Abstract—This paper demonstrates how economic theory can be combined with state-of-the-art empirics to make quantitative statements about optimal policy. Using a large administrative data set for West Germany as well as nonparametric and semiparametric estimation techniques, I parameterize the model proposed by Pavoni and Violante (2005, 2007) for the optimal choice of different financial and activation measures for unemployed workers. The parameterized model is used to study the role of job search assistance in optimal welfare-to-work programs and assess the optimality of the West German policy in the period 2000 to 2002.

I. Introduction

In most countries, labor market policies for the unemployed rely on two sets of measures. Passive income support during unemployment provides consumption insurance and gives job seekers the opportunity to look for appropriate jobs. Active measures aim at increasing job seekers’ exit rates to employment, for example, by improving search effectiveness or skills. Welfare-to-work (WTW) programs are government programs that combine active and passive measures. Expenditures on such programs are substantial, ranging from 1% to 5% of GDP in OECD countries (OECD, 2007).

This paper studies the role of job search assistance programs (JA) in optimal WTW schemes. These programs offer a combination of intensified counseling and job search training, as well as some availability checks or other forms of weak monitoring that make job seekers search more effectively. The objective is to increase the exit rate from unemployment to employment. Job search assistance programs belong to the most widely used activation measures in OECD countries (OECD, 2007). They are relatively inexpensive and have proven to be effective in many countries (Fay, 1996; Heckman, LaLonde, & Smith, 1999; Martin & Grubb, 2001; Klueve & Schmidt, 2002; Thomsen, 2009).

The analysis is based on the framework developed by Pavoni and Violante (2007) in its extended form with training (Pavoni & Violante, 2005). This framework is particularly well suited to study optimal policy for the following reasons. Benefit levels, wage taxes or subsidies, and policy instruments are chosen endogenously for each period of unemployment. The environment is dynamic, with duration dependence in the returns to search. Both the cost-effectiveness of single measures and the possibility of alternative uses of funds to obtain a given policy objective are taken into account. Thus, cost-effectiveness is a necessary but not sufficient condition for a policy measure to be part of the optimal scheme. As a first contribution, I derive a necessary condition for cost-effective JA to be part of the optimal scheme if there is no negative duration dependence in the exit rates to employment that provides a natural role for using JA.

The second contribution of my paper is to demonstrate how this theoretical framework can be combined with state-of-the-art econometrics to quantitatively study the optimal design of active labor market policies. My approach is inspired by the literature on sufficient statistics (see the survey by Chetty, 2009) where credible reduced-form estimates are used in combination with economic theory to inform optimal policy choice. However, I am interested in the optimal combination of various policy instruments in a dynamic setting, where it is hard to follow a pure sufficient-statistic approach without imposing strong restrictions. Therefore, I combine the ideas from this literature with classical numerical simulation approaches. I use administrative data for West Germany to nonparametrically or semiparametrically estimate the returns to search and to participation in JA as well as the features of the actual policy. In the empirical analysis, I impose as few assumptions as the data allow and ensure that all parameters are estimated for the same population of interest by addressing the relevant selection problems. The estimates are used to parameterize the model to resemble West Germany in the period 2000 to 2002. West Germany is comparable to most industrialized OECD countries, and it is the largest among the typical continental European economies where JA is an important and intensively used instrument to combat unemployment.

Although the majority of studies finds positive effects, there is quite substantial variability in the size of the estimated effects of JA.1 Therefore, as a third contribution, I use the parameterized model to analyze the optimal use of JA as a function of its effectiveness in raising exit rates to employment. I find that the minimum effectiveness required for JA to be used at all in the optimal policy is at the lower bound of existing estimates of positive program effects. Hence, JA will usually be part of optimal WTW programs. If sufficiently effective, JA is used to delay or prevent situations in which it is no longer optimal to incentivize the worker to provide a positive search effort because the returns to search are too low.

Therefore, it should be targeted at workers who enter unemployment with low search skills or at long-term unemployed workers whose search skills have depreciated substantially.

The fourth contribution of my paper is a comparison of the policy actually implemented in West Germany in the period 2000 to 2002 with the optimal scheme. The results indicate a substantial inefficiency of the German system. It provides too much insurance and puts too little emphasis on search incentives. The benefit schedule is much too flat compared to the optimal one. Moreover, expenditures on JA were wasted, as JA was ineffective in the period 2000 to 2002. I also show that there would have been considerable gains from implementing tight monitoring in West Germany. If available, monitored search would crowd out unmonitored search to a large extent in the optimal policy, especially for workers with low expected exit rates.

Most existing studies that take a normative perspective on WTW programs have entirely focused on the optimal time profile of benefits if a search effort is unobservable (some of the main contributions are Shavell & Weiss, 1979; Wang & Williamson, 1996; Davidson & Woodbury, 1997; Hopenhayn & Nicolini, 1997; Cahuc & Lehmann, 2000; Fredriksson & Holmlund, 2001). This literature shows that benefits have to decline with unemployment duration in order to provide search incentives. Most of the still limited normative literature that takes into account both passive and active measures studies the role of job search monitoring. For example, Fredriksson and Holmlund (2006) and Boone et al. (2007) compare optimal unemployment insurance with and without monitoring and sanctions in a general equilibrium framework. Assuming random assignment of monitoring, they show that welfare increases when monitoring is used. In a partial equilibrium framework, Setty (2009) and Wunsch (2011) endogenize monitoring intensity and the size of the benefit sanction. They derive their optimal values as a function of unemployment duration and, in the case of Wunsch (2011), worker characteristics too. Other active measures have been analyzed by Coles and Masters (2000), who consider job creation subsidies and retraining, and Blanchard and Titré (2008), who study layoff taxes. Moreover, Cardullo and Van der Linden (2007) and Van der Linden (2003a, 2003b) explicitly analyze the interactions between the design of the unemployment insurance system and the use of active measures like employment subsidies, monitoring, and training.

The study that is most closely related to my work is Spinnewijn (2013) who analyzes the optimal use of training that increases work-related skills in a framework with human capital depreciation during unemployment. He shows that effective training can be used to prevent reemployment wages from deteriorating to levels where providing search incentives is no longer optimal. Combining this result with my results for job search assistance implies complementary roles for training that affects reemployment wages and JA. In the presence of negative duration dependence in both exit rates to employment and reemployment wages, only the combination of both measures can prevent the returns to search from deteriorating to prohibitively low levels during unemployment. Another related study comes from Ribi (2009), who analyzes the optimal size of job search assistance programs in a static general equilibrium framework with homogeneous agents. She shows a trade-off between the positive effects of JA on the participants and the marginal increase in taxes needed to finance JA and distorts employment decisions. Moreover, she finds that the optimal size of the program decreases with increasing generosity of the unemployment system because additional funds have to be raised.

The remainder of the paper is organized as follows. The next section summarizes the key features of the model. Section III contains the quantitative analysis for West Germany. I describe in detail how the model is parameterized using the administrative data and discuss the results. The last section concludes. The appendix contains the proof of the theoretical result and additional details on the estimation of the parameters. Further information on the data and the quantitative analysis is provided in an appendix that is available on the Internet.4

II. The Model

A. The Setup

The setup that Pavoni and Violante (2005, 2007) consider is based on the dynamic principal-agent problem that Hopenhayn and Nicolini (1997) studied. A risk-neutral planner (the principal), such as the representative of the unemployment insurance or the welfare system, faces a risk-averse unemployed worker (the agent) whose effort with respect to his activity, job search, or participation in training is unobservable to the planner.

The objective of the planner is to design a contract that insures the worker against the failure of his activity by providing some income during unemployment but also incentivizes the worker to exert an appropriate level of effort. This contract maximizes the expected discounted net fiscal revenue of the insurer, subject to providing the agent with at least an expected discounted utility level of $U_0$. From a policy perspective, this is a more relevant case than the dual problem of utilitarian expected welfare maximization subject to a budget constraint. Moreover, it allows writing the optimal

1 However, the empirical literature is quite pessimistic regarding the cost-effectiveness of the type of training studied by Spinnewijn (2013). These programs are rather costly and with only few exceptions (Lechner et al., 2011), their effects on labor market outcomes are at most small (see Fay, 1996; Heckman et al., 1999; Martin & Grubb, 2001; Dolton & O’Neill, 2002; Klue & Schmidt, 2002; Blundell et al., 2004; OECD, 2005). Wunsch (2007) shows that the minimum effectiveness required for these programs to be part of the optimal policy exceeds the few existing estimates of positive program effects. Therefore, my study focuses on job search assistance.

2 In other periods, JA was also effective in Germany. See Thomsen (2009).

3 However, the empirical literature is quite pessimistic regarding the cost-effectiveness of the type of training studied by Spinnewijn (2013). These programs are rather costly and with only few exceptions (Lechner et al., 2011), their effects on labor market outcomes are at most small (see Fay, 1996; Heckman et al., 1999; Martin & Grubb, 2001; Dolton & O’Neill, 2002; Klue & Schmidt, 2002; Blundell et al., 2004; OECD, 2005). Wunsch (2007) shows that the minimum effectiveness required for these programs to be part of the optimal policy exceeds the few existing estimates of positive program effects. Therefore, my study focuses on job search assistance.


5 Note that this is equivalent to minimizing net expenditure.
contract in a convenient recursive form (see Spear & Srivastava, 1987). The promised utility, \( U_0 \), is exogenously given (for example, the outcome of voting). It can be regarded as a measure of the generosity of the welfare system. Its level may depend on worker characteristics. In particular, the planner can give different weights to different types of individuals in the social welfare function by requiring different \( U_0 \). Net fiscal revenue is tax revenue if the worker is employed and expenditures on benefit payments, monitoring, and JA if the worker is unemployed.

Workers are infinitely lived and maximize expected lifetime utility discounted at the rate \( \beta \in (0, 1) \). They have time-separable preferences over consumption, \( c \geq 0 \), and effort, \( a \), with period utility given by \( u(c, e) - a \). Effort is assumed to be either high (\( a = e > 0 \)) or low (\( a = 0 \)), underlining the role of fixed costs and the extensive margin of participation decisions. The disutility of effort during employment is generally not restricted to be the same as during unemployment, but it must be ensured that accepting a job offer always dominates staying unemployed.\(^6\)

Unemployed workers find jobs with probability \( \pi(p, a) \), where \( p \) denotes the worker’s level of job search skills or search effectiveness. For simplicity and to focus on the role of job search assistance programs in optimal WTW programs, I abstract from heterogeneity in other dimensions of skills.\(^7\) The exit rates are normalized to \( \pi(p, 0) \equiv 0 \), and I assume that \( \pi(p, e) \equiv \pi(p) \in (0, 1) \) is continuous and increasing in \( p \). I allow for depreciation in search skills during unemployment. This induces negative duration dependence in \( \pi \), an empirically well-established fact in labor economics (see the survey of the relevant literature by Machin & Manning, 1999). It can be explained by the stock-flow matching approach (Coles & Smith, 1998). At the beginning of unemployment, job seekers start with the existing stock of vacancies and may look for a particular type of job. If their job search activities fail, they will have exhausted this stock of vacancies after some time and will start sampling from a new flow of vacancies. They may also have to change or extend the range of jobs they are looking for.

Job search assistance may be used to increase initially low levels of search effectiveness or restore search skills that depreciate during unemployment. The outcome of JA is stochastic. Search skills increase during JA with probability \( \theta(a) \), which depends on the effort exerted during program participation, where \( \theta \equiv \theta(e) > \theta(0) \equiv 0 \). For simplification, to avoid the necessity of modeling time allocation within a period, it is assumed search and program participation are mutually exclusive activities within a period, implying that participants cannot exit to employment directly from the program.\(^8\) However, since the length of a period can be arbitrarily small, this assumption is not restrictive.

The planner can observe whether the worker is employed or unemployed, his activity (job search or participation in JA) during unemployment, and the outcome of the worker’s activity, \( y \in \{s, f\} \), where \( s \) denotes success and \( f \) failure. But the planner cannot observe the worker’s effort choice, \( a \), so he faces a moral hazard problem. For each period \( t \) and contingent on all observable histories up to \( t \), the contract specifies the transfers to the worker, the policy instrument to be used, and the corresponding recommended effort choice of the worker.

For simplification it is assumed that workers do not have access to savings, insurance, or credit markets. In particular, it is assumed that workers cannot self-insure against the random outcome \( y \in \{s, f\} \) of their activity (for example, by saving). Pavoni and Violante (2005) show that when workers can save through credit markets but still face a no-borrowing constraint, a reasonable assumption for unemployed workers, the same optimal contract can be implemented by introducing a linear, time-invariant interest tax.

### B. The Planner’s Problem

Pavoni and Violante (2005, 2007) show a recourse formulation of the optimization problem of the planner. The state variables are the promised continuation utility, \( U \), and search effectiveness, \( p \). At the beginning of each period, the planner chooses the optimal policy instrument, \( i(U, p) \), for an unemployed worker who enters the period with state \( (U, p) \) by solving

\[
V(U, p) = \max_{i \in \{U, JM, JA\}} V^i(U, p),
\]

where \( V \) is the upper envelope of the values associated with different available policy instruments. In each case, the planner chooses an effort recommendation \( a(U, p) \), the transfer \( c(U, p) \), and the continuation utilities \( U^i(U, p) \) conditional on the outcome \( y \in \{f, s\} \) of the worker’s activity. There is always a promise-keeping (PK) constraint, which ensures delivery of the promised utility \( U \) to the worker. Moreover, whenever positive effort is required and effort cannot be verified, a second incentive compatibility (IC) constraint ensures that it is optimal for the worker to choose this effort level.

Following Hopenhayn and Nicolini (1997), employment is assumed to be absorbing without informational asymmetries in order to concentrate on the (current) unemployment experience.\(^9\) The planner can also tax or subsidize the wage

---

\(^6\) For example, assuming that effort is the same during employment and search would imply that the worker searches for a job with the same intensity and time input as he works.

\(^7\) Wunsch (2007) shows that the key features of the framework by Pavoni and Violante (2005, 2007) and their main results on optimal transfers and policies are insensitive to both the distinction between and joint consideration of work-related and search-related skills.

\(^8\) This corresponds to a so-called lock-in effect of program participation, which has been documented in many empirical studies (Gerth & Lechner, 2002; Van Ours, 2004; Lechner et al., 2011; Sianesi, 2004, 2008; Jespersen, Munch, & Skipper, 2008). The assumption is relaxed in Spinnewijn (2013).

\(^9\) Qualitative results for the same unemployment spell do not change as long as the job separation rate is exogenous. Optimal contracts with endogenous job separation are studied by Zhao (2002) and Hopenhayn and Nicolini (2009), who show that in this case, the optimal contract has to take into account the worker’s full employment history.
of employed workers. He solves \( W(U) = \max_{c,U^s} w - c + \beta W(U^s) \) subject to \( U = u(c) - e^W + \beta U^s \), where \( e^W \) denotes the disutility of work. The absence of the information asymmetry and the absorbing nature of employment imply that \( W(U) = \frac{w - c}{1 - \beta} \), where \( c^W(U) = u^{-1}((1 - \beta)U + e^W) \). The implicit wage tax (if positive) or subsidy (if negative) that is imposed on employed workers is thus constant and given by \( \tau(U) = w - e^W(U) \).

Unemployment insurance (UI) requires the worker to search with effort \( e \), and he finds a job with a probability \( \pi(p) \). With probability \( 1 - \pi(p) \), the worker remains unemployed, and his job search skills depreciate to \( p^f \leq p \):

\[
V^{UI}(U, p) = \max_{c,U^s} \left[ -c + \beta[\pi(p)W(U^s)] + (1 - \pi(p))V(U^f, p^f) \right],
\]

\[s.t. \ U = u(c) - e + \beta[\pi(p)U^s + (1 - \pi(p))U^f] \]

\[U \geq u(c) + \beta U^f.\]

Using the PK constraint, the cost of having to obey the IC constraint during UI can be expressed as \( U^s - U^f \geq \frac{e}{\beta \pi(p)} \). It is increasing in the required effort level and decreasing in the unemployment hazard, which in turn is increasing in \( p \). Hence, this cost increases as search effectiveness depreciates during UI. Because \( a(U, p) = e \), there is also an effort compensation cost.

Job search monitoring (JM) is a technology that enables the planner to verify the worker’s search effort against payment of a cost \( \kappa^{JM} \). He may, for example, pay a caseworker to closely monitor the worker’s job search activity. Verifiability of search effort implies that no IC constraint is needed. Hence, there are no incentive costs, but the effort compensation cost remains, as high effort is required from the worker. The planner will never combine monitoring with low effort because any deviation from no search can be observed at no cost.

The planner will never combine monitoring with low effort because any deviation from no search can be observed at no cost. The implicit wage tax (if positive) or subsidy (if negative) that is imposed on employed workers is thus constant and given by \( \tau(U) = w - e^W(U) \).

Hence, there are no incentive costs, but the effort compensation cost remains, as high effort is required from the worker. The planner will never combine monitoring with low effort because any deviation from no search can be observed at no cost. The implicit wage tax (if positive) or subsidy (if negative) that is imposed on employed workers is thus constant and given by \( \tau(U) = w - e^W(U) \).

Unemployment insurance (UI) requires the worker to search with effort \( e \), and he finds a job with a probability \( \pi(p) \). With probability \( 1 - \pi(p) \), the worker remains unemployed, and his job search skills depreciate to \( p^f \leq p \):

\[
V^{UI}(U, p) = \max_{c,U^s} \left[ -c + \beta[\pi(p)W(U^s)] + (1 - \pi(p))V(U^f, p^f) \right],
\]

\[s.t. \ U = u(c) - e + \beta[\pi(p)U^s + (1 - \pi(p))U^f] \]

\[U \geq u(c) + \beta U^f.\]

C. Optimal Sequence of Transfers

Because of the absence of the IC constraint under JM (no information asymmetry) and SA (no effort required), it is easy to see that promised utility remains constant and full insurance will be provided. Consequently, benefits remain constant during JM and SA. For UI and JA, the presence of the IC constraint requires promised utility and, hence, benefits to fall on failure of job search or JA (as punishment) and to increase on their success (as reward). The wage tax or subsidy on reemployment is one instrument to incentivize the worker. Because promised utility remains constant during JM, the wage tax or subsidy on reemployment remains constant as well. On the other hand, the wage tax (subsidy) increases (decreases) during UI and JA because promised utility falls if search effectiveness remains constant. Once it depreciates, the opposite may happen.

D. Optimal Choice and Sequence of Policies

Pavoni and Violante (2007) show that SA is always absorbing because it is used at a point from which it is no longer optimal to incentivize the worker. They also show that if both reemployment wages and the exit rates to employment remain constant during unemployment, UI and JM are absorbing as well because no policy change is required as long as the

\[U = u(c) + \beta U^f.\]
returns to search remain unchanged during unemployment. Hence, human capital dynamics are necessary for policy transitions to occur. This holds in the model with any type of training. Here, human capital dynamics can occur for two reasons: skill depreciation, during unemployment and skill accumulation during training. Consequently, in the absence of skill depreciation, a necessary condition for policy transitions to occur is the optimality of training for some types of workers at some unemployment duration. In appendix A, I show that this requires \( \theta \leq \pi(p) \) if \( V \) is submodular. The reason is that JM and SA are absorbing without skill depreciation and that \( V^U \) is more negatively sloped than \( V^J \) because the incentive costs of JM, \( U^J - U^J \geq \frac{\epsilon}{\theta} \), are higher than the incentive costs of UI, \( U^U - U^J \geq \frac{e}{\beta \pi(p)} \), if \( \theta \leq \pi(p) \). Consequently, if JA is not used in the first period of unemployment at \( (U_0(p_0), p_0) \), it will never be used if \( \theta \leq \pi(p_0) \). Thus, JM can be optimal only if initial search effectiveness, \( p_0 \), and, hence \( \pi(p_0) \), is not too high and if JA is sufficiently effective, that is, \( \theta \) is sufficiently high. But \( \theta > \pi(p_0) \) is not sufficient for JA to be optimal because there is also a positive fixed cost, \( \kappa^{JA} \).

When search skills depreciate during unemployment, neither UI nor JM is absorbing. Since search effectiveness behaves in the same way during JM and UI, the main difference between JM and UI is that the incentive cost incurred during UI is replaced by the fixed monitoring cost. Since the incentive cost increases because of depreciation of \( p \) while the monitoring cost remains fixed, UI will be used only before JM as long as the incentive cost is lower than the fixed cost, \( \kappa^{JM} \), which becomes less likely with increasing unemployment duration. Once \( \kappa^{JM} \) is lower than the incentive cost during UI, a switch back to UI may occur only after successful use of JA because, otherwise, the incentive cost keeps rising for UI. Essentially JM is a policy instrument to save incentive costs in the presence of negative duration dependence in the exit rate to employment. Consequently, the typical sequence of policies in the absence of JA is UI, followed by JM and then by SA.

Now consider the role of JA. Apart from the effort compensation cost, JA has three main cost components. The first one, the property that participants cannot exit to employment directly from JA, is absent under UI and JM. However, since the length of a period can be arbitrarily small, this is not a big issue. The second cost, component is the incentive cost, which is absent under JM but not under UI. Yet the incentive cost under UI, \( U^U - U^J \geq \frac{\epsilon}{\beta \pi(p)} \), increases with the unemployment duration, while the incentive cost of JA, \( U^J - U^J \geq \frac{e}{\beta} \), remains constant. Thus, the use of JA relative to UI becomes more likely the longer the worker remains unemployed. Finally, there is the fixed cost, \( \kappa^{JA} \), which is absent under UI but, although potentially with a different amount, present under JM. Hence we have the situation that the cost of UI is increasing over time, while that of JA and JM is constant. JM will be preferred to UI once the incentive cost of UI is higher than the incentive cost of JA plus the fixed cost of JA. The advantage of JM relative to JA is the absence of the incentive cost and a potentially smaller fixed cost. However, in contrast to JA, JM cannot prevent the deterioration of search effectiveness and thus the use of SA. Because the worker no longer searches for a job under SA and because SA is absorbing, the optimality of SA implies extremely high net expenditures in the long run. In the absence of JA, JM is used to save costs when the deterioration of search skills and the use of SA cannot be avoided. Once JA is available, it can be used to delay or even prevent the use of SA, which creates a potential for considerably larger cost savings than with the use of JM. Thus, JA is likely to crowd out JM.

III. Quantitative Analysis

In order to derive more detailed insights into the optimal choice over policy instruments, the model is parameterized to resemble the West German economy in the period 2000 to 2002, and then solved numerically for the optimal policy. West Germany is an interesting case to study because it is comparable to most industrialized OECD countries and is the largest among the typical Continental European economies, where WTW programs are important and intensively used instruments to combat unemployment. Moreover, in the course of substantial reforms of the German unemployment insurance system, there have been heated debates about its optimal design in recent years.

A. General Approach

The main conclusion from section IID is that the key determinants for the optimal choice between policy instruments are the exit rates to employment, \( \pi(p) \), the success probability of JA, \( \theta \), and the fixed costs of JM, \( \kappa^{JM} \), and JA, \( \kappa^{JA} \). Acknowledging the importance of these parameters, I estimate them with as few assumptions and restrictions as possible using state-of-the-art identification and estimation methods. The key objective is to bring empirical realism and, hence, more usefulness for actual policy design into a literature dominated by illustrative numerical simulations. Estimation of the parameters is based on a large and very informative administrative database. Thus, they refer to the same time period and economic and institutional environment. I also ensure that the parameters are estimated for the same population of interest by addressing the relevant selection problems. Moreover, I link the simulation results to the population observed in the data, which considerably improves the interpretation and understanding of the results.

Estimation of the preference parameters would require imposition of very strong assumptions that are unlikely to hold in reality (Chetty, 2009). Therefore, I follow the literature when parameterizing them. Although preference parameters are important to determine the optimal policy, they are the same across policy instruments and periods. Thus, they are less important for the relative attractiveness of different policy instruments than their direct costs and returns. However, acknowledging a potential sensitivity of
the quantitative analysis to the choice of preference parameters, I investigate this issue by varying the preference parameters in section IID. As a preview, results do not change for a range of parameter values consistent with empirical estimates.

B. Data and the Population of Interest

The administrative database used is a 2% random sample from all individuals who have been subject to German social insurance since 1990. It covers the period 1990 to 2005 and combines spell information from social insurance records, program participation records, and the benefit payment and job seeker registers of the public employment service (PES). The database comprises detailed and accurate information on several dimensions. It provides information on employment status on a daily basis for the period 1990 to 2005. Personal characteristics include education, age, gender, marital status, number of children, age of youngest child, profession, nationality, disabilities, and health. The benefit payment register provides information on type and amount of benefits received, remaining benefit claims, and imposition of sanctions. The job seeker register includes information on the type of employment looked for, compliance with benefit conditions, and the number of placement propositions by the PES. Moreover, the data comprise information on job characteristics such as type of employment, industry, occupation, qualification, and wages. With respect to program participation, the data cover the type of the program as well as the planned and actual durations. Detailed regional information, which includes federal state, local unemployment rate, GDP, migration, demographic and industry structure, infrastructure and urbanity, complements the database (see online appendix A for a full list of variables).

For the numerical analysis, the model is parameterized to resemble the population of West German workers who entered unemployment from employment between January 2000 and December 2002 and received benefits from the unemployment insurance system. To concentrate on the main body of the workforce, apprentices, young men on civilian or military service, and elderly workers in special forms of employment are excluded. The final sample consists of 76,304 observations and is referred to as the reference population of interest (RPI) in the remainder of the paper. All parameters of interest will be estimated for this population.

C. Parameterization of the Model

Table 1 summarizes the baseline parameterization of the model. The details are provided in the following subsections.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>0.5 months</td>
<td>Smallest reasonable level of aggregation in data</td>
</tr>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu(c)$</td>
<td>ln($c$)</td>
<td>Assumption</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.998</td>
<td>To match interest rate of 4%</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.84</td>
<td>Disutility of effort: fraction of time spent working (BMAS, 2003; Statistisches Bundesamt, 2000–2002; Chari, Christiano, &amp; Kenhoe, 1995)</td>
</tr>
<tr>
<td>Returns to search</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w$</td>
<td>860 euros</td>
<td>Median last wage per half-month: nonparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>$\pi(p)$</td>
<td>Figure 1</td>
<td>Exit rates to employment: semiparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>$b$</td>
<td>0.026</td>
<td>Depreciation in exit rates: semiparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>Job search assistance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s_{JA}$</td>
<td>190 euros</td>
<td>Cost of JA per half-month: from public statistics (BA, 2001–2005)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0</td>
<td>Success rate of JA: semiparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>Monitoring cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s_{JM}$</td>
<td>100 euros</td>
<td>Cost of JM per half-month from public statistics (BA, 2001–2005)</td>
</tr>
<tr>
<td>Actual policy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_{UB}$</td>
<td>345 euros</td>
<td>Unemployment benefit per half-month: nonparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>$b_{JA}$</td>
<td>305 euros</td>
<td>Unemployment assistance per half-month: nonparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>$d_{UB}$</td>
<td>24</td>
<td>Median UB claim in half-months: nonparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>$f_{JA}$</td>
<td>0–1%</td>
<td>Fraction participating in JA in half-month $t$: nonparametric estimate from IEB, 1990–2005</td>
</tr>
<tr>
<td>$f_{JM}$</td>
<td>0</td>
<td>Fraction monitored in half-month $t$: effectively 0 due to capacity constraints</td>
</tr>
<tr>
<td>$\tau$</td>
<td>310 euros</td>
<td>Wage tax per half-month for $w = 860$ euros from legislation</td>
</tr>
</tbody>
</table>

Table 1.—Summary of the Baseline Parameterization

12 This excludes recipients of social assistance, which is administered by local authorities for which no common register exists. For this reason, individual data on these persons are unavailable in Germany.
utility over consumption is assumed to be logarithmic, \( u(c) \equiv \ln(c) \), but specifications with intertemporal elasticities of substitution below and above 1 are tested as well.

To calibrate the disutility of effort, \( e \), I follow the approach that Pavoni and Violante (2007) used, which originates from common practice in macroeconomics. Denote by \( \phi \) the relative weight on leisure versus consumption. Assuming a standard Cobb-Douglas production function and logarithmic utility, the static optimality condition of the worker yields a value of \( \phi = 2.35 \) given a labor share of 0.73 (BMAS, 2003), a consumption-income ratio of 0.72 (Statistisches Bundesamt, 2000–2002) and a fraction of time spent working of \( n = 0.3 \). This implies a value for the disutility of work effort of \( e = \phi[\ln(1) - \ln(1 - n)] = 0.84 \) (Chari et al. 1995), which lies within the range of empirical estimates of \( e \in [0.31, 0.97] \) (see Hausman, 1980; Cogan, 1981; Eckstein & Wolpin, 1989; Keane & Wolpin, 1997). In the baseline case, it is assumed that the disutility of effort during unemployment and work is equal. As a sensitivity check, the case where only half the time spent on work is spent on search and program participation is considered as well.

One important issue is that the model assumes that job search effort is positive and equal to \( e \) whenever the exit rates to employment are positive. In order to be consistent with the model, the parameterization must be consistent, with high job search effort \( e \) being provided by the worker under the actual policy. To investigate whether this is satisfied, I check for each period of unemployment whether incentive compatibility is satisfied given the parameterization of the model and the parameters of the actual policy observed in the data. I find that worker utility conditional on providing high job search effort \( e \) always exceeds worker utility without search when exit rates are positive. Moreover, the utility of working always exceeds the utility of staying unemployed, which implies that all job offers will be accepted.

Returns to search. The reemployment wage is set to the median level of the full-time wage from the employment spell directly before entering unemployment in the RPI in the data (860 euros per half-month or 1,720 euros per month). For part-time workers, the wage is scaled up to obtain an approximate full-time equivalent. This paper abstracts from ex ante heterogeneity and duration dependence in reemployment wages in order to simplify the analysis. In Wunsch (2007, 2010), I show that qualitative results do not change when allowing for this type of heterogeneity and dynamics.

To parameterize the exit rates to employment, I use the following procedure. First, I estimate initial exit rates directly after entering unemployment for each individual in the RPI. They are obtained from a probit model for the probability of exiting unemployment within three months of becoming unemployed. As explanatory variables, I use personal characteristics, detailed information on ten years of individual employment histories, characteristics of the last job, type of job looked for, remaining UI claim, amount of benefit, compliance with benefit conditions, regional characteristics, and as well as year and quarter dummies (see online appendix B for the exact specification and the estimated coefficients of the probit model). The distribution of the predicted probabilities from this model in the RPI is used to obtain a ranking of individuals in the data in terms of unobserved search effectiveness \( p \) (see online appendix B, figure 1, for the distribution of predicted probabilities in the RPI). As described in more detail below, this allows linking a particular parameterized value of \( \pi(p) \) to a specific part of the distribution of \( p \) in the data, which is useful for interpretation of the results.

Second, I calibrate the maximum initial level of \( \pi(p) \) used in the simulation. I use the fraction of individuals in the highest five percentiles of the distribution of \( p \) in the sample (obtained from the above probit model) who exit unemployment within the first month of unemployment.\(^{13} \) Note that there are no selection issues at the beginning of unemployment so that no further adjustments of the raw fraction are needed. This procedure yields \( \pi_{\text{max}} = 19\% \).

Third, starting from \( \pi_{\text{max}} \), I discretize the function \( \pi(p) \) to thirty grid points. The grid points are geometrically spaced at

\[^{13}\text{Because of labor market frictions, the exit rate is very low in the first half-month, and so the one-month exit rate is a better proxy.}\]
rate \( \gamma = 0.1 \) or 10% between \( \pi_{\text{max}} = 19\% \) and \( \pi_{\text{min}} = 0.14 \.

The values of \( p \) corresponding to these grid points are set to \( p = \{1, 2, \ldots, 30\} \). Thus, \( p \) is just a label for the grid points of \( \pi(p) \).

Hence, the model is parameterized directly in terms of the exit rates to employment without making assumptions about the unobserved level of search skills \( p \) and its relation to the exit rates. For each grid of \( p \), the value functions for all policies with respect to \( U \) are computed using Chebychev polynomials up to the twentieth order. Panel a of figure 1 displays the exit rates to employment for the thirty grid points resembled by \( p \). The link to the distribution of search effectiveness in the population of interest will be provided by the distribution of predicted exit probabilities obtained from the data in the first step.

Fourth, the duration dependence of the exit rates to employment is estimated semiparametrically as follows. Ideally we would like to observe the exit probability of everyone in the RPI for different unemployment durations without participation in activation measures. However, in reality, a substantial part of the RPI participates in some program. Moreover, the exit rate between time \( t \) and \( t + 1 \) is observed only for the subpopulation of the RPI that is still unemployed at \( t \).

The problem is that program participation is highly selective and that unemployment duration is endogenous. Program participants and nonparticipants, as well as individuals with different realized unemployment durations, differ systematically from the RPI.

To correct for this selection bias, all individuals who remain unemployed for at least \( t \) periods and do not participate in any program are reweighted in terms of the variables that cause selection bias to resemble the distribution of these characteristics in the RPI. The data ensure that the reweighting procedure accounts for all key factors that determine the job finding probability. It not only takes into account the standard set of sociodemographic characteristics but also health and disability and their effect on employability, the kind of job that a person is looking for, occupation and industry-specific experience, incentives provided by the UI system in terms of benefit claims and benefit sanctions, compliance with benefit conditions as a proxy for job search motivation and intensity, and past performance on the labor market over the past ten years, which allows controlling indirectly for ability, motivation, productivity, and individual employment-related preferences. Finally, differences in regional labor markets are accounted for in a very detailed way.

The reweighing is implemented using a robust radius matching estimator with regression adjustment proposed by Lechner, Miquel, and Wunsch (2011) and generalized by Lechner and Wunsch (2009) that uses a triangular kernel to weight observations (see online appendix E for the matching protocol). Selection correction is based on the predicted selection probability that is estimated from a probit model (see online appendix C for the exact specification and the estimated coefficients). The unobserved counterfactual average probability of the RPI to find employment within a half-month after \( t \) half-months of unemployment is estimated for \( t \in \{6, 12, 18, 24, 30, 36, 42, 48, 54\} \) half-months. Thus, I do not impose any a priori restrictions on the functional form of the duration dependence. It turns out that the time dependence of the exit rate to employment can be approximated well by the function \( \pi_{t+1} = (1 - \delta)\pi_t \), where \( \delta = 0.026 \) per half-month (see online appendix C). The resulting dependence of the exit rate to employment on unemployment duration is displayed in panel b of figure 1 for different levels of initial search effectiveness. The chosen levels of initial search effectiveness correspond to the mean exit rates in the first month of unemployment for the upper, middle, and lower third of the distribution of \( p \) in the sample, which is obtained in the first step described above, thus illustrating the value of this exercise.

In the simulation, the decline of the exit rates is implemented as being stochastic. The parameterized depreciation rate of \( \delta = 0.026 \) is converted into the probability \( p_0 \) that the exit rate falls by one grid point in the next period: \( (1 - \delta)\pi(p) = p_0(1 - \gamma)\pi(p) + (1 - p_0)\pi(p) = (1 - p_0\gamma)\pi(p) \). Thus, the stochastic depreciation probability is \( p_0 = \delta / \gamma = 0.026 / 0.1 = 0.26 \).

**Job search assistance.** To simplify the numerical analysis, the potential effect of JA on the exit rate to employment (and correspondingly on search effectiveness \( p \)) is fixed to one grid point. The overall effectiveness of JA is calibrated to the actual policy in terms of the implied value for the success probability \( \theta \). Thus, \( \theta \) is the probability that the exit rate to employment increases by one grid point within one period of JA. Let \( \pi^1(p) \) and \( \pi^0(p) \) denote the exit rate to employment of a worker of type \( p \) when, respectively, participating or not participating in JA for one period. Identify and estimate from the data the average effect of JA on the exit rate to employment can be approximated to the actual policy in terms of the implied value for the success probability \( \theta \).

Thus, the implied value of \( \theta \) can be calculated from the estimated effect of JA divided by the value of \( \gamma / (1 - \gamma) + p_0\gamma \pi(p) \), which can be calculated from the estimates of its components. In the estimation of the effect of JA, I allow effect heterogeneity, that is, different values of \( \theta \), for different
types of workers in terms of search skills $p$, as well as different durations of JA. So I allow the first half-month in JA to have a different effect from the second half-month and so on.

When I estimate these effects, two kinds of selection problems arise. First, participants in JA differ systematically from the RPI in ways that are also related to the job finding probability. Second, actual program durations are potentially endogenous. In the RPI, actual durations differ from planned durations by more than 15% in 11% of the cases. Thus, actual program durations cannot be regarded as exogenous. Selectivity in both dimensions becomes evident from online appendix A, which shows the means of all variables in the RPI and subsamples defined by (not) participating in JA of different durations.

Given sufficiently rich data, both selection problems can be solved using a so-called dynamic (or sequential) evaluation approach as suggested by Lechner and Miquel (2010). The idea is to consider a $\tau$-period program as a sequence of $\tau$ one-period programs and to control for selection at the beginning of each of the $\tau$ periods. Selectivity between program participants and the RPI can then be controlled for at the beginning of the sequence, whereas endogeneity of program durations can be accounted for by selection correction in the $\tau - 1$ following periods. The effect of a program of length $\tau$ relative to a program of length $\tau - 1$ can then be obtained by comparing the sequence “participating up to $\tau$” with the sequence “participating up to $\tau - 1$ and not participating in $\tau$.” Selection correction at each stage follows a similar procedure as used for estimating the exit rates to employment for different unemployment durations. The details are therefore relegated to appendix B (justification of the approach) and online appendix D (specification and estimation results of the selection models).

To estimate the effects of different sequences of programs, Lechner (2009) proposes a sequential nearest-neighbor matching estimator where the population of interest is defined within one of the sequences (for example, everyone who participates in the first part of the sequence). I improve this estimator in two dimensions. First, I extend the setup to the case where the population of interest is defined outside the considered sequences as this is the case for the RPI. Second, I incorporate the improvements suggested by Lechner et al. (2011) for a static evaluation approach into the dynamic setting. Hence, this is also the first application of this adapted sequential matching estimator. (For more details, see appendix B.)

Table 2 shows the estimated effects of JA of different durations on the probability of finding employment within three months from the point in time where participation status is defined (for example, after two half-months of JA when interest is in the effect of completing three versus only two half-months of JA). Three months have been chosen because individuals stay in the program for up to three months: A shorter window potentially confounds outcome measurement as some individuals would not be employed by construction. Note that this way of measuring outcomes also takes into account the implicit cost of not being able to exit to employment while in the program, as is assumed in the model. The effects are estimated separately for the entire RPI as well as for subsamples defined by the lower, middle, and upper third of the distribution of $p$ in the RPI.

It turns out that in West Germany in the period 2000 to 2002, JA was essentially ineffective independent of program duration.16 There seems to be some gain in employment from participating in the program for more than three rather than only three half-months, but the first half-month of participation seems to have negative impacts. Thus, the actual policy in that period yields $\theta = 0$. The fixed cost of JA, $k^J$, is 190 euros per half-month, which is calculated from total expenditures per year (excluding benefit payments), the number of participants, and the average program durations from official statistics (BA, 2001–2005). Given this cost, it can be presumed that under these particular circumstances, JA would not be part of the optimal policy. However, in addition to assessing the actual policy relative to the optimal policy, a key objective of this paper is also to gain more general insights into how the optimal use of JA depends on the effectiveness of JA. For this purpose, $\theta$ will also be varied between 0.1 and 1, which covers the range of positive empirical estimates of the effectiveness of JA (see the survey by Thomsen, 2009).

### Job search monitoring

To parameterize the cost of tight monitoring, the average gross salary of a caseworker per

---

15 An alternative way of estimating the effect of programs of different durations would be to apply the methodology of Hirano and Imbens (2004) for continuous treatments (see, for example, Flores-Lagunes, Gonzalez, & Neumann, 2012). However, this method assumes that selection into programs of different durations is fully determined at the beginning of the program. That is, different program durations are not the outcome of intermediate decisions after the program started.

16 This is consistent with the findings of Wunsch and Lechner (2008), who analyze the effectiveness of different forms of job search assistance using the same database but a static evaluation approach.
half-month (about 1,200 euros, according to BA, 2001–2005) is divided by the number of job seekers a caseworker can reasonably take care of in a half-month. A conservative number would be 20, which yields a value of 60 euros. Allowing for some administrative cost, the value for \( \kappa^{JM} \) is set to 100 euros.

**Actual policy in West Germany, 2000–2002.** In Germany, unemployment insurance is mandatory, and employees who have contributed for at least 12 months within the 3 years before entering unemployment are eligible for unemployment benefits (UB) if they register with the public employment service (PES). The minimum UB entitlement is 6 months. In the period under consideration, maximum benefit duration increased stepwise with the total contribution time in the 7 years preceding unemployment, as well as age, up to a maximum of 32 months at age 57 or above with previous contributions of at least 64 months. Since 1994, the replacement rate is 67% (60%) of previous average net earnings from insured employment with (without) dependent children. The median worker in the RPI received 345 euros per half-month for 24 half-months. After exhaustion of UB, workers received so-called unemployment assistance (UA) subject to a means test (until 2005). The replacement rate was lower than that of UB: 57% and 53% with and without dependent children, respectively. The median worker in the RPI received 305 euros per half-month for the rest of the unemployment spell.

Actual payment of benefits is conditional on active job search, regular show-up at the PES, and participation in labor market programs. In case of noncompliance with benefit conditions, benefits could be reduced or suspended temporarily or withdrawn after repeated noncompliance. However, in the period under consideration, search activities were essentially not monitored due to capacity constraints within the PES. I therefore parameterize the fraction of workers subject to JM under the actual policy to \( \theta \) for all periods.

In Germany, job search assistance has become by far the most important activation measure in terms of the number of participants (1.1 million during the period 2000 to 2002). Expenditures are moderate (on average, about 570 euros per participant) because durations are short (up to 3 months but usually about 1.5 months). German job search assistance combines some counseling with a substantial training component, as well as some availability checks. The training component focuses on how to locate job vacancies, fill out job applications, and prepare for job interviews. The courses are usually full time, which fits very well to how job search assistance is modeled in the theoretical part. Of the RPI in the data, about 1% participates in JA every period, but this fraction declines gradually with unemployment duration.

### D. Results

**Optimal choice of policy instruments as a function of \( U, p, \) and \( \theta \).** Figure 2 displays the optimal choice between policy instruments SA, UI, JM, and JA in the \((U, p)\) space for different exemplary levels of the effectiveness parameter of JA, \( \theta \). The latter is varied between 0.1 and 1. The implied effects on the exit rate to employment cover the typical range of positive empirical findings (Fay, 1996; Heckman et al., 1999; Martin & Grubb, 2001; Dolton & O’Neill, 2002; Kluve & Schmidt, 2002; Blundell et al., 2004; OECD, 2005; Thomsen, 2009). The full set of graphs for \( \theta \in \{0.1, 0.2, … , 1\} \) can be found in figure 3 in online appendix F. Moreover, figure 5 in online appendix G shows that the results are not very sensitive to relatively large changes in the fixed cost of JA: \( \kappa^{JA} \in [100, 300] \).

Absorbing SA is used when \( p \) and, hence, the job finding probability, \( \pi(p) \), is too low and promised utility is too high to incentivize the worker. Relative to both SA and JM where incentive costs are absent, UI becomes more attractive the higher \( p \) and the lower \( U \) because the incentive cost under UI, \( Uf_{UI} - Uf_{UB} \geq \frac{\kappa^{JM}}{\beta \pi(p)} \), falls as \( p \) increases for given \( U \) and increases in \( U \). For sufficiently high \( \theta \) (> 0.5), both JM and UI increasingly dominate JA the higher \( p \), because the return to paying the fixed cost of JA falls as \( p \) becomes larger since \( p_{max} \) increases closer. In the case of UI, this effect is reinforced by incentive costs falling with increasing \( p \) relative to constant incentive costs under JA \((Uf_{JA} - Uf_{JA} \geq \frac{\kappa^{JM}}{\beta \pi(p)} )\). For given \( p \) and \( \theta > \pi(p) \), which implies lower incentive costs under JA than UI, JA becomes more attractive relative to UI as \( U \) increases because the difference in incentive costs grows with \( U \).

When considering the optimal use of JA as a function of its effectiveness \( \theta \), four cases can be distinguished. They are displayed in the different panels of figure 2. Panel a shows the first case: JA is not optimal for any \( U \) and \( p \) if \( \theta \leq 0.15 \). Clearly, costly JA is not used whenever \( \theta \leq \pi(p) \) because in this case, the incentive cost exceeds that of UI and there is also the fixed cost. Yet even if \( \theta > \pi(p) \), the case for \( p \leq 27 \) if \( \theta = 0.15 \), the fixed cost of JA prevents its use. Thus, \( \theta = 0.15 \) can be interpreted as the minimum effectiveness required for JA to be used in the optimal policy at all. The implied required effect of JA on the exit rate to employment is small (see Fay, 1996, Heckman et al., 1999; Martin & Grubb, 2001; Dolton & O’Neill, 2002; Kluve & Schmidt, 2002; Blundell et al., 2004; OECD, 2005; Thomsen, 2009), implying that existing programs that exhibit positive effects are sufficiently effective to be optimal for some workers.

Panel b of figure 2 is an example for the second case. When \( 0.15 < \theta < 0.5 \), JA is used only if \( p \) and, hence, the returns to search are very low and if promised utility is not too large. The reason is that even at relatively low success rates and returns of JA, it is worthwhile to prevent the use of very costly and absorbing SA where workers would no longer search for a job.

For \( \theta = 0.5 \), panel c shows the third case, where JA is also optimal for intermediate values of \( p \). The reason that JA is not optimal for the full range of low values of \( p \) is the following: As can be seen from panel a of figure 1, the exit rate to employment is very flat in \( p \) for low levels of \( p \). Hence, the immediate return to paying the fixed cost of JA is relatively low if \( \theta \) is not high enough. Therefore, in addition to preventing SA at the lowest level of \( p \), JA is optimal only...
for parts of the distribution of $p$ where the impact on the exit rate is larger and the incentive costs of UI are still too high relative to those of JA.

Panel d of figure 2 illustrates the last case, which applies for $\theta \geq 0.6$. Here, as long as $U$ is not too high, JA is optimal for all levels of $p$ up to some threshold that increases with $\theta$. In this case, $\theta$ is sufficiently high for JA to be optimal even in the flat parts of $\pi(p)$. At higher levels of $p$, UI dominates because of relatively low incentive costs combined with limited returns to JA, as discussed above. The highest level of $p$ for which JA is optimal, which is reached for $\theta = 1$, is $p = 24$. This still lies below the median search effectiveness in the data ($p = 25$), implying that JA should be targeted exclusively at workers with low exit rates.

The implications for the optimal sequence of policies are as follows. If JA is not effective enough to be part of the optimal policy as in panel a, the typical sequence of policies is UI followed by JM and then absorbing SA for the upper two-thirds of the distribution of initial search effectiveness. The lower third will start immediately with JM. For sufficiently effective JA, the optimal policy will also start with UI for most unemployed workers as long as the initial promised utility is not too high. Eventually a switch to JA will become optimal. This switch will occur the earlier the lower initial search effectiveness and the higher that $\theta$ is. Then JA will be used for some periods—usually 1 month but no more than 3.5 months—until $p$ has increased sufficiently (see figure 4 in online appendix G). Afterward the optimal policy will alternate between UI and JA as $p$ depreciates during UI and is raised again during JA. Note that in case $\theta$ depends on the frequency or duration of previous use of JA, the program will be used less frequently and after increasing periods of UI the longer the worker stays unemployed. If $\theta$ is relatively low and initial $U$ is relatively high, the optimal policy will start with JM and switch to JA relatively late in the unemployment spell. UI is unlikely to be used because promised utility stays constant during JM and falls only if JA is not successful. If $\theta$ is sufficiently high and $p$ is relatively low, the optimal policy will start with JA and then alternate between UI and JA for the same reasons as discussed above.

Note that in comparison to the optimal policy without JA, as in panel a, the importance of JM in the optimal scheme is reduced drastically once JA is part of the optimal policy. JM is crowded out considerably by both JA and UI. JM is a tool to save incentive costs under UI when the returns to search are low. However, if $\theta$ is sufficiently high the incentive...
Assessing the West German policy, 2000–2002. Panel a of Figure 3 shows the sequence of policies that would have been optimal for West Germany in the period 2000 to 2002, given that the same level of initial utility $U_0(p_0)$ is provided as under the actual policy. Since JM was essentially not used and JA was ineffective in this period, UI and SA are the only policy instruments that appear in the graph. JA was ineffective in this period, UI and SA are the only as under the actual policy. Since JM was essentially not used and JA was ineffective in this period, UI and SA are the only policy instruments that appear in the graph. JM becomes optimal after some time because the incentive costs under UI become higher than the fixed monitoring cost despite falling promised utility during UI. Naturally UI is used for longer periods, while JM is used for shorter periods despite falling promised utility during UI.

As can be seen from panel a of figure 2, absorbing SA is used only when at $p = 1$, where the returns to search are essentially 0. Consequently, for $p_0 > 1$, the optimal policy start with UI but switch to SA once search effectiveness deteriorates to $p_i = 1$. The higher the initial search effectiveness, $p_0$, the longer it will take in the unemployment spell to reach this point. For low ($p_0 = 19$), medium ($p_0 = 25$), and high search effectiveness ($p_0 = 29$), as defined by the mean initial exit rate in the lower, middle, and upper third of the distribution of $p$ in the data, UI will be optimal for unemployment durations of up to 39.5, 49, and 55 months, respectively.

It is important to note, though, that this does not mean that maximum unemployment insurance claim durations should be that long. It just means that it is optimal to condition benefits payments on job search effort for this period, regardless of whether they are unemployment insurance payments or welfare payments financed by tax revenue. In the model, “UI” is just a label for benefits conditional on job search; it has nothing to do with the source of the funding.

For a representative worker with medium reemployment wage and search effectiveness, panel b of figure 3 shows that the German benefit scheme is much too flat and that long-run benefits before the 2005 reform were much too high compared to the optimal scheme. Consequently, incentives receive too little weight relative to insurance. The optimal replacement rate starts at 63% and declines to 24%, when SA becomes optimal. From this point on, the benefit remains constant at a level of about 400 euros per month. Interestingly, this is very close to the level of unemployment benefits II (345 euros), which were introduced in 2005 as the baseline welfare payment that can be received after exhaustion of UI. Thus, the postreform benefit scheme is considerably closer to the optimal policy than the prereform scheme.

Potential budget savings from implementing the optimal policy result from three sources: the waste of expenditures on ineffective JA, post-Ul benefits that are too high, and somewhat higher wage taxes that can be imposed under optimized search incentives (see figure 3b). For low, medium, and high initial search skills, $p_0$, the total expected budget savings amount to 18%, 4%, and 1%, respectively. The savings are higher for lower $p_0$ because SA, which is considerably lower than actual UA, is paid earlier in the unemployment spell the lower $p_0$.

The value of tight monitoring. Panel a of Figure 4 shows that the availability of JM for the optimal policy in West Germany leads to substantial crowding out of UI by JM. For sufficiently high initial search effectiveness ($p_0 > 21$), or about 70% of the persons in the data, the optimal policy starts with UI. As search skills depreciate during unemployment, JM becomes optimal after some time because the incentive costs under UI become higher than the fixed monitoring cost despite falling promised utility during UI. Naturally UI is used for longer periods, while JM is used for shorter periods the higher $p_0$. At low initial levels of $p_0$, JM will be optimal from the beginning of unemployment.

The results are insensitive to whether 0 or a small, positive number is chosen for $\pi(1)$.

$18$ The optimal wage tax falls as $p$ decreases because a larger utility spread $U^* - U^b \geq \frac{\pi(p)}{1+p}$ is needed to incentivize the worker. However, as depreciation of $p$ is implemented stochastically in the simulation, the wage tax will increase as long as $p$ remains constant but promised utility and hence benefits fall during UI.
Panel b of figure 4 displays optimal and actual benefits and wage taxes for a representative worker. As predicted by the model, benefits remain constant once JM is used because full insurance is provided. Moreover, because the information asymmetry is removed under JM, both higher-than-actual benefits can be paid and higher wage taxes can be imposed. The additional budget savings from JM relative to the optimal policy with SA, UI, and JM range from 2% to as much as 28%, as θ varies from 0.3 to 1. For workers with medium initial search effectiveness (p0 = 25), the savings are considerably smaller, ranging from below 1% to at most 4%. The budget savings for a worker with high initial search effectiveness (p0 = 29) are negligible, as they are below 1% for all values of θ.

**Sensitivity to the preference parameters.** To check the sensitivity of the results with respect to the preference parameters of the model, I first vary the utility function of the worker. Rather than log utility, \( u(c) = \frac{c^{\sigma - 1}}{1-\sigma} \) is chosen with intertemporal elasticity of substitution once above and once below 1 (\( \sigma \in [0.8, 1.2] \)). Note that if \( \sigma \) is below (above) 1, the initial utility provided by the actual policy is higher (lower) than in the baseline case, because the same level of consumption implies higher (lower) utility. Moreover, for the same reason a given utility difference can be obtained with a smaller (larger) difference in benefits, implying that benefits must decline more slowly (faster) on failure of the activity when the incentive constraint binds under UI or JA. However, the optimal sequence of policies remains almost unchanged. The only change is for \( \sigma < 1 \), where JM and JA are used slightly earlier in the unemployment spell, that is, for slightly higher \( p \).

Second, the disutility of effort, \( e \), is allowed to vary between work and search. The simulation is repeated for the case where the disutility of effort is assumed to be higher than the disutility of work. The higher value of \( e = 0.93 \) rather than 0.84 in the baseline case is obtained when assuming that the time the worker would like to spend on a search is only half of the time he would spend on work. Higher \( e \) implies higher incentive costs during UI and JA, but also makes work relatively more attractive. In the simulation, these effects seem to offset each other because the optimal choice of policies remains unchanged.

**How to implement the optimal policy.** An important issue is how the optimal policy could be implemented. The key feature of the optimal scheme is that both assignment to different instruments and benefit levels depend on the expected individual reemployment wage and job finding probabilities at each possible unemployment duration. A profiling approach could be used to obtain estimates for these quantities. The first step would be to estimate from historical data how individual characteristics affect \( w_t \) and \( \pi_t \) for any unemployment duration \( t \). Based on the estimated coefficients, the predicted values should be calculated for each individual at

\[ u(1) = 0 \] as under log utility.

---

The results displayed in panel b are for a worker with median reemployment wage and medium initial search effectiveness (\( w = 1,720 \) euros per month, \( p_0 = 25, n(25) = 11\%).

**The value of effective JA.** For an initial promised utility as implied by the actual West German policy during the period 2000 to 2002, the required minimum effectiveness of JA is \( \theta = 0.22 \), which corresponds to a relatively low expected effect on the exit rate of about 0.03 to 0.5 percentage points. If sufficiently effective, JA would be used as described in section IIID. The budget implications of the availability of effective JA for the optimal policy in West Germany depend on both the initial search effectiveness, \( p_0 \), and the success probability, \( \theta \). Naturally the budget savings are larger the larger \( \theta \). Moreover, as costly SA without tax revenue is the dominant policy in the absence of JA early in the unemployment spell when \( p_0 \) is small and since JA can be used to prevent this, they are also higher the lower \( p_0 \). Consequently, for workers with low initial search effectiveness (\( p_0 = 19 \)), the additional budget savings relative to the optimal policy with SA, UI, and JM range from 2% to as much as 28%, as \( \theta \) varies from 0.3 to 1. For workers with medium initial search effectiveness (\( p_0 = 25 \)), the savings are considerably smaller, ranging from below 1% to at most 4%. The budget savings for a worker with high initial search effectiveness (\( p_0 = 29 \)) are negligible, as they are below 1% for all values of \( \theta \).
the beginning of unemployment and should be updated whenever individual characteristics (like health or family situation) change over time.

A potential problem is that the discrimination in benefit levels across individuals and time required under the optimal policy might be perceived as unjust. Ex ante heterogeneity in benefit levels as a function of past wages and (un)employment experience is common practice in most industrialized OECD countries, making this less of an issue. However, in contrast to existing policies, benefit claims should be defined in ex ante expected value terms rather than as a fixed amount per period. Conditional on the predicted individual reemployment wage and job finding rates, they should then be distributed differently over time during unemployment for different types of workers.

IV. Conclusion

One of the key challenges when studying optimal policy is how to make quantitative statements about the optimal values of the policy parameters of interest. This is particularly difficult in complex settings that take into account a variety of policy instruments and dynamics. Structural estimation of all necessary parameters in rich theoretical models is usually infeasible without overly strong assumptions that are hard to justify. The virtue of sufficient-statistic approaches is that fewer parameters have to be quantified and that in the best case, state-of-the-art reduced-form estimates can be used to quantify the remaining parameters. However, in complex settings like the one proposed by Pavoni and Violante (2005, 2007), it is extremely hard, if not impossible, to derive a sufficient statistic for every policy parameter of interest that is a function of quantities that can be estimated without assumptions that are similarly strong as those required for structural models.

My paper demonstrates one way to estimate key parameters of the theoretical model studied by Pavoni and Violante (2005, 2007) with relatively few restrictions and assumptions. One advantage of the approach is that it can also be used in more complex settings with more policy instruments, more worker heterogeneity, and more dynamics (see Wunsch, 2007, for an example). It may also be combined with more structural estimation approaches. This paper deliberately focuses on a limited set of policy instruments and only one source of heterogeneity and dynamics. The objective is to illustrate the value of the approach as well as the importance of the framework developed by Pavoni and Violante (2005, 2007) in terms of potential applications beyond the cases analyzed there.

By studying the optimal use of job search assistance as a function of its effectiveness in raising exit rates to employment, it is, on the one hand, possible to derive the minimum effectiveness that is required for the program to be used at all in the optimal policy. This goes beyond traditional cost-effectiveness because there may be other less costly or more effective instruments. Comparing this with empirical estimates of program effectiveness allows for a much more comprehensive evaluation than standard cost-benefit analyses. On the other hand, given that the performance of these programs often varies over time and for different types of job search assistance, it allows for drawing more general conclusions about the role of this program in optimal welfare-to-work programs. By repeating this type of analysis for a variety of activation measures and by jointly studying the full set of labor market policies for the unemployed based on realistic estimates of the model parameters, it will be possible to obtain a comprehensive picture of the optimal design of welfare-to-work programs that allows for deriving specific recommendations for policymakers.

REFERENCES


APPENDIX A

Proof

Pavoni and Violante (2007) show that with constant reemployment wages and in the absence of negative duration dependence in the exit rate to employment JM and SA are absorbing once they are optimal at some \((U, p)\). Thus, JA will never be used thereafter. Now consider the case where UI is optimal at \((U, p)\). Pavoni and Violante (2007) show that in the absence of JA and SA, UI will replace JA after UI. If \(p\) remains constant during UI, it is sufficient to consider the dynamics with respect to \(U\). The following lemma establishes that whenever \(w\) and \(p\) do not depreciate and UI is optimal at \((U, p)\), \(V^{U}\) is more negatively sloped than \(V^{J}\) if \(\theta(p) \leq \pi(p)\) and, hence, JA will never be optimal in the next period.

Lemma 1: For every \((U, p)\) at which UI is optimal and \(\theta \leq \pi\), we have that \(V^{U}(U, p) \geq V^{J}(U, p)\) if \(V\) is submodular and \(w\) and \(p\) do not depreciate.

Proof. From the first-order and envelope conditions for UI and JA, we have

\[ V^{U}(U, p) = -\frac{1}{u'(c_{U})} = \pi(p) V_{U}(U_{J}^{U}, p) + (1 - \pi(p)) V_{U}(U_{J}^{U}, p) \]

or

\[ V^{J}(U, p) = -\frac{1}{u'(c_{J})} = \theta(p) V_{U}(U_{J}^{U}, p) + (1 - \theta(p)) V_{U}(U_{J}^{U}, p). \]

where \(c_{U}\), \(U_{J}^{U}\), and \(U_{J}^{U}\) are optimal consumption and continuation utilities under UI, and \(c_{J}\), \(U_{J}^{U}\), and \(V_{U}(U_{J}^{U}, p)\) are the respective values under JA. If \(\theta(p) \leq \pi(p)\), the envelope condition and the concavity of \(u\) imply the desired result.
So consider the case, where \( \theta \leq \pi(p) \) but \( e^{U,J} > e^{U,A} \). In this case, we have that

\[
- \frac{1}{u'(c_{UJ})} = \theta V_{UJ}(U_{UJ}, p) + (1 - \theta) V_{UJ}(U_{UJ}, p) \\
\leq \theta V_{UJ}(U_{UJ}, p) + (1 - \theta) V_{UJ}(U_{UJ}, p) \tag{A3}
\]

\[
\leq \theta V_{UJ}(U_{UJ}, p) + (1 - \theta) V_{UJ}(U_{UJ}, p) = \frac{1}{u'(c_{UJ})} \tag{A5}
\]

\[
\leq \theta V_{UJ}(U_{UJ}, p) + (1 - \theta) V_{UJ}(U_{UJ}, p) \tag{A6}
\]

Inequality (A4) follows from \( \theta \leq \pi(p) \), \( V_{UJ} < 0 \) and the concavity of \( V \), while inequality (A5) is implied by \( e^{UJ} > e^{UA} \). Inequality (A6) follows from submodularity, which implies that \( V_{UJ}(U, p) \leq V_{UJ}(U, p) \). Pavoni and Violante (2007) show that the IC constraint binds when UI is optimal, that is, \( U = \alpha e^{UJ} + \beta U_{UJ} \). Satisfying the IC constraint under JA, \( U \geq \alpha e^{UJ} + \beta U_{UJ} \) with the same \( U \) requires \( U_{UJ} > U_{JA} \) if \( e^{UJ} > e^{UJA} \). Consequently, a necessary (but not sufficient) condition for

\[
\theta V_{UJ}(U_{UJ}, p) + (1 - \theta) V_{UJ}(U_{UJ}, p) \leq \theta V_{UJ}(U_{UJ}, p) + (1 - \theta) V_{UJ}(U_{UJ}, p) \tag{A7}
\]

to be true is that \( U_{UJ} < U_{UA} \) since \( V_{UJ}(U_{UJ}, p) \leq V_{UJ}(U_{UJ}, p) \) because \( V \) is concave and \( U_{UA} > U_{JA} \). However, the PK constraints and the IC constraint being bound under UI imply that

\[
U_{JA} - U_{UJ} \geq \frac{\epsilon}{\beta \pi(p)} \geq \frac{\epsilon}{\beta \pi(p)} \Rightarrow U_{JA} - U_{UJ} \geq U_{UJ} - U_{UJ} > 0, \tag{A8}
\]

which implies the contradiction that \( U_{JA} > U_{UJ} \). Thus, when \( \theta \leq \pi(p) \), \( e^{UJ} > e^{UJA} \) cannot hold implying that \( e^{UJ} \leq e^{UJA} \) must hold.

Since search effectiveness does not change during UI while promised utility falls, this completes the proof of the proposition.

**Appendix B**

**Identification and Estimation of the Effects of JA**

The population for which the effects of JA are to be estimated is an entry sample into unemployment that is defined independent of any program participation later in the unemployment spell. Lechner and Miquel (2010) show that in this case, the effects of interest are identified under the so-called weak dynamic conditional independence assumption (W-DCIA). This assumption states that, first, conditional on confounding variables at \( t = 0 \), potential outcomes measured from time \( t \geq 1 \) onward are independent of program participation in period \( t = 1 \). This would account for the endogeneity of program durations. The third part of the W-DCIA is a common-support requirement that demands overlap in the control variables between the populations involved in each of the selection steps.

To judge whether W-DCIA is plausible in this particular application, the confounding variables for program participation relative to the RPI have to be identified. This comparison is mainly driven by the difference between participants in JA and nonparticipants. In addition, the variables that jointly influence the job finding probability and the changes in treatment status (i.e., the decision between staying in JA for another period and leaving) have to be detected.

Selection into programs from the RPI is driven by program eligibility, selection by caseworkers, and self-selection by the unemployed. By construction of the sample, all unemployed workers are eligible because they receive unemployment insurance payments. Caseworkers select on the basis of an assessment of the employment prospects and the specific qualification needs of the unemployed. According to the German legislation, they also have to take into account the chances of successful completion of the program and the local labor market conditions. Similar arguments apply to self-selection by the unemployed because they also compare their employment prospects with and without a program, as well as the corresponding costs in terms of effort or potentially forgone benefits in case of refusal to participate. Thus, selection is basically driven by the same factors as the ones affecting the exit rate to employment discussed in detail above.

Decisions to leave or stay in the program are driven by factors that change after entering a program. The most important factors are probably the arrival of job offers, exhaustion of benefit claims, and significant changes in health conditions. Other factors may be noncompliance with the benefit conditions, changes in family status, moving to another place, and take-up or loss of a minor job, which, including exhaustion of benefits, are directly observed in the data. Changes in health conditions are observed if they are severe enough to affect unemployment insurance status. The arrival of job offers is not directly observed in the data, but the number of placement propositions beyond the ES is known per spell. To approximate the arrival of job offers at or up to a specific point in time, a Heckman (1979)-type selection model for the log number of placement propositions per day, that accounts for zero proposi-

In summary, most of the potentially confounding factors are directly observed in the data. Moreover, those that are not directly observed either can be controlled for indirectly by information on past employment histories and compliance with benefit conditions (ability, motivation, preferences), or they can be approximated by the use of observed variables, like arrival of job offers. Thus, the data are sufficiently rich to capture the main sources of selection bias at different points in time before and during program participation. In fact, this is the first application of a dynamic treatment evaluation approach for estimating the effects of programs of different (endogenous) durations.

Lechner (2009) proposes a sequential nearest-neighbor matching estimator where the matching is based on propensity scores to estimate the effects of different sequences of programs for a population defined within the RPI. Consider the case where the interest lies in estimating the effect of the sequence \( S^0 = (S_1^0, S_2^0) \) compared to \( S^1 = (S_1^1, S_2^1) \) for the population defined by \( S_2^1 \). In the first step, the population defined by \( S_1^1 \) is matched to the population of interest, \( S_1^0 \), based on the estimated propensity score of the corresponding selection equation. Then the population defined by \( S_1^1 \) is matched to the observations in \( S_1^0 \) that served as matches in the previous step based on the propensity scores from this and the first selection step. In a similar vein, the population defined by \( S_2^1 \) is matched to the one defined by \( S_2^0 \) based on the corresponding propensity score. To obtain an estimate of the effect of interest, the reweighed outcome of the population defined by \( S_2^1 \) is then subtracted from the reweighed outcome of the population defined by \( S_2^0 \).

Here, a modified version of this estimator is used. First, the effects are estimated for a population defined outside the sequences under consideration. This implies that an additional matching step has to be performed in each comparison. The populations defined by the first element of each sequence under consideration have to be matched to the RPI before matching with each of the sequences is performed (for the exact specifications of the selection models in each step and the corresponding results see see appendix D in the online appendix). Second, radius matching as proposed by Lechner et al. (2011) for static evaluation problems rather than nearest-neighbor matching is used to increase efficiency and potentially robustness given that the RPI is potentially large compared to the populations defined by the last element of the sequences under consideration. Huber, Lechner, and Wunsch (2013) show in a recent simulation study that this estimator belongs to the ones performing best when conditioning on a large set of covariates is necessary.

20 To provide additional work incentives, benefit recipients can earn additional labor income without losing their claim if they work fewer than fifteen hours per week.

21 A placement proposition is a job vacancy proposed to the job seeker by the caseworker.

22 Another recent paper that uses this approach but focuses on the effects of sequences of different programs, and the timing of program participation is Lechner and Wieher (2013).

23 See Lechner and Wunsch (2009) for the basic idea of estimating effects for populations defined independently of treatment status.

24 If the comparison population is small relative to the population for which the effects are estimated, using more than one similar observation prevents one particular observation from getting too much weight.