employment Agency (FEA). In addition, we have used a large set of variables that describe the local labor market. Finally, we have considered unique information on the internal organization of the welfare agencies in our sample.

We find that decentralized welfare agencies have a negative effect on male welfare recipients with respect to integration into employment. Given the low transition intensity from welfare receipt into employment in general, the magnitudes of the effects for men are substantial. The integration quota of decentralized welfare agencies is up to 24% lower than the quota of centralized agencies. For women, we also find negative treatment effects, which are, however, smaller in magnitude than for men and which are not statistically significant. Gender differences are found within all subgroups considered (randomly chosen individuals, singles, and non-singles). These might result from different activation intensity between men and women. Evidence suggests that, irrespective of agency type, the activation intensity of women is far lower compared to men. If welfare agencies concentrate their activation efforts predominantly on men rather than on women, it is harder to uncover significant differences in the relative performance of decentralized and centralized agencies for the latter subgroup.

We have further explored channels through which our results may have emerged. Because welfare agencies have significant discretionary power with respect to internal organization, we have checked whether the organization of tasks at individual welfare agencies is responsible for the result of decentralization. Although the effects are slightly weakened by the inclusion of the additional organizational strategies, the overall result is not affected. We conclude that the negative effect of decentralization is not due to differences in the adoption of strategies between centralized and decentralized welfare agencies and is not subject to their choices regarding the internal organization of tasks. The remaining differences are related to the very nature of (de)centralized organization. Examples are the application of central best practice guidelines of the FEA concerning the use of instruments of activation, as well as the (de)centralized controlling system.

This is the first paper that provides quantitative evidence on the effects of (de)centralization of public welfare on employment transitions. Our findings point to the importance of the organizational aspects of welfare administration to the integration of welfare recipients into employment. Despite their importance, this topic has been largely neglected by existing literature on employment policy. Identifying successful and less successful strategies for the organization of welfare administration remains a difficult yet highly relevant task.

References

Abadie, A. and Imbens, G. (2008). On the Failure of the Bootstrap for Matching Estimators, *Econometrica, 76*(6), p. 1537-1557.

Aldashev, A., Thomsen, S. L. and Walter T. (2010). *Short-term Training Programs for Immigrants: Do Effects Differ from Natives and Why?*, ZEW Discussion Paper No. 10-021, Mannheim.

Arntz, M., Wilke, R. and Winterhager, H. (2006). *Regionenmatching im Rahmen der Evaluation der Experimentierklausel des § 6c SGB II: Methodische Vorgehensweise und Ergebnisse*, ZEW Discussion Paper No. 06-061, Mannheim.

Beaudry, R. (2002). *Workfare and Welfare: Britain's New Deal*, Working Paper, Department of Economics, University of York.

Bergemann, A. and van den Berg, G. (2008). Active Labor Market Policy Effects for Women in Europe: A Survey, *Annales d'Economie et de Statistique, 91/92*, p. 385-408.

Bernhard, S., Gartner, H. and Stephan, G. (2008). *Wage Subsidies for Needy Job-Seekers and Their Effect on Individual Labor Market Outcomes after the German Reforms*, IZA Discussion Paper No. 3772, Bonn.

Besley, T. and Coate, S. (2003). Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach, *Journal of Public Economics, 87*. p. 2611-2637.

Blank, R. M. (2002). Evaluating Welfare Reform in the United States, *Journal of Economic Literature, 40*(4), p. 1105-1166.

Bloom, D. and Michalopoulos, C. (2001). *How Welfare and Work Policies Affect Employment and Income: A Synthesis of Research*, Manpower Demonstration Research Corporation, New York, NY.

Boockmann, B., Thomsen, S. L. and Walter, T. (2009). *Intensifying the Use of Benefit Sanctions - An Effective Tool to Shorten Welfare Receipt and Speed Up Transitions to Employment?*, ZEW Discussion Paper No. 09-072, Mannheim.

Boockmann, B., Thomsen, S. L. and Walter, T. (2011). Aktivierung der erwerbsfähigen Hilfebedürftigen mit arbeitsmarktpolitischen Maßnahmen - Wer wird gefördert? Empirische Ergebnisse auf Grundlage eines neuen Surveydatensatzes, *AStA Wirtschafts- und Sozialstatistisches Archiv*, 4(4), p. 269-292.

Boockmann, B., Zwick, T., Ammermüller, A. and Maier, M. (forthcoming). Do Hiring Subsidies Reduce Unemployment Among Older Workers? Evidence From Natural Experiments, *Journal of the European Economic Association*.

Card, D. and Sullivan, D. (1988). Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment, *Econometrica, 56*, p. 497-530.

Dostal, J. M. (2008). The Workfare Illusion: Re-examining the Concept and the British Case, *Social Policy and Administration, 42*, p. 19-42.

Deutscher Bundestag (2008). *Bericht zur Evaluation der Experimentierklausel nach § 6c des Zweiten Buches Sozialgesetzbuch*, Forschungsbericht F390, Bundestagsdrucksache, Nr. 16/11488. http://www.bmas.de/portal/31070/f390__forschungsbericht.html.

Eichhorst, W., Grienberger-Zingerle, M. and Konle-Seidl, R. (2010). Activating Labor Markets and Social Policies in Germany: From Status Protection to Basic Income Support, *German Policy Studies, 6*(1), p. 65-106.

Finn, D. (2000). Welfare to Work: the Local Dimension, *Journal of European Social Policy, 10*, p. 43-57.

Grogger, J. and Karoly, L. (2005). *Welfare Reform: Effects of a Decade of Change*, Harvard University Press, Cambridge, MA.

Halvorsen, R. and Jensen, P. H. (2004). Activation in Scandinavian Welfare Policy, *European Societies, 6*, p. 461-483.

Heckman, J. J., Ichimura, H., Smith, J. and Todd, P. (1998). Characterizing Selection Bias Using Experimental Data, *Econometrica, 66* (5), p. 1017-1098.

Heckman, J. J., Lalonde, R. and Smith, J. (1999). The Economics and Econometrics of Active Labor Market Programs, in *Handbook of Labor Economics*, ed. by Ashenfelter, O. and Card, D., Vol. 3A, Elsevier: Amsterdam. p. 1865-2097.

Hohmeyer, K. and Wolff, J. (2007). *A Fistful of Euros. Does One-euro-job Participation Lead Means-tested Benefit Recipients Into Regular Jobs and Out of Unemployment Benefit II Receipt*, IAB Discussion Paper, No. 32/2007, Nürnberg.

Huber, M., Lechner, M., Wunsch, C. and Walter, T. (2010). Do German Welfare-to-Work Programmes Reduce Welfare Dependency and Increase Employment?, *German Economic Review, 12*(2), p. 182-204.

Hutchcroft, P. D. (2001). Centralization and Decentralization in Administration and Politics: Assessing Territorial Dimensions of Authority and Power, *Governance, 14*(1), p. 23-53.

IAQ, FIA and GendA (2009). *Bewertung der SGB II-Umsetzung aus gleichstellungspolitischer Sicht.*, Abschlussbericht, Bundesministerium für Arbeit und Soziales, Berlin.

IAW and ZEW (2008). *Evaluation der Experimentierklausel nach § 6c SGB II. Vergleichende Evaluation des arbeitsmarktpolitischen Erfolgs der Modelle der Aufgabenwahrnehmung "zugelassener kommunaler Träger" und "Arbeitsgemeinschaft". Untersuchungsfeld 1: "Deskriptive Analyse und Matching"*, Abschlussbericht, Bundesministerium für Arbeit und Soziales, Berlin.

IFO and IAW (2008). *Evaluation der Experimentierklausel nach § 6c SGB II - Untersuchungsfeld 4: Makroanalyse und regionale Vergleiche. Vergleichende Evaluation des arbeitsmarktpolitischen Erfolgs der Modelle Aufgabenwahrnehmung "zugelassene kommunale Träger" und "Arbeitsgemeinschaft"*, Abschlussbericht, Bundesministerium für Arbeit und Soziales, Berlin.

Kildal, N. (2001). *Workfare tendencies in Scandinavian welfare policies*, International Labor Office, Geneva.

Kluve, J. (2010). The effectiveness of European active labor market programs, *Labour Economics, 17*(6), p. 904-918.

Knijn, T. and van Wel, F. (2001). Careful or Lenient: Welfare Reform for Lone Mothers in the Netherlands, *Journal of European Social Policy, 11*, p. 235-251.

Lechner, M. (2001). Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption, in: *Econometric Evaluation of Labor Market Policies*, ed. by M. Lechner, and F. Pfeiffer, Vol. 13 of ZEW Economic Studies, 43-58. Physica-Verlag, Heidelberg.

Lechner, M. and Wunsch, C. (2011). *Sensitivity of Matching-Based Program Evaluations to the Availability of Control Variables*, IZA Discussion Paper No. 5553, Bonn.

Leuven, E. and Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*, Statistical Software Components, Boston College Department of Economics.

Lindsay, C. and McQuaid, W. (2008). Inter-agency Co-operation in Activation: Comparing Experiences in Three Vanguard 'Active' Welfare States, *Social Policy and Society, 7*(3), p. 353-365.

Martin, P. and Grubb, D. (2001). What Works and for Whom: A Review of OECD Countries Experiences with Active Labor Market Policies, *Swedish Economic Policy Review, 8*, p. 9-56.

Moffitt, R. (2002). Welfare Programs and Labor Supply, in: Auerbach, A. and Feldstein, M. (eds.). *Handbook of Public Economics, 4*(4), p. 2393-2430.

Moulton, B. (1986). Random Group Effects and the Precision of Regression Estimates, *Journal of Econometrics, 32*(3), p. 385-397.

Moulton, B. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units, *Review of Economics and Statistics, 72*(2), p. 334-338.

Oertel, M., Schneider, A. and Zimmermann, R. (2009). *Kundenbefragung zur Analyse der Organisationsstrukturen in der Grundsicherung nach SGB II. Dokumentation der Scientific-Use-Files*, FDZ Datenreport 4/2009, Nürnberg.

Rosenbaum, P. and Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika, 70*(1), p. 41-50.

Rosenbaum, P. and Rubin, D. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score, *The American Statistician, 39*(1), p. 33-38.

Richardson, H. A., Vandenberg, R. J., Blum, T. C. and Roman, P. M. (2002). Does Decentralization Make a Difference for the Organization? An Examination of the Boundary Conditions Circumscribing Decentralized Decision-making and Organizational Financial Performance, *Journal of Management, 28*(2), p. 217-244.

Rubin, D. (1986). Statistics and Causal Inference: Comment. Which Ifs Have Causal Answers, *Journal of the American Statistical Association, 81*(396), p. 961-962.

Sianesi, B. (2004). An Evaluation of the Active Labor Market Programmes in Sweden, *The Review of Economics and Statistics, 86*(1), p. 133-155.

Smith, J. and Todd, P. (2005). Rejoinder, *Journal of Econometrics, 125*(1/2), p. 365-375.

Tergeist, P. and Grubb, D. (2006). *Activation Strategies and the Performance of Employment Services in Germany, the Netherlands and the United Kingdom*, OECD Social, Employment and Migration Working Papers, No. 42, OECD Publishing.

Thomsen, S. L. and Walter, T. (2010a). Temporary Extra Jobs for Immigrants: Merging Lane to Employment or Dead-end Road in Welfare?, *LABOUR: Review of Labour Economics and Industrial Relations 24*(s1), p. 114-140.

Thomsen, S. L. and Walter, T. (2010b). Der Zugang zu arbeitsmarktpolitischen Maßnahmen in der Grundsicherung für Arbeitsuchende von Personen mit und ohne Migrationshintergrund im Vergleich", in: Knuth, M. (ed.), *Eingliederung in Arbeit ohne Integrationspolitik? Migrant/innen im SGB II-Leistungsbezug und ihre Aktivierung durch die Grundsicherungsstellen*, pp. 160-183. Nomos Verlagsgesellschaft, Baden-Baden.

Torfing, J. (1999). Workfare with Welfare: Recent Reforms of the Danish Welfare State, *Journal of European Social Policy, 9*, p. 5-28.

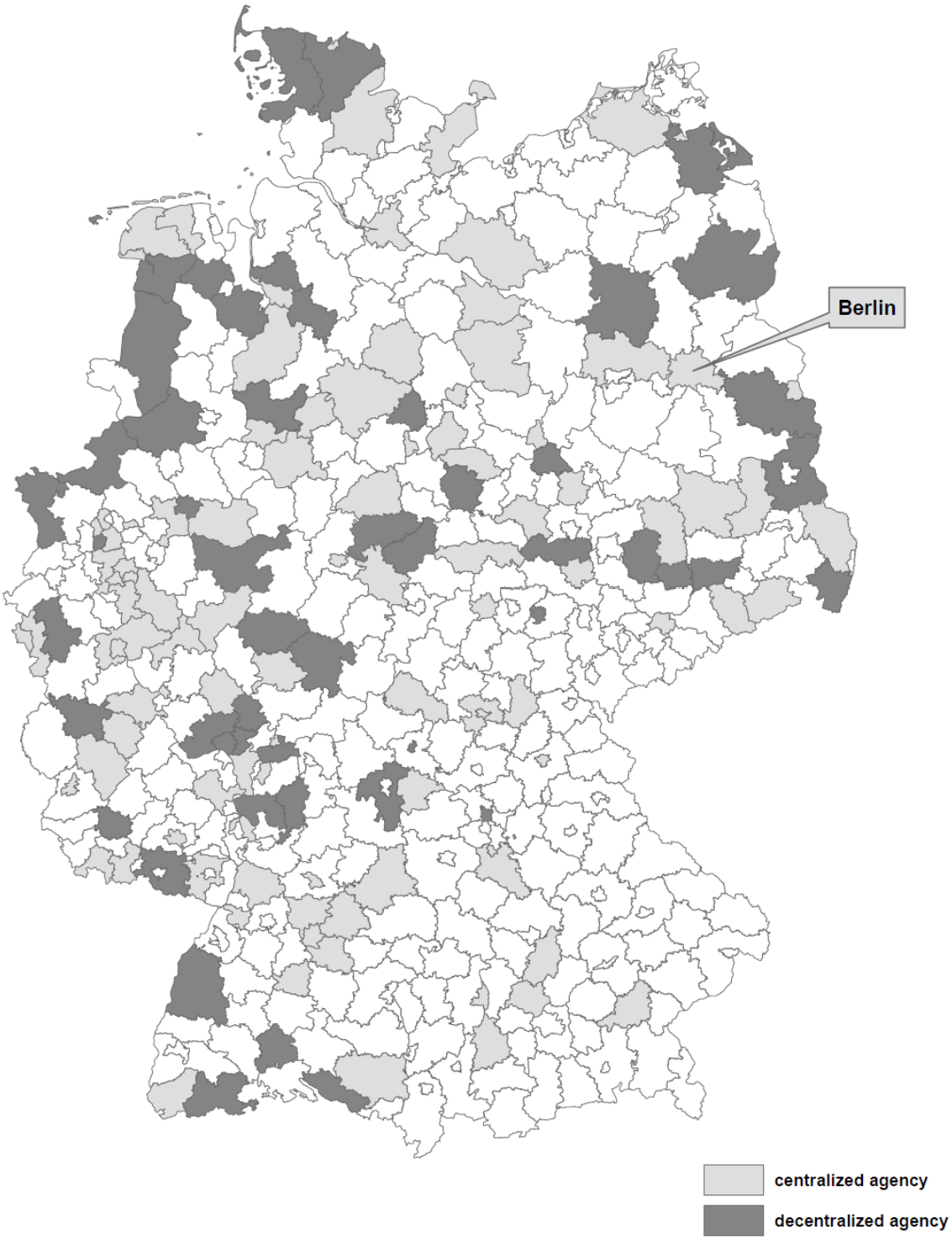
Wolff, J. and Jozwiak, E. (2007). *Does Short-term Training Activate Means-tested Unemployment Benefit Recipients in Germany*, IAB Discussion Paper No. 29/2007, Nürnberg.

WZB, infas and FHS Frankfurt/M. (2008). *Evaluation der Experimentierklausel nach § 6c SGB II – Vergleichende Evaluation des arbeitsmarktpolitischen Erfolgs der Modelle der Aufgabenwahrnehmung "Zugelassener kommunaler Träger" und "Arbeitsgemeinschaft". Untersuchungsfeld 2: "Implementations- und Governanceanalyse"*, Abschlussbericht, Bundesministerium für Arbeit und Soziales, Berlin.

ZEW, IAQ and TNS Emnid (2008). *Evaluation der Experimentierklausel nach §6c SGB II – Vergleichende Evaluation des arbeitsmarktpolitischen Erfolgs der Modelle der Aufgabenwahrnehmung "Optierende Kommune" und "Arbeitsgemeinschaft". Untersuchungsfeld 3: "Wirkungs- und Effizienzanalyse"*, Abschlussbericht, Bundesministerium für Arbeit und Soziales, Berlin.

Zhao, Z. (2004). Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence, *The Review of Economics and Statistics, 86*(1), p. 91-107.

Appendix 1: Map of the 154 Welfare Agencies in the Sample



Appendix 2: Balancing of Regional Variables Among the Sampled Welfare Agencies

	Centralized agencies	Decentralized agencies	p-value
Unemployment rate (Source: FEA)	11.309	11.412	0.906
Unemployment rate of the young (age < 25) (Source: FEA)	10.628	10.505	0.860
Unemployment rate of foreigners (Source: FEA)	23.285	24.340	0.567
Personnel expenditure per unemployed in the stock (classified)	373.435	375.562	0.867
Personnel expenditure per unemployed (inflow)	2442.675	2440.875	0.969
Material expenses per unemployed (classified)	45.188	48.830	0.100
Placement expenditure as a share of total expenditure (classified)	0.024	0.028	0.315
Ratio of caseworkers to unemployed (classified)	0.016	0.016	0.837
Ratio of placement officers with fixed-term contract to unemployed	0.002	0.002	0.895
Ratio of unemployed to applicants	0.804	0.802	0.570
Ratio of male to female unemployed	1.300	1.296	0.892
Ratio of young (< 25) to old (> 50) unemployed (in percent)	49.478	50.966	0.339
Share of unemployed under age 25 (in percent)	12.142	12.211	0.802
Share of unemployed over age 50	0.123	0.116	0.067
Unemployment-Vacancy (UV) relation in textile industry	73.592	84.213	0.301
UV relation in construction sector	37.124	35.640	0.702
UV relation in engineering	16.267	17.857	0.567
UV relation in commerce sector	24.820	27.332	0.462
UV relation in service sector	20.753	24.232	0.212
UV relation in metal industry	15.261	14.610	0.661
UV relation in healthcare	6.346	6.356	0.983
UV relation in social sector	11.433	11.121	0.728
UV relation overall	30.208	32.386	0.471
Share of employees with fixed-term contract	0.811	0.780	0.788
Share of long-term unemployed	0.332	0.333	0.896
Share of severely disabled unemployed	0.040	0.039	0.809
Ratio of welfare recipients receiving no unemployment benefits to all unemployed	0.163	0.101	0.198
Ratio of welfare recipients receiving no unemployment benefits to all unemployed welfare recipients	0.347	0.374	0.057
Rate of long-term unemployed	0.332	0.333	0.893
Rate of long-term unemployed under age 25	0.072	0.069	0.492
Rate of long-term unemployed over age 50	0.508	0.505	0.769
FF per unemployed	0.007	0.009	0.408
FF per male unemployed	0.008	0.010	0.479
FF per female unemployed	0.006	0.008	0.337
FF per unemployed over age 50	0.004	0.005	0.405
FF per unemployed under age 25	0.014	0.019	0.253
Wage subsidies per unemployed	0.032	0.033	0.753
Wage subsidies per unemployed over age 50	0.062	0.065	0.763
Total transitional allowance per unemployed	0.018	0.017	0.787
Bridging allowance per unemployed over age 50	0.008	0.009	0.638
Bridging allowance per unemployed under age 25	0.008	0.007	0.735
Wage subsidies for long-term unemployed per unemployed	0.002	0.003	0.168
Wage subsidies for long-term unemployed per male unemployed	0.002	0.003	0.149
Wage subsidies for long-term unemployed per female unemployed	0.002	0.003	0.131
ABM+SAM/unemployed+ABM+SAM	0.025	0.029	0.444
ABM+SAM/unemployed+ABM+SAM (men)	0.027	0.031	0.407

ABM+SAM/unemployed+ABM+SAM (women)	0.023	0.026	0.508
ABM/unemployed+ABM	0.017	0.019	0.430
ABM/unemployed+ABM (women)	0.016	0.018	0.488
ABM/unemployed+ABM (men)	0.017	0.020	0.389
FbW/(unemployed+FbW)	0.058	0.060	0.205
FbW/(unemployed+FbW) (men)	0.049	0.052	0.310
FbW/(unemployed+FbW) (women)	0.069	0.071	0.264
FbW/(unemployed+FbW) (age > 50)	0.014	0.015	0.360
FbW/(unemployed+FbW) (age < 25)	0.054	0.055	0.741
TM/(unemployed+TM)	0.022	0.022	0.637
TM/(unemployed+TM) (women)	0.023	0.023	0.763
TM/(unemployed+TM) (men)	0.022	0.021	0.539
TM/(unemployed+TM) (age > 50)	0.010	0.010	0.883
TM/(unemployed+TM) (age < 25)	0.036	0.035	0.828
Share of ESF-assisted unemployed	0.005	0.006	0.720
Share of persons in minor employment	0.164	0.193	0.084
JUMP per unemployed (age <25)	0.121	0.136	0.209
Total unemployment rate (Source: Statistical Office)	12.435	12.520	0.927
Female unemployment rate (Source: Statistical Office)	11.601	11.839	0.808
Male unemployment rate (Source: Statistical Office)	13.161	13.120	0.964
Export turnover in manufacturing per employee	52.876	55.487	0.672
Commuter balance per 1000 employees	-64.233	-172.431	0.034
Migration balance/ gross population	0.001	0.001	0.571
Rate of social assistance recipients	0.036	0.028	0.004
Rate of social assistance recipients (men)	0.033	0.025	0.004
Rate of social assistance recipients (women)	0.039	0.031	0.004
Rate of social assistance recipients (natives)	0.032	0.025	0.006
Rate of social assistance recipients (foreigners)	0.083	0.068	0.035
Total business founding intensity per 10000 employable persons	45.947	43.676	0.268
Business foundations per 10000 inhabitants aged 15 to 64	149.643	146.700	0.517
Population density (inhabitants per square meter)	833.656	339.509	0.001
GDP per economically active person	51.657	51.343	0.826
Employment rate	0.465	0.424	0.075
Share of foreigners in total population	0.084	0.065	0.032
Rate of economically active men	0.357	0.361	0.450
Rate of economically active women	0.284	0.285	0.823
Rate of economically active population	0.320	0.322	0.535
Avg. number of years in apprenticeship per employee (subject to social insurance contribution)	14.707	14.651	0.019
Available infant care places per infant	0.637	0.655	0.339
Available child care places per child	0.281	0.285	0.777
Universities per inhabitant (classified)	0.563	0.373	0.121

Remarks: All variables measured in December 2003. FF is the discretionary budget of a local employment office. ABM denotes the number of participants in job creation schemes, SAM participants in structural adjustment measures. FbW persons participating in long-term training, TM persons participating in short-term training, JUMP the number of participants in a program for the activation of young unemployed persons. FEA = Federal Employment Agency. ESF = European Social Fund.

Appendix 3: Means of Variables Included in the Propensity Score Specification Before (first row) and After (second row) Matching

	Men			Women				
	Treated	Controls	p-value	Treated	Controls	p-value	Data source	Propensity score specification
Age								
18 to 24 years	0.185	0.194	0.423	0.229	0.250	0.056	Survey	1
	0.185	0.186	0.922	0.231	0.231	0.980		
25 to 34 years	0.166	0.208	0.000	0.218	0.233	0.173	Survey	1
	0.166	0.168	0.848	0.217	0.221	0.767		
35 to 44 years	0.222	0.201	0.052	0.220	0.212	0.394	Survey	1
	0.221	0.219	0.872	0.220	0.216	0.749		
45 to 57 years	0.427	0.397	0.026	0.332	0.306	0.025	Survey	1
	0.427	0.426	0.931	0.332	0.332	0.999		
Schooling								
Secondary general school	0.465	0.500	0.008	0.421	0.449	0.023	Survey	1
	0.465	0.472	0.622	0.423	0.425	0.921		
Intermediate secondary school	0.303	0.264	0.002	0.386	0.350	0.003	Survey	1
	0.303	0.297	0.690	0.384	0.378	0.664		
University entrance diploma	0.167	0.151	0.100	0.144	0.133	0.182	Survey	1
	0.167	0.164	0.804	0.144	0.146	0.869		
Other or missing	0.066	0.084	0.010	0.049	0.068	0.002	Survey	1
	0.066	0.067	0.907	0.049	0.052	0.633		
Migration background								
Migrant	0.244	0.260	0.174	0.248	0.258	0.379	Survey	1
	0.244	0.249	0.725	0.248	0.255	0.575		
Household size								
1 person	0.418	0.464	0.001	0.234	0.244	0.323	Survey	1
	0.418	0.422	0.816	0.233	0.236	0.771		
2 persons	0.203	0.182	0.045	0.353	0.342	0.351	Survey	1
	0.203	0.205	0.841	0.353	0.348	0.751		
3 or more persons	0.379	0.354	0.049	0.414	0.414	0.965	Survey	1
	0.379	0.373	0.687	0.414	0.415	0.955		
Number of children								

No children	0.708	0.736	0.020	0.487	0.486	0.990	Survey	1
	0.708	0.714	0.672	0.485	0.490	0.746		
1 child	0.129	0.125	0.631	0.307	0.294	0.269	Survey	1
	0.129	0.130	0.925	0.308	0.304	0.775		
2 or more children	0.163	0.139	0.013	0.207	0.220	0.212	Survey	1
	0.163	0.156	0.542	0.207	0.206	0.941		
Obstacles to employment								
Disabled	0.162	0.127	0.000	0.076	0.067	0.169	Survey	1
	0.161	0.159	0.820	0.075	0.076	0.949		
Care obligation	0.024	0.020	0.229	0.043	0.042	0.748	Survey	1
	0.024	0.024	0.983	0.043	0.044	0.893		
Status before welfare receipt								
(Minor) employment	0.294	0.323	0.020	0.316	0.305	0.347	Survey	1
	0.294	0.298	0.800	0.317	0.311	0.652		
Labor market history from 2001 to 2004								
Number of half-months unemployed in 2004	12.264	12.300	0.888	9.582	8.871	0.004	Admin	1
	12.272	12.252	0.948	9.563	9.566	0.991		
Number of half-months unemployed in 2003	10.215	10.307	0.728	7.973	7.118	0.000	Admin	1
	10.224	10.171	0.863	7.914	7.828	0.757		
Number of half-months unemployed in 2002	8.105	8.059	0.856	6.093	5.532	0.009	Admin	1
	8.107	7.997	0.707	6.048	5.982	0.795		
Number of half-months unemployed in 2001	6.346	6.275	0.757	5.171	4.556	0.002	Admin	1
	6.346	6.193	0.564	5.122	5.000	0.607		
Number of half-months out of labor force from 2001 to 2004	17.056	19.778	0.000	23.952	28.477	0.000	Admin	1
	17.072	17.411	0.678	24.110	24.433	0.714		
Mean duration out of labor force from 2003 to 2004 in half-months	4.393	5.208	0.006	7.117	9.048	0.000	Admin	1
	4.397	4.450	0.765	7.173	7.291	0.775		
Number of programs from 2003 to 2004	0.351	0.384	0.060	0.278	0.262	0.259	Admin	1
	0.351	0.361	0.608	0.274	0.274	0.978		
Mean duration of programs from 2003 to 2004 in half-months	2.351	2.322	0.845	1.973	1.730	0.075	Admin	1
	2.354	2.366	0.946	1.917	1.941	0.883		
Current welfare spell								
Months in welfare before 10/2006	13.862	13.757	0.659	14.532	14.595	0.770	Survey	1
	13.861	13.747	0.678	14.539	14.485	0.829		
Start after 10/2006 or missing	0.158	0.157	0.965	0.129	0.126	0.721	Survey	1

	0.158	0.158	0.998	0.128	0.126	0.827		
Regional information								
Unemployment ratio (high)	0.250	0.223	0.017	0.282	0.231	0.000	Regional	1
	0.250	0.236	0.278	0.279	0.264	0.261		
Urban district	0.167	0.369	0.000	0.163	0.371	0.000	Regional	1
	0.167	0.164	0.849	0.165	0.165	0.962		
Further regional variables								
GDP per employed person (high)	0.265	0.308	0.001	0.262	0.320	0.000	Regional	2
	0.266	0.300	0.013	0.263	0.299	0.005		
Population density (high)	0.210	0.393	0.000	0.205	0.392	0.000	Regional	2
	0.210	0.209	0.972	0.207	0.211	0.704		
Labor market conditions above average	0.358	0.284	0.000	0.346	0.304	0.000	Regional	2
	0.358	0.373	0.309	0.347	0.367	0.141		
Labor market conditions on average	0.311	0.315	0.750	0.318	0.304	0.233	Regional	2
	0.312	0.284	0.051	0.315	0.278	0.006		
Labor market conditions below average	0.331	0.388	0.000	0.336	0.382	0.000	Regional	2
	0.331	0.338	0.612	0.339	0.350	0.395		
East Germany	0.262	0.212	0.000	0.294	0.223	0.000	Regional	2
	0.262	0.247	0.283	0.290	0.277	0.312		
Further sociodemographic variables								
At least one child aged below 3 in the household	0.115	0.108	0.460	0.180	0.165	0.101	Survey	3
	0.115	0.113	0.889	0.180	0.182	0.900		
Lone parent status	0.023	0.019	0.300	0.295	0.299	0.752	Survey	3
	0.023	0.022	0.938	0.296	0.291	0.720		
Professional qualification								
None	0.229	0.272	0.000	0.272	0.327	0.000	Survey	3
	0.230	0.232	0.858	0.274	0.274	0.987		
In-firm training	0.464	0.456	0.548	0.425	0.387	0.002	Survey	3
	0.465	0.465	0.986	0.423	0.423	0.998		
Off-the-job training	0.174	0.153	0.032	0.208	0.185	0.018	Survey	3
	0.174	0.175	0.967	0.208	0.205	0.807		
University degree	0.071	0.067	0.613	0.053	0.061	0.197	Survey	3
	0.071	0.069	0.837	0.054	0.056	0.783		
Other or missing	0.061	0.052	0.103	0.042	0.041	0.870	Survey	3
	0.061	0.059	0.839	0.042	0.042	0.891		

Self-assessment of overall state of health								
Good	0.556	0.576	0.130	0.593	0.620	0.029	Survey	3
	0.557	0.558	0.956	0.596	0.593	0.853		
Satisfactory	0.245	0.235	0.346	0.230	0.210	0.048	Survey	3
	0.245	0.241	0.778	0.228	0.227	0.969		
Poor	0.194	0.186	0.440	0.175	0.168	0.511	Survey	3
	0.194	0.197	0.825	0.175	0.178	0.778		
Missing	0.004	0.003	0.430	0.002	0.002	0.896	Survey	3
	0.004	0.005	0.935	0.002	0.002	0.996		
Impairments to health								
Gastro-intestinal diseases	0.152	0.155	0.701	0.152	0.176	0.010	Survey	3
	0.152	0.156	0.691	0.153	0.152	0.964		
Cardiovascular diseases	0.175	0.190	0.150	0.217	0.224	0.495	Survey	3
	0.175	0.178	0.856	0.217	0.218	0.925		
Rheumatism and other articular trouble	0.288	0.288	0.971	0.261	0.247	0.198	Survey	3
	0.289	0.287	0.913	0.259	0.263	0.767		
Sleep disorders	0.230	0.244	0.222	0.260	0.280	0.062	Survey	3
	0.230	0.231	0.946	0.262	0.264	0.837		
Nervous disorders	0.171	0.177	0.585	0.224	0.232	0.454	Survey	3
	0.172	0.173	0.872	0.223	0.223	0.994		
Allergies	0.173	0.168	0.608	0.252	0.272	0.077	Survey	3
	0.173	0.171	0.887	0.254	0.252	0.890		
Back complaint	0.418	0.405	0.340	0.423	0.414	0.468	Survey	3
	0.418	0.415	0.858	0.422	0.422	0.973		
Other complaints	0.048	0.045	0.583	0.040	0.037	0.437	Survey	3
	0.048	0.050	0.846	0.040	0.040	0.998		
No health problems	0.282	0.289	0.550	0.273	0.257	0.158	Survey	3
	0.281	0.288	0.669	0.273	0.257	0.217		
Self-assessment of daily working capacity								
Less than 3 hours	0.042	0.041	0.908	0.039	0.044	0.377	Survey	3
	0.042	0.043	0.781	0.039	0.041	0.737		
3 to 6 hours	0.077	0.076	0.794	0.183	0.178	0.643	Survey	3
	0.078	0.077	0.987	0.183	0.182	0.966		
6 to 8 hours	0.131	0.124	0.437	0.235	0.223	0.232	Survey	3
	0.130	0.128	0.859	0.234	0.231	0.749		

8 or more hours	0.706	0.726	0.105	0.514	0.528	0.269	Survey	3
	0.706	0.708	0.925	0.514	0.518	0.821		
Missing	0.044	0.034	0.045	0.029	0.028	0.754	Survey	3
	0.044	0.043	0.863	0.029	0.028	0.869		
Self-assessment of basic skills measured from 1 (= very good) to 6 (= fail); Missings are set to 3,5								
Reading and Writing (in mother tongue)	2.121	2.080	0.138	1.920	1.885	0.155	Survey	3
	2.121	2.115	0.854	1.916	1.914	0.934		
Mathematics	2.370	2.326	0.108	2.549	2.595	0.090	Survey	3
	2.369	2.360	0.792	2.546	2.539	0.823		
Emails and Internet	2.993	2.984	0.845	3.113	3.079	0.415	Survey	3
	2.993	3.000	0.898	3.120	3.118	0.970		
Other skills								
Driver's license	0.700	0.637	0.000	0.635	0.586	0.000	Survey	3
	0.700	0.695	0.744	0.634	0.635	0.954		
Further information on the labor market history from 2001 to 2004								
Number of half-months employed in 2004	4.645	4.516	0.531	5.260	5.232	0.894	Admin	3
	4.642	4.700	0.808	5.299	5.279	0.935		
Number of half-months employed in 2003	6.488	6.105	0.108	6.353	6.463	0.630	Admin	3
	6.484	6.574	0.751	6.406	6.396	0.970		
Number of half-months employed in 2002	7.784	7.539	0.343	7.284	7.507	0.355	Admin	3
	7.785	7.897	0.714	7.299	7.324	0.928		
Number of half-months employed in 2001	8.562	8.649	0.747	7.507	7.877	0.133	Admin	3
	8.570	8.680	0.724	7.550	7.621	0.801		
Number of half-months seeking for a job while employed in 2004	0.467	0.479	0.831	0.442	0.517	0.182	Admin	3
	0.467	0.476	0.898	0.446	0.464	0.777		
Number of half-months seeking for a job while employed in 2003	0.263	0.268	0.895	0.260	0.310	0.222	Admin	3
	0.263	0.260	0.948	0.261	0.266	0.907		
Number of half-months seeking for a job while employed in 2002	0.199	0.209	0.795	0.188	0.173	0.634	Admin	3
	0.199	0.205	0.879	0.189	0.186	0.943		
Number of half-months seeking for a job while employed in 2001	0.143	0.151	0.820	0.169	0.147	0.472	Admin	3
	0.143	0.141	0.952	0.166	0.153	0.709		
Number of half-months in a program in 2004	1.818	1.756	0.594	1.551	1.278	0.006	Admin	3
	1.819	1.809	0.939	1.498	1.494	0.970		
Number of half-months in a program in 2003	1.411	1.401	0.927	1.132	1.074	0.545	Admin	3
	1.413	1.412	0.994	1.103	1.148	0.685		

Number of half-months in a program in 2002	1.459	1.529	0.562	1.387	1.188	0.055	Admin	3
	1.461	1.430	0.817	1.376	1.368	0.948		
Number of half-months in a program in 2001	1.527	1.507	0.867	1.394	1.150	0.018	Admin	3
	1.523	1.544	0.882	1.376	1.350	0.833		
Number of employment spells in 2004	0.306	0.310	0.779	0.298	0.316	0.168	Admin	3
	0.306	0.308	0.880	0.299	0.300	0.964		
Number of employment spells in 2003	0.216	0.240	0.075	0.168	0.199	0.007	Admin	3
	0.216	0.221	0.738	0.170	0.171	0.900		
Number of employment spells in 2002	0.232	0.244	0.389	0.219	0.234	0.222	Admin	3
	0.232	0.239	0.655	0.219	0.212	0.616		
Number of employment spells in 2001	0.282	0.296	0.366	0.225	0.258	0.011	Admin	3
	0.282	0.285	0.867	0.226	0.225	0.911		
Number of unemployment spells in 2004	0.764	0.802	0.035	0.588	0.590	0.906	Admin	3
	0.765	0.771	0.773	0.589	0.587	0.912		
Number of unemployment spells in 2003	0.339	0.384	0.008	0.252	0.265	0.322	Admin	3
	0.339	0.347	0.668	0.252	0.256	0.773		
Number of unemployment spells in 2002	0.361	0.405	0.011	0.270	0.279	0.551	Admin	3
	0.360	0.376	0.432	0.269	0.265	0.783		
Number of unemployment spells in 2001	0.348	0.366	0.292	0.256	0.249	0.597	Admin	3
	0.348	0.344	0.818	0.253	0.247	0.717		
Number of spells of job seeking while employed in 2004	0.081	0.098	0.052	0.077	0.089	0.094	Admin	3
	0.081	0.082	0.947	0.077	0.079	0.884		
Number of spells of job seeking while employed in 2003	0.065	0.059	0.410	0.054	0.058	0.540	Admin	3
	0.065	0.065	0.966	0.055	0.054	0.940		
Number of spells of job seeking while employed in 2002	0.050	0.047	0.610	0.042	0.038	0.454	Admin	3
	0.050	0.054	0.665	0.042	0.044	0.808		
Number of spells of job seeking while employed in 2001	0.030	0.033	0.505	0.040	0.033	0.106	Admin	3
	0.030	0.030	0.955	0.039	0.039	0.956		
Number of programs in 2002	0.142	0.125	0.092	0.114	0.094	0.022	Admin	3
	0.142	0.140	0.842	0.112	0.110	0.837		
Number of programs in 2001	0.122	0.121	0.950	0.122	0.088	0.000	Admin	3
	0.122	0.121	0.935	0.117	0.113	0.750		
Number of spells out of labor force in 2004	0.222	0.262	0.003	0.253	0.326	0.000	Admin	3
	0.222	0.223	0.956	0.254	0.257	0.856		
Number of spells out of labor force in 2003	0.168	0.214	0.000	0.145	0.172	0.011	Admin	3

	0.168	0.171	0.849	0.145	0.149	0.726		
Number of spells out of labor force in 2002	0.189	0.218	0.026	0.173	0.191	0.119	Admin	3
	0.189	0.197	0.596	0.175	0.169	0.659		
Number of spells out of labor force in 2001	0.175	0.218	0.001	0.150	0.195	0.000	Admin	3
	0.175	0.178	0.852	0.151	0.155	0.701		
Observations before matching	2066	4194		2423	4603			
Observations of the matched sample	2064	4194		2401	4603			

Remarks: Treated are those individuals who are registered at decentralized welfare agencies, whereas the controls are registered at centralized welfare agencies. The p-values derive from t-tests on equality of means of the displayed variables for treated and controls before (first row) and after (second row) matching. The data sources for the variables are the survey of welfare recipients (Survey), administrative data of the Federal Employment Agency (Admin) and regional data (Regional). Variables marked by 1 in the final column of the table are included in the preferred specification of the propensity score as well as in the sensitivity analyses. Variables indicated by 2 are used for the propensity score specifications in the sensitivity analyses with additional regional variables. Variables marked by 3 are only included in the propensity score specification in the sensitivity analysis with all covariates. The displayed means and the number of obsevations after matching refer to this specification with all covariates. Due to the common support restriction, 2 treated individuals in the sample of men and 21 treated individuals in the sample of women had to be excluded from the matching analysis. The macroeconomic variables (unemployment ratio, GDP per employed person, population density) are binary dummy variables. They are measured in December 2003 and have been collected for all 439 German districts. Districts that face an unemployment ratio larger than the 75^{th} percentile of all unemployment ratios across districts (agencies) are considered to be districts (agencies) with a high unemployment ratio and individuals registered at these agencies are classified respectively. The same procedure applies for GDP and population density. The classification of labor market conditions (above average, on average, below average) is based on the results of Arntz et al. (2006). The upper tercile of districts, which face the best economic conditions with respect to the regional variables relevant for labor market transitions of the long-term unemployed, are classified to have above average labor market conditions. The middle tercile of districts is subsumed to have average labor market conditions, and the lower tercile has below average conditions.

Appendix 4: Preferred Propensity Score Specifications

	Total sample		Singles		Nonsingles	
	Men	Women	Men	Women	Men	Women
Age (reference: 25 to 34 years)						
18 to 24 years	0.0433**	-0.0017	0.0379	0.0674	0.0323	-0.0161
	(0.0220)	(0.0199)	(0.0344)	(0.0430)	(0.0311)	(0.0215)
35 to 44 years	0.0551***	0.0042	0.0542*	-0.0080	0.0559**	0.0069
	(0.0205)	(0.0188)	(0.0284)	(0.0484)	(0.0262)	(0.0194)
45 to 57 years	0.0527***	0.0166	0.0571**	-0.0021	0.0444	0.0294
	(0.0184)	(0.0199)	(0.0270)	(0.0369)	(0.0279)	(0.0216)
Schooling (reference: secondary general school)						
Intermediate secondary school	0.0376**	0.0147	0.0370	0.0136	0.0356*	0.0144
	(0.0171)	(0.0184)	(0.0257)	(0.0325)	(0.0210)	(0.0203)
University entrance diploma	0.0642***	0.0440	0.0765***	0.0701	0.0529*	0.0350
	(0.0200)	(0.0301)	(0.0281)	(0.0543)	(0.0274)	(0.0310)
Other or missing	-0.0164	-0.0458*	0.0203	-0.0034	-0.0404	-0.0562**
	(0.0231)	(0.0239)	(0.0323)	(0.0661)	(0.0354)	(0.0261)
Migration background (reference: non-migrants)						
Migrant	0.0025	0.0334	-0.0073	0.0565	0.0047	0.0301
	(0.0241)	(0.0262)	(0.0299)	(0.0458)	(0.0273)	(0.0260)
Household size (reference: 2 persons)						
1 person	-0.0239	-0.0024				
	(0.0175)	(0.0198)				
3 or more persons	-0.0281	0.0058				
	(0.0267)	(0.0184)				
Number of children (reference: 1 child)						
No children	-0.0270	-0.0224			-0.0091	-0.0244
	(0.0266)	(0.0188)			(0.0214)	(0.0196)
2 or more children	0.0341	-0.0149			0.0284	-0.0136
	(0.0227)	(0.0181)			(0.0232)	(0.0165)
Obstacles to employment						
Disabled	0.0539***	0.0416	0.0278	0.1213***	0.0796***	-0.0095
	(0.0179)	(0.0261)	(0.0229)	(0.0384)	(0.0255)	(0.0332)
Care obligation	0.0360	0.0184	-0.0427	0.0558	0.1051*	0.0102
	(0.0457)	(0.0322)	(0.0566)	(0.0809)	(0.0637)	(0.0344)
Status before welfare receipt						
(Minor) employment	-0.0254*	0.0101	-0.0180	0.0527*	-0.0303*	-0.0057
	(0.0151)	(0.0116)	(0.0234)	(0.0304)	(0.0165)	(0.0143)
Labor market history from 2001 to 2004						
Number of half-months unemployed in 2004	-0.0016*	-0.0007	-0.0018	-0.0039**	-0.0014	0.0004
	(0.0009)	(0.0011)	(0.0013)	(0.0018)	(0.0013)	(0.0014)
Number of half-months unemployed in 2003	-0.0009	0.0013	-0.0000	0.0043**	-0.0018	0.0001
	(0.0009)	(0.0010)	(0.0012)	(0.0019)	(0.0013)	(0.0012)
Number of half-months unemployed in 2002	0.0003	-0.0008	0.0016	-0.0008	-0.0010	-0.0011
	(0.0009)	(0.0012)	(0.0014)	(0.0018)	(0.0014)	(0.0014)
Number of half-months unemployed in 2001	-0.0002	0.0005	-0.0013	0.0011	0.0009	0.0004
	(0.0009)	(0.0010)	(0.0015)	(0.0016)	(0.0012)	(0.0012)
Number of half-months out of labor force from 2001 to 2004	-0.0008**	-0.0006*	-0.0010**	-0.0010	-0.0006	-0.0006
	(0.0004)	(0.0003)	(0.0005)	(0.0007)	(0.0005)	(0.0004)
Mean duration out of labor force from 2003 to 2004 in half-months	-0.0009	-0.0007	-0.0005	-0.0013	-0.0011	-0.0005
	(0.0008)	(0.0006)	(0.0009)	(0.0015)	(0.0011)	(0.0007)
No. of programs from 2003 to 2004	-0.0238	-0.0074	-0.0412**	-0.0122	-0.0043	-0.0048

	(0.0157)	(0.0184)	(0.0185)	(0.0236)	(0.0196)	(0.0203)
Mean duration of programs from 2003 to 2004 in half-months	0.0003	0.0009	-0.0002	0.0012	0.0004	0.0006
	(0.0013)	(0.0013)	(0.0018)	(0.0023)	(0.0018)	(0.0015)
Current welfare spell						
Months in welfare before 10/2006	0.0006	0.0000	0.0003	0.0043**	0.0009	-0.0014
	(0.0012)	(0.0012)	(0.0015)	(0.0021)	(0.0014)	(0.0014)
Start after 10/2006 or missing	0.0278	0.0208	0.0202	0.1160**	0.0330	-0.0055
	(0.0247)	(0.0266)	(0.0367)	(0.0527)	(0.0317)	(0.0311)
Regional information						
Unemployment ratio (high)	0.0151	0.0427	0.0204	0.0429	0.0120	0.0405
	(0.0973)	(0.1001)	(0.0968)	(0.1036)	(0.1001)	(0.1007)
Urban district	-0.2127**	-0.2236**	-0.1942**	-0.2038**	-0.2273***	-0.2319***
	(0.0861)	(0.0872)	(0.0860)	(0.0916)	(0.0875)	(0.0862)
Observations	6,260	7,026	2,810	1,690	3,450	5,336
Pseudo R^2	0.047	0.046	0.044	0.051	0.050	0.049
Log-Likelihood	-3,783.36	-4,318.83	-1,656.71	-1,022.46	-2,118.80	-3,279.64

Remarks: Displayed are marginal effects and standard errors in brackets. The dependent variable is defined to be 1 if an individual is registered at a decentralized welfare agency. Otherwise, the variable is 0. The unemployment ratio variable is a binary dummy variable. It is measured in December 2003 and has been collected for all 439 German districts. Districts that face an unemployment ratio larger than the 75^{th} percentile of all unemployment ratios across districts (agencies) are considered to be districts (agencies) with a high unemployment ratio and individuals registered at these agencies are classified respectively. Singles are defined to be persons living on their own, while non-singles are individuals living together with at least one other person in the same household. Because of this definition, we have to drop the variables household size and number of children (in the household) for the propensity score estimation of singles. For non-singles, we have to drop the variable household size only. *** denotes $p<0.01$, ** denotes $p<0.05$ and * denotes $p<0.1$.

Appendix 5: Indicators for Matching Quality, Singles

	Men	Women
Before Matching		
McFadden-R^2	0.044	0.051
LR-Test	152.850	110.170
p-value	0.000	0.000
Mean standardized difference in percent	7.026	7.826
After Matching		
McFadden-R^2	0.001	0.001
LR-Test	1.520	0.910
p-value	1.000	1.000
Mean standardized difference in percent	1.038	0.789
Smith and Todd (2005) balancing test		
p-values > 0,05	20	18
p-values > 0,01	20	22

Remarks: McFadden-R^2 derives from a probit estimation of the propensity score on all covariates considered. The LR-statistic and the corresponding p-value derive from a likelihood-ratio test of the joint insignificance of all covariates. The mean standardized difference in percent has been calculated as an unweighted average of all covariates. The Smith-Todd test displays the number of covariates passing the test at the indicated significance level. There are 22 covariates included in the preferred specification.

Appendix 6: Indicators for Matching Quality, Non-singles

	Men	Women
Before Matching		
McFadden-R^2	0.050	0.049
LR-Test	224.190	337.040
p-value	0.000	0.000
Mean standardized difference in percent	6.594	7.761
After Matching		
McFadden-R^2	0.001	0.002
LR-Test	3.340	8.650
p-value	1.000	0.998
Mean standardized difference in percent	1.353	1.228
Smith and Todd (2005) balancing test		
p-values > 0,05	22	15
p-values > 0,01	22	19

Remarks: McFadden-R^2 derives from a probit estimation of the propensity score on all covariates considered. The LR-statistic and the corresponding p-value derive from a likelihood-ratio test of the joint insignificance of all covariates. The mean standardized difference in percent has been calculated as an unweighted average of all covariates. The Smith-Todd test displays the number of covariates passing the test at the indicated significance level. There are 24 covariates included in the preferred specification.