

WHAT IS THE VALUE ADDED BY CASEWORKERS?

Michael Lechner⁺, Jeffrey A. Smith^{++*}

⁺ *University of St. Gallen
Swiss Institute for International Economics and Applied Economic Research (SIAW),
CEPR, IZA and ZEW*

⁺⁺ *University of Maryland, NBER, IZA and PSI*

First version: February 2000

Current, revised version: September 6, 2004

Date this version has been printed: 06 September 2004

Comments welcome

Addresses for correspondence

Michael Lechner

Swiss Institute for International Economics and Applied Economic Research (SIAW)

University of St. Gallen

Dufourstr. 48, CH-9000 St. Gallen, Switzerland

Michael.Lechner@unisg.ch, www.siaw.unisg.ch/lechner

Jeffrey Smith

Department of Economics

University of Maryland

3105 Tydings Hall

College Park, MD 20742-7211 USA

smith@econ.umd.edu, www.glue.umd.edu/~jsmithz

* Financial support from the Swiss National Science Foundation (NFP 12-53735.18) is gratefully acknowledged by Michael Lechner. Financial support from the Social Science and Humanities Research Council of Canada and from the CIBC Chair in Human Capital and Productivity at the University of Western Ontario is gratefully acknowledged by Jeffrey Smith. The data are a subsample from a database generated for the evaluation of the Swiss active labour market policy together with Michael Gerfin. We are grateful to the Department of Economics of the Swiss Government (*seco*; *Arbeitsmarktstatistik*) for providing the data and to Michael Gerfin for his help in preparing them. We thank Tom Crossley, David Margolis and seminar participants at the University of Maryland and the U.S. Bureau of the Census for helpful comments.

Abstract

This paper contributes to the small (but growing) literature on the efficacy of caseworkers in allocating individuals to government programs and in allocating participants to particular services within those programs. We investigate caseworker allocation of unemployed individuals to different subprograms within Swiss active labour market policy in 1998. Our analysis compares the caseworker allocation to alternatives including random assignment to services and allocation via statistical treatment rules based on observable participant characteristics. Using unusually informative data from administrative unemployment and social security records, we find that Swiss caseworkers obtain roughly the same post-program employment rate as random allocation to services, while statistical treatment rules, even when subject to capacity constraints, do substantially better. Our findings suggest one of three possible conclusions. First, caseworkers may be trying to solve the problem of allocating the unemployed to maximize their subsequent employment, but may lack the skills or knowledge to do this. Second, caseworkers may have a goal other than efficiency, such as allocating the most expensive services to the least well-off clients, that is not explicit in the law regulating active labour market policies. Third, the distortions of the local decision process could be due to federal authorities imposing strict minimum participation requirements for the programs at the regional level.

Keywords

Targeting, Statistical Profiling, Statistical Treatment Rule, Active Labour Market Policy, Caseworkers

JEL classification: J68, H00

1. Introduction

This paper considers the problem of how best to assign unemployed persons to one of a set of available employment and training programs. Several different methods exist to do this. The most common one consists of having the unemployed person meet with a caseworker. Together, the unemployed person and the caseworker come to an agreement about the services that the person should receive based on the person's interests, the caseworker's evaluation of his or her capabilities and the availability of slots in particular programs in the local area. Caseworker allocation is based on the idea that optimal assignment requires knowledge of the characteristics of the unemployed person, the local labour market and local service providers, combined with the presumed professional expertise of the caseworker.

Three other allocation schemes have also been used in practice. The first scheme consists of random assignment to services, a practice typically confined to experimental evaluations.¹ The second scheme consists of deterministic assignment, in which everyone in a particular status gets the same service. For example, everyone on social assistance might be required to receive job search assistance.

The third allocation scheme consists of using statistical treatment rules to assign persons to services (or to any service). This scheme is sometimes called profiling or targeting. It is presently used to assign Unemployment Insurance (UI) claimants in the United States to mandatory employment and training services.² In this context, the profiling is based on a statistical prediction of each claimant's probability of benefit exhaustion or expected benefit receipt duration. Claimants with higher predicted probabilities of exhaustion (or longer expected

¹ For example, in the Canadian Self-Sufficiency Project experiment, treated persons were randomly assigned to receive only a wage subsidy or both a wage subsidy and employment and training services.

² See, e.g., Black, Smith, Berger and Noel (2003), Eberts, O'Leary and Wandner (2002) or Manski (2004).

durations of benefit receipt) must receive services to continue receiving benefits. As discussed in Berger, Black and Smith (2000), this scheme assigns treatment based on the predicted outcome in the absence of treatment, rather than on the predicted impact of the treatment, where the impact is the *difference* between the outcomes the individual would receive with and without treatment. Assignment on the basis of poor predicted outcomes rather than of impacts may serve equity goals (such as allocating the least employable among the unemployed to the most intensive services), but does not serve efficiency goals unless outcomes correlate negatively with impacts.

In this paper, we consider the use of statistical treatment rules to assign treatments on the basis of their predicted impacts. In particular, we use data on the Active Labour Market Policies (ALMPs) in place in Switzerland following their UI reform in 1996 to examine the relative performance of alternative allocation rules. We employ these Swiss data for four reasons. First, the Swiss ALMPs include a wide variety of different treatments, of which we consider eight here. This variety allows substantial scope for caseworker discretion in treatment assignment. Second, the highly decentralized nature of the Swiss government means that caseworkers typically have substantial discretion to use their professional expertise in assigning persons to services. Third, the rich data available in the Swiss context give credibility to the non-experimental estimation methods we use to generate our impact estimates. Finally, the Swiss programs are sufficiently similar in terms of design and services offered to those of other developed countries to make it credible to generalize our findings beyond the Swiss border.

We find that Swiss caseworkers achieve about the same employment rate one year after program initiation as would result from assigning the unemployed randomly to the available treatments in their existing proportions. Furthermore, even if we impose capacity constraints by keeping the overall fraction in each treatment at its current level, our results suggest substantively

important increases in post-program employment rates associated with assignment using statistical treatment rules based on observable characteristics.

The remainder of the paper develops as follows. In Section 2, we describe the policy environment in Switzerland at the time our data were collected. This includes a detailed description of the available employment and training programs. Section 3 describes the existing caseworker assignment mechanism and the basic patterns of assignment to the various treatments. Section 4 outlines our econometric strategy and Section 5 presents our findings on the performance of alternative allocation schemes. Section 6 concludes.

2. The Policy Environment

Switzerland is unique among European countries in its low unemployment rates throughout much of the post-war period. In the 1970s, the Swiss unemployment rate never exceeded one percent, and it did not exceed 1.1 percent in the 1980s. In the 1990s, however, it began to rise to historically high levels, with a peak of 5.2 percent in 1997. This rise in the level of unemployment, though still remarkably low by European standards, prompted the Swiss government to enact a series of unemployment law reforms and to extend their active labour market policies in the 1990s.

Under the 1996 unemployment law reform in Switzerland, which is the one in place at the time our data were generated, individuals may be required to participate in employment and training services once they have been unemployed for 150 days (or 30 weeks) out of their two-year benefit entitlement.³ If they are requested to participate after the deadline and do not

³ The two-year entitlement is available to persons who contributed to the UI system in at least six of the past 24 months. After the two-year entitlement has been exhausted, obtaining a new entitlement requires 12 months of employment within the three years after the previous unemployment spell. The usual replacement rate in the Swiss UI system is 0.70 or 0.80, depending on the recipient's family status.

comply, then their benefits can, and sometimes actually are, cut off. Claimants may participate in services before the deadline, while in other cases the deadline is not enforced by caseworkers.

Table 1 describes the different employment and training services provided under the Swiss UI reform in 1996 (and defines the abbreviations we use for them in the remaining tables). There are three general categories: classroom training of various sorts, work experience in public and private sector jobs that may be created specifically as part of the active labour market policy, and (partial) wage subsidies for temporary regular jobs in the private sector (where the latter may sometimes, but are not supposed to, substitute for permanent regular jobs). The training courses offered under the Swiss ALMP do not include occupational retraining – only further training within the current occupation. Courses last from one day to six months, but only courses at least two weeks in length are counted in our empirical work. Employment programs typically last six months, although participants are required to continue their job search while participating and to accept appropriate offers. Wages on the employment programs can in principle exceed the UI benefit level, but in practice usually do not. Neither courses nor employment programs count toward further UI eligibility. Temporary wage subsidies are not formally a part of the ALMP, but caseworkers appear to treat them as if they were. We follow the caseworkers in doing so here. Local placement offices arrange only about 20 percent of temporary wage subsidy placements, with the remainder arranged through direct contact between employers and employees. The local placement office must confirm placements in the latter category in order for them to receive the subsidy. Time spent employed on a temporary wage subsidy counts toward further UI eligibility.

The general categories of programs offered in Switzerland mirror those available in other developed countries. With the exception of the wage subsidies for temporary jobs, which represent the one unique aspect of the service mix in the Swiss system, Swiss ALMP resembles

that in Germany quite strongly. The New Deal for Young People in the United Kingdom also provides classroom training, subsidized employment and work experience, where the last of these corresponds to the New Deal's Voluntary Sector and Environmental Task Force options. The Swiss services also resemble those provided as Employment Benefits and Support Measures to unemployed persons in Canada. They are somewhat less similar to the service structure of the U.S. Workforce Investment Act program, given the emphasis in the latter on services related to job search, at least as a first step.

3. Data

Our data consist of administrative records on all persons who were registered unemployed in Switzerland as of December 31, 1997. Our analysis sample consists of the subsample of this population that results from imposing a number of exclusion criteria. In particular, we keep only unemployed persons with the following characteristics: age between 25 and 55 (inclusive), not disabled, at least 100 Swiss Francs of past earnings, valid value of mother tongue variable, Swiss citizen or foreigner with annual or permanent work permit, not working at home, not a student, not an apprentice, unemployed less than one year, no program duration longer than 14 days in 1997, no employment program (at all) in 1997, and no program start on January 1, 1998 (such a start date implies a continuing program). The analysis sample includes over 19,000 persons.⁴

We code the first major spell of program participation starting after January 1, 1998, where we define a major spell of participation as one lasting at least 14 days. We code persons not participating in any single program for more than two weeks between January 1, 1998 and January 1, 1999 as non-participants. In order to code time-varying variables defined relative to the program start date for the non-participants, we need to assign each one a fictitious start date.

We do this by drawing at random from the empirical distribution of start dates among participants. Non-participants whose assigned start date occurs after the end of their unemployment spell are dropped from the sample.⁵ We condition on the elapsed time between the start of the unemployment spell and the start date (real or assigned) in our empirical work

In coding service receipt, we have to deal with the familiar problem that participants often participate in more than one program in a given unemployment spell. As in other countries, these additional programs sometimes represent part of a planned sequence but often represent an endogenous response to a poor match between the claimant and the initial program. In our data, about 30 percent of those participating in at least one program also participated in another; however, for the majority of these, the second program was of the same type (in the typology of Table 1). As such, we follow Gerfin and Lechner (2002) by coding persons based on the first program they participate in for more than two weeks during a given unemployment spell.

4. The Caseworker Allocation

Currently, Swiss ALMPs rely on caseworkers to assign unemployed persons to employment and training services. In the Swiss system, each caseworker handles 75 to 150 clients and has an in-depth interview with each client every month. This represents substantially more personal contact than participants would receive in most other developed countries. It also means that Swiss caseworkers have the opportunity to gain a large amount of information about the claimant's needs and abilities, information that, in principle, they should be able to use in effectively matching claimants to services.

⁴ See Appendix A.2 of Gerfin and Lechner (2002) for more detail about the sample definition.

⁵ See Lechner (1999) for an extended discussion of the strengths and weaknesses of this method of handling the temporal alignment of non-participants. See Fredricksson and Johansson (2003), Abbring and van den Berg (2003) and Sianesi (2004) for more general discussions and alternative methods for addressing this issue.

Table 2 presents information on the allocation chosen by the caseworkers. The first column of Table 2 shows the number of sample observations in each service type, followed in the second column by the sample proportion.⁶ It reveals temporary wage subsidies as the most common service, followed by language courses. The predominance of the latter reflects the over-representation of foreigners among the Swiss unemployed. The third column indicates the mean duration of the program for persons receiving each service. In general, employment-related services tend to last longer than classroom-based services. The fourth and fifth columns indicate the mean days of unemployment prior to the start of services and the fraction of persons for whom the services started prior to the 150-day deadline. The sixth column indicates the mean qualification of persons receiving each service type, with qualifications measured on a scale from one (skilled) to three (unskilled). Perhaps not surprisingly, participants in language courses have the lowest mean level of qualifications, while participants in computer courses have the highest. The opposite pattern holds in the sixth column, which indicates the percentage of foreigners in each service type. The highest percentage is now found for language courses, and the lowest for computer courses.

The final column in Table 2 gives employment rates one year after the program begins; this measure serves as our outcome variable. The highest employment rate corresponds to temporary wage subsidies and the lowest to private employment programs. Of course, these employment rates reflect a combination of non-random assignment to services based on employment-related characteristics and the impact of the services themselves.

⁶ The “other programmes” include a heterogeneous set of services. The primary ones consist of “practice firms” to prepare the unemployed for self-employment (30%) or for jobs in the health (17%) or tourism (17%) sectors. In these pretend firms, participants perform business functions including buying and selling with other practice firms, doing the books and so on. The “other programmes” category also includes practical courses for young unemployed (7%), in which the youth try out a particular occupation for a few weeks, along with a heterogeneous group of services whose content is not specified in greater detail (30%).

We draw three main lessons from Table 2. First, Swiss caseworkers are making use of the flexibility available to them to assign unemployed persons in large numbers to all of the treatment types we consider here. Second, the caseworkers do not allocate persons at random with respect to their observed characteristics. Mean unemployment durations, mean qualifications and percent foreign all differ among the service types. Assuming that most services have only modest impacts (consistent with the survey in Heckman, LaLonde and Smith, 1999, and with Gerfin and Lechner, 2002, for Switzerland), we also observe strong differences in mean employment chances in the absence of treatment. Third, the caseworker allocation shows evidence of systematic, reasonable patterns. It makes sense to assign foreigners to language courses and the most qualified among the unemployed to computer courses.

5. Econometric Strategy

We seek to compare alternative strategies for allocating unemployed workers to services (including non-participation) within Swiss ALMP. This requires estimates of the potential outcome associated with each alternative service for each of the unemployed individuals in our data. We observe one potential outcome, that corresponding to the option assigned by the caseworker, for each individual. To identify the other eight potential outcomes, we rely on our unusually informative data and adopt a “selection on observables” approach. That is, we assume that, conditional on the rich set of variables in our dataset, the potential outcomes associated with each alternative service are independent of the particular service assigned by the caseworker. In formal terms, we rely on a Conditional Independence Assumption (CIA) for the case of multiple programs, as outlined in Imbens (2000) and Lechner (2001).

This is a strong assumption, which requires that our data include all of the variables that affect both (not either, but both) service assignment and labour market outcomes. Though very

strong the CIA is plausible given the richness of our data. The data include a wealth of demographic variables (age, sex, first language, other languages, marital status, nationality), variables related to pre-program skill levels (skill level assessed by the caseworker, past program participation), characteristics of the last job (earnings, skill level, occupation and industry, size of town), variables from social security data related to past employment, variables related to current and past UI receipt (duration of spell up to participation, remaining eligibility, number of previous spells, length of preceding spell, past sanction days), variables related to job search intensity (mobility, seeking part-time or full-time work, presently employed part-time, canton of residence) and, perhaps most importantly, the individual's chance of finding a job *as assessed by the caseworker* at the time of their first meeting (very easy, easy, medium, difficult and special case). This latter variable, in particular, should summarize otherwise unobservable factors related to the individual's motivation, presentation, social skills and so on.⁷ Lechner and Smith (2003, Table 3) show that these variables strongly predict service assignment.

We obtain our predicted potential outcomes using two alternative estimators, each based on the multi-treatment version of the CIA. The first estimator is the semi-parametric multi-treatment matching estimator implemented using these data in Gerfin and Lechner (2002). Their estimator relies on non-parametric single nearest neighbour matching with replacement to construct estimates of the impact of receiving one treatment rather than another. The matching variables include a balancing score estimated using a multinomial probit model that includes the covariates described in the preceding paragraph, and a handful of variables of particular importance included both directly in the matching and in the balancing score.⁸ We borrow their

⁷ A complete list of variables appears in Appendix Table A.1 in Lechner and Smith (2003). For an even more extensive statement of the case for the CIA in this context see Gerfin and Lechner (2002).

⁸ The variables included in the matching both directly and indirectly through the balancing score are a dummy for a non-Swiss native language, sex, the calendar date of program start and the duration of the unemployment spell

estimates of the estimated employment rate resulting from assigning all of the unemployed to one of the nine treatments we examine: non-participation and the eight services listed in Table 1.⁹

Consideration of assignment rules other than assigning everyone to receive one service requires the estimation of the labour market outcomes associated with each of the nine service alternatives (including non-participation). The matching estimator as used in Gerfin and Lechner (2002), though preferable in general because it is semi-parametric, does not estimate such person-specific probabilities with sufficient precision. As a result, we use a flexible parametric estimator that also relies on the CIA for identification to produce the estimated post-program employment rates associated with the remaining allocation schemes.

In particular, we estimate a binary probit model with employment in day 365 after program start as the dependent variable for each of the nine subsamples defined by the observed treatment choice. As conditioning variables in the probits we include the estimated conditional choice probabilities (allowing for functional flexibility by including some of their linear indices as well), as they jointly form balancing scores and thus correct for selection bias, along with the additional variables in the matching protocol employed by Gerfin and Lechner (2002). The specification has been tested against omitted variables and functional misspecification using standard score tests. We also performed specification tests against heteroscedasticity, information matrix tests, and a normality test.¹⁰ These probits allow construction of the conditional probability of employment for each sample member in each treatment and thereby

prior to program start. Including these variables directly in the matching increases, in finite samples, the weight they receive in determining the nearest neighbors used to construct the counterfactuals of interest. They are chosen on the basis of our a priori views about the key variables affecting treatment choice and labor market outcomes in this context.

⁹ Lechner and Smith (2003) and Gerfin and Lechner (2002) present more details regarding the implementation of the multi-treatment matching estimator.

¹⁰ Lack of omitted variables, conditional homoscedasticity and normality of the probit latent error terms are tested using conventional specification tests (Bera, Jarque, and Lee, 1984, Davidson and MacKinnon, 1984, and White, 1982). The information matrix test (IMT) statistics are computed using the second version suggested in Orme (1988), which appears to have good small sample properties.

allow us to predict the overall employment rates associated with all possible allocation schemes involving these treatments.

Five important caveats apply to the econometric approach adopted here. First, we assume that the potential outcomes for each person do not depend on the distribution of the unemployed into the different service alternatives. Put differently, we assume the absence of spillovers, where one individual's treatment choice affects the potential outcomes of other individuals, as well as the absence of scale effects or other general equilibrium phenomena. The formal name for this assumption is the Stable Unit Treatment Value Assumption, or SUTVA; see Rubin (1980). It is common to all partial equilibrium analyses, not just those based on the CIA. This is a strong assumption in our context. For example, assigning all of the unemployed to, say, vocational training, would raise the quantity of labour with certain skills, and thereby depress its price, relative to a situation in which only a modest fraction of the unemployed receive such training. This is one reason (the other being supply constraints) that we consider allocation schemes that reallocate the unemployed among the various alternative services while keeping the proportion of the unemployed assigned to each alternative the same as what we actually observe.

Second, we do not have information on direct costs for the different services, so our results rely on estimates of gross rather than net impacts. Our estimates do (partly) capture differences in indirect cost savings among alternative services due to reductions in the amount of time spent collecting UI benefits. Third, because we condition on functions of X in our employment probits, rather than conditioning more finely on X itself, our results understate the performance of the econometric assignment models. Fourth, because we take the maxima and minima of sets of estimated values to determine assignments with no consideration of the variance of these estimates, sampling variation will lead us to overstate somewhat the

performance of the econometric assignment models; see, e.g., the discussion in Horrace and Schmidt (2000).

Fifth, our outcome variable measures employment on one specific day – the day 365 days after the start of the program. If the different service alternatives imply different times paths of employment probabilities, then our one-day measure may provide a biased guide to the discounted present value of the time spent employed associated with each service (and, likewise, to the discounted present value of earnings which would represent the object of interest in North American ALMP). In light of these caveats, we view our estimates not as definitive statements of expected gains, but rather as suggestive of the improvements that could be achieved by supplementing or replacing caseworker judgement with econometric forecasts.

6. Results

Table 3 presents the employment rates associated with alternative allocations of the unemployed workers in our data to the nine available services (including non-participation) we consider. The table includes employment rates for both the full sample of the unemployed, and for that subsample (about 60 percent) who report as their native language one of the three primary Swiss national languages (German, French or Italian).¹¹ This separate analysis allows us to determine whether caseworkers do better with the unemployed immigrants who make up the non-Swiss language group.

The first two rows of Table 3 present the estimated employment rate given random assignment of the unemployed to the nine service alternatives in their existing proportions, and the observed overall mean employment rate associated with the caseworker allocation. They

¹¹ The fourth official Swiss language, Romansch, is spoken by only a tiny fraction of the population.

show that for both the full sample and the Swiss language sample, the caseworkers do just a bit worse in their allocation than random assignment would do. This finding may seem surprising, particularly to those who have interacted with caseworkers confident of their abilities. However, we note that Swiss caseworkers, like similar caseworkers in most countries, receive very limited feedback with which to update the beliefs underlying their allocation decisions. The third row in Table 3 shows that replacing assignment to the four least effective treatments – as determined by Gerfin and Lechner (2002) - with assignment to non-participation would have a small positive effect on the employment rate.

The next nine rows present the estimated employment rates associated with assigning everyone to each of the nine service alternatives in turn. These simple allocations would presumably save on program administration costs. For five of the service alternatives, assigning everyone to that alternative leads to a lower estimated employment rate than either the current caseworker allocation or random assignment to services in the existing proportions. In contrast, in the remaining four cases – non-participation, vocational training, other training, and temporary wage subsidies – assigning everyone to the service dominates both the caseworker allocation and random assignment. The non-participation case holds special interest, as it represents simply getting rid of the ALMP. It requires zero direct costs, but still dominates all of the one-service-for-all alternatives except for other training and temporary wage subsidies.

The next four rows consider allocations that maximize and minimize the predicted employment rate. These allocations (like the ones that assign all of the unemployed to one particular service) relax the constraint imposed by the existing service proportions. The first of the four allocations assigns each person to that one of the nine alternatives for which he or she has the highest predicted employment probability. The resulting mean post-program employment

rates of 55.5 overall and 61.9 for the Swiss language sub-sample represent large increases over those implied by either random assignment or the observed caseworker allocation. The implied distributions of the unemployed among the various services for this allocation and for the other three allocations in this group appear in Table 4. The allocation that maximizes the predicted employment rate assigns far more of the unemployed to vocational training, other training and temporary wage subsidies than does the observed caseworker allocation, and far fewer to non-participation, basic courses and language courses.

The second of the four allocations resembles the first, only it rules out non-participation as an alternative (and also drops the non-participants from the sample). Not surprisingly, given that the first allocation assigned only 1.6 percent of the unemployed to non-participation, ruling out this option makes little difference to the resulting estimated overall post-program employment rate. These two allocations capture the spirit of the Canadian Service Outcomes and Measurement System (SOMS) described in Colpitts (2002) and the American Frontline Decision Support System (FDSS) described in Eberts, O’Leary and DeRango (2002). These systems sought (in the case of SOMS) or seek (in the case of FDSS) to promote efficiency in allocation through the assignment of individuals based on predicted impacts.

The next pair of allocations turns the previous pair on its head by assigning individuals to that alternative for which they have the lowest predicted probability of employment, with or without non-participation included in the set of available options (and non-participants in the sample). We find that allocating services so as to minimize the post-program employment rate leads to overall rates of 25.7 percent with non-participation as an option and 26.7 percent without non-participation as an option. These figures are far below (over 10 percentage points) the employment rates resulting from either the observed caseworker allocation or random assignment with existing service proportions. Relative to the observed caseworker allocation, the allocation

that minimizes the estimated employment rate assigns more of the unemployed to temporary employment in the public (especially) and private sectors, and to language training. It assigns almost no one to temporary wage subsidies.

The final six assignment schemes in Table 3 impose “supply constraints” at either the national (in the first three rows in this group) or regional (in the second three rows) level. By supply constraints, we just mean that we force the allocation to adopt the observed distribution of services either for the country as a whole or separately for unemployed workers in each region. The cantons included in each region for this purpose appear in the notes to Table 3. The point of imposing these constraints on the allocations we consider is realism; in many cases, there may be no way, particularly in the short to medium term, to substantially increase the number of slots in computer courses, or to substantially increase the number of temporary wage subsidies which, after all, require a willing employer. By considering both cases of unlimited flexibility (with no supply constraints) and no flexibility (where we impose the existing distribution of services) we bracket the true situation, which involves some limited amount of flexibility.

The supply constraints raise the problem that who gets assigned to what now depends on the order in which we consider the unemployed in our data. Those who get assigned first will get their preferred service alternative, but those who get assigned later may find that all the slots for their preferred service have already been filled. We deal with this issue by utilizing the following two schemes to order the sample:

1. *“Effect-based” ordering*: First we calculate for each sample member the estimated mean impact on the probability of post-program employment, relative to non-participation, associated with each service alternative, where some (or all) of these estimated impacts may be negative. We then sort the sample members by the difference between the most positive (or least

negative) impact and the second most positive (or least negative) impact. Assignment to services then proceeds in order by this difference, until one service becomes full. At that point, we reset the estimated impact for the service with no remaining slots to a very large negative number (for purposes of the allocation), and the unassigned observations are re-sorted. Allocation then proceeds based on the resorted order until a second service becomes full, and so on.

2. *“Need-based” ordering*: First we estimate the probability of employment conditional on non-participation for each sample member. Next we sort the sample based on this probability. Then we assign services in order starting with the lowest value of this probability, which we take as a measure of need.¹²

Separate from the ordering scheme is the choice of which service alternative to assign to each person when they come up. We consider two alternatives here: (1) assignment to the alternative with the largest predicted employment rate (denoted the “largest predicted gross impact” in Table 3); and (2) assignment to the alternative with the smallest predicted employment rate (denoted the “smallest predicted gross impact” in Table 3). The first represents a best-case assignment that maximizes, given the available estimates and subject to the indicated supply constraints, the efficiency of service allocation. The second is a worst-case scenario, from an efficiency standpoint, again given the available estimates and subject to the supply constraints.

Now return to the final six assignment schemes in Table 3. The first three represent assignment to the service with the largest gross impact with effect-based ordering, assignment to the service with the smallest gross impact with effect-based ordering and assignment to the service with the largest gross impact with need-based ordering, all with supply constraints

¹² We also considered a “first come, first served” allocation scheme, wherein individuals were assigned in order by their observed (and assigned, for the non-participants) start dates. The results for this scheme were basically the same as for the need-based scheme, so we do not consider them in detail.

imposed at the national level. The next three assignments are the same but with the supply constraints imposed at the regional level.

These six allocations provide several useful lessons. First, comparing the constrained and unconstrained allocations based on gross impacts for the full sample, we see that imposing the national supply constraints makes a large difference, by reducing the estimated post-program employment rate from 55.5 to 49.3. In contrast, imposing the supply constraints at the regional rather than the national level leads to only a small further reduction from 49.3 to 47.2. Thus, supply constraints matter, and without further information regarding the flexibility of the supply of subsidized jobs and training slots, the data leave us with a fairly wide range of potential employment rates associated with service assignment based on estimated impacts.

Second, comparing the estimates based on assignment to the largest and smallest gross impacts (with effect-based ordering) shows that imposing the supply constraints moderates the difference in estimated employment rates between these best and worst cases, relative to that found for the unconstrained case. In addition to the decrease in the employment rate associated with allocation based on the largest predicted impacts, the employment rate associated with allocation based on the smallest predicted impacts increases from 25.7 to 37.0 for the full sample when we impose the supply constraints. The supply constraints strongly limit the number of unemployed allocated to either relatively effective or relatively ineffective services.

Third, the way in which we order the respondents makes very little difference. For the full sample, switching from effect-based ordering to need-based ordering lowers the estimated post-program employment rate from 49.3 to 47.8 with the national supply constraints and from 48.4 to 47.2 for the regional supply constraints. In this case, adding an equity dimension to the allocation has only a small efficiency cost.

In Table 5 we consider the same allocations as in Table 3, but with the estimated employment rates broken down into subgroups based on regional characteristics. The first two columns present results for urban and rural regions as defined by the size of the town the regional employment office (RAV) is located in. The third, fourth and fifth columns present estimates separately for Type I, Type II and Type III RAVs, as defined Atag, Ernst and Young Consulting (1999). These types relate to estimated inflow and outflow rates from unemployment for each office, conditional on local economic conditions. Type I RAVs have low inflow rates and high outflow rates, Type II RAVs have high inflow rates and high outflow rates and Type III RAVs have high inflow rates and low outflow rates. There are no cantons with low inflow rates and low outflow rates. The final two columns break the cantons down by whether their primary language is German or not (in which case it is French or Italian).

The patterns observed for the full sample, and for the Swiss language sample, largely carry over for all of the subgroups in Table 5. We note two additional findings of interest. First, the difference between the employment rates implied by the observed caseworker assignment and by random assignment remains remarkably stable for the various subgroups. It varies between 1.2 (for rural RAVs and Type I RAVs) and 0.0 (for Type II RAVs). Caseworkers do not appear to vary very much on a geographic basis in their ability to allocate the unemployed. Second, the gains from moving from caseworker allocation to unconstrained (or constrained) allocation based on estimated impacts appears noticeably larger for Type I RAVs, and for primarily German-speaking cantons. We do not have a clear explanation for these patterns.

8. Conclusions

Most active labour market policies in the developed world feature a variety of different employment and training services. With a few notable exceptions, such as the WPRS system for the unemployed in the U.S.¹³, individuals seeking help in the labour market get allocated to these services with the assistance of caseworkers.

In this paper, we show, using recent data on the Swiss unemployed, that caseworkers do about as well at allocating clients to services as random assignment with existing service proportions, when performance consists of employment rates measured one year after the start of the program. By examining allocations based on assigning each person to that service with the largest, or smallest gross impact (relative to non-participation), we show that things could either be much better, or much worse. Taking our estimates for the full sample without supply constraints, we estimate that assigning individuals to the service with the largest impact would raise post-program employment rates by 14.0 percentage points. At the other end, deliberately assigning the unemployed to the service with the lowest predicted impact reduces the estimated employment rate for the same group by 15.8 percentage points. Adding supply constraints moderates these figures (in both directions), but still suggests an increase of nearly 7.0 percentage points in the employment rate under regional constraints. Thus, caseworkers do not add much value, but they do not subtract much either, in their role as service allocators.

Our findings generally comport with the (very) small literature that has examined related questions. Using the same data that we do here and the multi-treatment matching methods described in Section 5, Gerfin and Lechner (2002) estimate the post-program employment rates associated with taking the unemployed in each of the nine service alternatives and assigning

them, as a group, to each of the other eight alternatives. They find that, in general, the existing service allocation typically lies in the middle of the possible post-program employment rates. Lechner and Smith (2003) amplify this result by showing that it holds for the subset of individuals in each service alternative with the highest probabilities of assignment to that service; that is, it holds even for individuals with characteristics for which caseworkers are in general agreement about the appropriate service.

Dehejia (2004) finds efficiency gains in a Bayesian analysis of a statistical treatment rule using predicted impacts in the context of California's Greater Avenues to Independence (GAIN) program. O'Leary, Decker and Wandner (2002) find efficiency gains from assigning the UI bonus treatment (a special payment to unemployed workers who find work quickly) only to claimants with long predicted spells. Caliendo, Hujer and Thomsen (2004) find some efficiency gains from using an impact-based statistical treatment rule in the context of assignment to German job creation schemes. Looking beyond ALMPs, Frölich (2003) examines treatment choice in the context of Swedish rehabilitation programs and finds large gains relative to caseworker assignment, while Benitez-Silva, Buchinsky and Rust (2004) find large benefits from using a statistical rule to screen applicants for disability insurance in the US.

Finally, Bell and Orr (2002) report on a study that asked caseworkers which applicants they thought would benefit most from the AFDC Homemaker-Home Health Aide program, which trained welfare mothers to become home health aides. This information was collected prior to the random assignment of applicants. By interacting the experimental treatment indicator with the caseworkers' ratings of potential benefits in the impact analysis, they show that caseworkers have, essentially, no idea who will benefit more or less from the program. This

¹³ Even WPRS represents only a partial example. The system uses a statistical treatment rule to assign the requirement of mandatory employment and training services to a subset of those collecting unemployment

suggests, in turn, that their choices regarding participation are unlikely to do as well as those of a statistical treatment rule based on predicted impacts.

Taken together, the available evidence, including that presented in this paper, does not make a strong case for the abilities of caseworkers at assigning individuals to services within ALMPs. Should the governments fire their caseworkers and replace them with statistical treatment rules? While the evidence presented here (and elsewhere) is suggestive, some important issues remain unresolved. Consider the Swiss context examined here. Swiss caseworkers perform a number of functions in addition to service allocation. These include monitoring the unemployed and encouraging them to look for work or training, networking with employers to develop opportunities for subsidized temporary jobs, keeping abreast of local training opportunities and so on. Our results do not pertain to these other functions, which caseworkers may perform either well or poorly.

In addition, as we have already noted, our analysis has some limitations that flow out of the data we use. First, we lack the cost data necessary to examine allocations based on gross rather than net impacts. Second, our dependent variable consists of employment at a particular point in time, rather than discounted sums of future earnings. Third, our impact estimates rely, of necessity, on non-experimental data. While the methods we employ have credibility in our context due to the wealth of covariate information available, data in which individuals were randomly assigned to services would make our analysis even more compelling.

Fourth, a decision about how to organize the assignment of the unemployed to services requires a full comparison of the benefits *and the costs* of the alternative methods under consideration. Caseworkers cost money, but so do statistical treatment rules. In particular, the latter require data collection, analysis, programming and so on. These are not cheap. At one

point in the late 1990s, the State of Kentucky shut down its WPRS profiling system because it cost less to serve all of its UI claimants than to serve only some and operate the profiling system.

Our findings raise several broader questions, which we note here but whose resolution awaits future work. First, why do caseworkers think they do a good job of allocating individuals to services when it appears they do not? Second, could a system of feedback be developed that would allow them to update their beliefs and to learn to do better? Third, would improved initial training allow the caseworkers to do better? Fourth, would a combination of caseworkers and statistical treatment rules dominate either mechanism on its own? Finally, who benefits when caseworkers fail to maximize the (economic) efficiency of their allocation? Does the failure of casework allocation that we document represent special interests at work, human errors of design, or the outcome of a compromise between many competing policy goals?

References

- Abbring, Jaap and Gerard van den Berg. 2003. "Analyzing the Effect of Dynamically Assigned Treatments Using Duration Models, Binary Treatment Models and Panel Data Models." IZA Working Paper No. 831.
- Atag, Ernst and Young Consulting. 1999. *RAV-Evaluation: Schlussbericht*. Bern.
- Bell, Stephen and Larry Orr. 2002. "Screening (and Creaming?) Applicants to Job Training Programs: The AFDC Homemaker Home Health Aide Demonstrations." *Labour Economics*. 9(2): 279-302.
- Benitez-Silva, Hugo, Moshe Buchinsky and John Rust. 2004. "How Large are the Classification Errors in the Social Security Disability Award Process." NBER Working Paper No. 10219.
- Bera, Anil, Carlos Jarque, and Lung-Fei Lee. 1984. "Testing the Normality Assumption in Limited Dependent Variable Models." *International Economic Review*. 25. 563-578.
- Berger, Mark, Dan Black and Jeffrey Smith. 2000. "Evaluating Profiling as a Means of Allocating Government Services." In Michael Lechner and Friedhelm Pfeiffer (eds.), *Econometric Evaluation of Labour Market Policies*. Heidelberg: Physica. 59-84.
- Black, Dan, Jeffrey Smith, Mark Berger and Brett Noel. 2003. "Is the Threat of Reemployment Services More Effective than the Services Themselves? Evidence from Random Assignment in the UI System." *American Economic Review*. 93(4): 1313-1327.
- Caliendo, Marco, Reinhard Hujer and Stephan Thomsen. 2004. "Effect Heterogeneity, Profiling and Targeting – How Can We Improve the Efficiency of Labour Market Policies?" Unpublished manuscript, J.W. Göthe Universität.
- Colpitts, Terry. 2002. "Targeting Reemployment Services in Canada." In Randall Eberts, Christopher O'Leary and Stephen Wandner (eds), *Targeting Employment Services*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. 283-302.
- Davidson, Russell and James MacKinnon. 1984. "Convenient Specification Tests for Logit and Probit Models." *Journal of Econometrics*. 25. 241-262.
- Dehejia, Rajeev. 2004. "Program Evaluation as a Decision Problem." *Journal of Econometrics*, forthcoming.
- Eberts, Randall, Christopher O'Leary and Stephen Wandner (eds). 2002. *Targeting Employment Services*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- Eberts, Randall, Christopher O'Leary and Kelly DeRango. 2002. "A Frontline Decision Support System for One-Stop Centers." In Randall Eberts, Christopher O'Leary and Stephen Wandner (eds), *Targeting Employment Services*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. 337-380

Fredricksson, Peter and Per Johansson. 2003. "Program Evaluation and Random Program Starts." Uppsala University Department of Economics Working Paper No. 2003:1.

Frölich, Markus. 2003. *Programme Evaluation and Treatment Choice, Lecture Notes in Economics and Mathematical Systems, Vol. 524*. Heidelberg: Springer.

Gerfin, Michael and Michael Lechner. 2002. "Microeconomic Evaluation of Active Labour Market Policy in Switzerland." *Economic Journal*. 112(482): 854-893.

Heckman, James, Robert LaLonde and Jeffrey Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics, Volume 3A*. Amsterdam: North-Holland. 1865-2097.

Horrace, William and Peter Schmidt. 2000. "Multiple Comparisons with the Best, with Economic Applications." *Journal of Applied Econometrics*. 15: 1-26.

Imbens, Guido. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika*. 87. 706-710.

Lechner, Michael. 1999. "Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany After Unification." *Journal of Business and Economic Statistics*. 17: 74-90.

Lechner, Michael. 2001. "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption." In Michael Lechner and Friedhelm Pfeiffer, eds., *Econometric Evaluation of Labour Market Policies*. Heidelberg: Physica/Springer. 43-58.

Lechner, Michael and Jeffrey Smith. 2003. "What is the Value Added by Caseworkers?" IZA Working Paper No. 728.

Manski, Charles. 2004. "Statistical Treatment Rules for Heterogeneous Populations." *Econometrica*. 75(4): 1221-1246.

O'Leary, Christopher, Paul Decker and Stephen Wandner. 2002. "Targeting Reemployment Bonuses." In Randall Eberts, Christopher O'Leary and Stephen Wandner (eds), *Targeting Employment Services*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. 161-182.

Orme, Chris. 1988. "The Calculation of the Information Matrix Test for Binary Data Models." *The Manchester School*. 56. 370-376.

Rubin, Donald. 1980. Discussion of paper by D. Basu. *Journal of the American Statistical Association*. 75. 591-3.

Sianesi, Barbara. 2004. "An Evaluation of the Active Labour Market Programmes in Sweden." *Review of Economics and Statistics*. 86(1): 133-155.

White, Halbert. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica*. 50. 1-25. [And associated "Corrigendum." *Econometrica*. 51. 513]

Table 1: Descriptions of programmes

COURSES	Courses must be necessary and adequate with the goal of improving individual employment chances; duration varies between one day and several months; here a minimum duration of two weeks is required
BASIC COURSES (BAC)	Short courses teach some basic skills not necessarily related to a particular occupation (<i>basic programme, courses to promote self-esteem and personality, courses for acquiring basic skills</i>)
LANGUAGE COURSES (LAC)	Language courses
COMPUTER COURSES (COC)	General computer courses, specific computer courses
FURTHER VOCATIONAL TRAINING (FVT)	Business and trade training (up to the level of a vocational degree), business and trade training (above the level of a vocational degree), manufacturing and technical training (up to the level of a vocational degree), and manufacturing and technical training (above the level of a vocational degree)
OTHER COURSES (OTC)	Practice firms, practical courses for the young unemployed, courses for jobs in the tourism sector, courses for jobs in the health care sector, and other courses
EMPLOYMENT PROGRAMMES	Goal: Work practise. These jobs should be as similar as possible to regular employment, but they should be <i>extraordinary</i> , i.e. employment programmes should not be in competition with other firms. Both public and private institutions offer employment programmes. During an employment programme the unemployed has to continue his job search and must accept any suitable job offer. Employment programmes usually last for six months.
PUBLIC (TE-PU)	Employment programmes within the public sector
PRIVATE (TE-PR)	Employment programmes within the private sector
TEMPORARY WAGE SUBSIDY (TEMP)	The objective of a TEMPORARY WAGE SUBSIDY is to encourage job seekers to accept job offers that pay less than their unemployment benefit by making up the difference with additional payments. The total income generated by this scheme is larger than the unemployment benefit. The TEMPORARY WAGE SUBSIDY scheme does not officially belong to the ALMP but there is compelling evidence that the placement offices intentionally use the subsidies as an active labour market policy instrument.

Note: We consider only participation in programmes of at least two weeks in duration. (NONP) denotes non-participation.

Table 2: Number of observations, selected characteristics, and outcomes of different groups

Group		obs.		duration of pro- gramme	unemployment before participation		qualifi- cation	foreign	employed 1 year after start
		per- sons	share in %	mean days	mean days	share of durations < 150 days	mean	share in %	share in %
NONPARTICIPATION	(NONP)	6918	36	0	240		1.8	47	41.3
BASIC COURSES	(BAC)	1491	8	46	236	36	1.8	45	35.7
LANGUAGE COURSES	(LAC)	1719	9	71	225	36	2.2	72	31.1
COMPUTER COURSES	(COC)	1394	7	36	214	40	1.3	22	45.6
FURTHER VOCATIONAL TRAINING	(FVT)	424	2	74	231	35	1.6	38	44.7
OTHER TRAINING COURSES	(OTC)	497	3	94	263	23	1.8	43	44.2
EMPLOYMENT PROGRAMMES (PUBLIC)	(EP-PU)	1124	6	153	302	18	1.7	41	32.9
EMPLOYMENT PROGRAMMES (PRIVAT)	(EP-PR)	1349	7	142	299	18	2.0	51	30.9
TEMPORARY WAGE SUBSIDY	(TEMP)	4390	23	114	228	35	1.7	46	51.2

Note: Qualification is measured as skilled (1), semiskilled (2), and unskilled (3). For nonparticipants random start dates are used.

Table 3: Allocation of participants using different assignment rules

Assignment	All		Native languages 'Swiss'	
	Mean	Std. error	Mean	Std. error
Random assignment in existing treatment proportions	42.2		46.9	
Case worker assignment	41.5	0.4	46.1	0.6
Case worker assignment with BAC, LAC, TE-PU, TE-PR assigned to NONP	43.9		47.1	
Assignment of everyone to				
NONP	42.9	0.8	46.4	1.1
BAC	32.2	2.1	36.9	2.6
LAC	38.8	2.2	46.9	3.4
COC	38.6	2.6	42.4	2.3
FVT	42.3	3.8	47.6	4.1
OTC	50.0	3.8	54.2	4.6
TE-PU	33.2	2.9	42.6	3.5
TE-PR	35.4	2.6	41.1	3.4
TEMP	50.0	1.2	53.4	1.4
Assignment to treatment based on ...				
... largest predicted gross impact for each person	55.5		61.9	
... largest predicted gross impact for each person without NONP	57.2		63.3	
... smallest predicted gross impact for each person	25.7		30.3	
... smallest predicted gross impact for each person without NONP	26.7		31.0	
... largest predicted gross impact for each person imposing national supply constraint – effect based	49.3		54.8	
... smallest predicted gross impact for each person imposing national supply constraint – effect based	37.0		40.2	
... largest predicted gross impact for each person imposing national supply constraint – need based	47.8		53.6	
... largest predicted gross impact for each person imposing regional supply constraint - – effect based	48.4		54.0	
... smallest predicted gross impact for each person imposing regional supply constraint – effect based	37.5		40.6	
... largest predicted gross impact for each person imposing regional supply constraint – need based	47.2		52.7	
Note:	The seven regions used to define the regional supply constraints are defined as follows: (SG, AI, AR, TH, GR, GL, SH), (LUZ SZ, UR, OW, NW, ZU), (BE, FR, JU, SO, NB), (WT, WS, GE), (BS, BL, AA), TE, ZR. "Swiss" languages are defined for the current study as German, French and Italian.			

Table 4: Allocation of participants to treatments when assignment rules allow a deviation from the observed shares

Assignment	NONP	BAC	LAC	COC	FVT	OTC	TE- PU	TE- PR	TEM P
Observed shares	All 38.8	7.8	10.1	8.1	2.4	2.3	4.1	4.9	21.5
Assignment to treatment based on ...									
... largest predicted gross impact for each person	1.6	0.02	1.6	6.3	25.2	27.0	2.2	3.1	33.0
... largest predicted gross impact for each person without NONP		0.02	1.7	6.4	25.4	27.4	2.2	3.1	33.8
... smallest predicted gross impact for each person	0.9	10.1	16.2	0.7	5.97	3.28	48.2	14.4	0.1
... smallest predicted gross impact for each person without NONP		10.2	16.5	0.7	6.2	3.4	48.3	14.5	0.1
	Native language 'Swiss'								
... largest predicted gross impact for each person	1.1	0.01	6.7	3.5	24.5	31.2	2.9	2.0	27.8
... largest predicted gross impact for each person without NONP		0.02	7.0	3.1	25.8	33.3	3.0	2.3	25.6
... smallest predicted gross impact for each person	0.5	22.9	5.0	3.0	17.3	7.3	25.0	18.9	0.2
... smallest predicted gross impact for each person without NONP		22.9	5.7	2.9	17.2	5.7	25.9	19.5	0.2

Table 5: Allocation of participants using different assignment rules - Heterogeneity

Assignment	RAV I		RAV II			Region (by language)	
	Rural	Urban	Type I	Type II	Type III	Ger	F, I
Random assignment in existing treatment proportions	43.6	41.1	45.4	41.1	38.8	45.2	37.9
Case worker assignment	42.4	40.5	44.2	41.1	38.1	44.9	36.8
Assignment of everyone to							
NONP	44.7	41.1	47.0	41.8	38.6	46.6	38.4
BAC	34.5	37.6	35.4	35.1	32.1	37.7	28.2
LAC	37.3	34.4	37.4	39.3	32.5	35.5	38.5
COC	42.5	42.1	43.4	38.1	41.6	44.5	36.1
FVT	46.7	37.2	61.3	37.9	35.6	58.7	28.4
OTC	44.0	48.2	52.1	42.1	39.2	56.2	33.8
TE-PU	31.0	33.0	41.8	22.5	25.3	36.9	23.9
TE-PR	42.6	25.8	43.1	27.2	32.1	39.9	31.4
TEMP	50.1	49.6	50.0	51.2	48.6	52.7	45.6
Assignment to treatment based on ...							
... largest predicted gross impact	60.4	61.1	72.5	61.1	62.7	67.4	54.7
... largest predicted gross impact w/o NONP	61.8	62.2	72.3	61.5	65.3	68.0	55.4
... smallest predicted gross impact	23.8	15.5	25.8	14.1	13.7	28.3	14.7
... smallest predicted gross impact w/o NONP	24.2	16.2	26.1	13.9	14.7	28.4	14.8
... largest predicted gross impact imposing national supply constraint – effect based	51.6	53.1	55.9	52.5	52.6	52.7	49.0
... smallest predicted gross impact imposing national supply constraint – effect based	33.3	36.0	35.8	32.3	29.1	38.5	30.2
... largest predicted gross impact imposing national supply constraint – need based	49.7	51.9	54.2	49.6	49.6	50.8	46.1
... largest predicted gross impact imposing regional supply constraint – effect based	50.3	52.0	55.2	51.7	50.6	52.0	48.2
... smallest predicted gross impact imposing regional supply constraint – effect based	33.9	36.4	36.2	32.3	29.9	39.0	30.5
... largest predicted gross impact for each person imposing regional supply constraint – need based	48.3	50.6	53.7	48.8	48.2	50.2	45.5

Note: Types I to III relate to a classification by Atag, Ernst and Young Consulting (1999). Supply constraints imposed as observed in the specific subgroup considered.