Employment, Mobility, and Active Labor Market Programs*

Peter Fredriksson† Per Johansson‡

24 January, 2003

Abstract

Using a unique micro panel data set we investigate whether active labor market programs improve employment prospects and increase mobility in the longer run. We consider two prototype programs: job creation programs and training programs. We find that both programs reduce the chances of finding a job substantially. Moreover, both programs are associated with a locking-in effect: the probability of finding a job outside the home region decreases after program participation. However, this effect appears to stem exclusively from the decrease in the overall job finding rate.

KEY WORDS: Subsidized employment, labor market training, program evaluation, employment, contracted mobility.
JEL-CLASSIFICATION: J61, J64, J68, C41

*We thank Torbjørn Hægeland, Laura Larsson for helpful comments. We also acknowledge the comments from seminar participants at: IFAU, Australian National University, University of New South Wales, and the Nordic Econometric meeting.

†Department of Economics, Uppsala University, Institute for Labour Market Policy Evaluation (IFAU), and CESifo. Adress: Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden.. Phone: +46-18–471 11 13. Email: peter.fredriksson@nek.uu.se. Fredriksson acknowledges the financial support from the Swedish Council for Working Life and Social Research (FAS).

‡Department of Economics, Uppsala University, and IFAU. Adress: IFAU, Box 513, SE-751 20 Uppsala, Sweden. Phone: +46-18-471 70 86. Email: per.johansson@ifau.uu.se. Johansson acknowledges the warm hospitality of the Department of Economics at the University of New South Wales.

1 Introduction

Active labor market programs (ALMPs) have become an integral part of the standard tool kit for combating unemployment in the OECD countries. According to OECD (2001), the OECD countries allocate around 40 percent of total labor market expenditures to active measures on average.

Although appropriately designed ALMPs should be a useful tool in fighting unemployment, the empirical evidence on the efficacy of ALMPs is far from conclusive; see e.g. Calmfors et al. (2002). There are several reasons for the apparent failure of ALMPs to generate positive employment effects; to mention but a few in the comprehensive list of Calmfors (1994), ALMPs are usually associated with displacement and "locking-in" effects. A locking-in effect is said to exist if participation in ALMPs reduces the time available for search or if ALMPs decreases the incentives to change occupation or region of residence.[1]

In this paper we examine whether ALMPs are associated with locking-in effects. We have access to unique micro data where we can differentiate between outflows to employment in the home region and outflows to employment in other regions; for a description of the data see Edin and Fredriksson (2000). Hence, we can ask whether participation in programs affects the job finding probability in the home region as well as in other regions.

Previous research has documented a direct locking-in effect in the sense that search activity is lower among program participants than the openly unemployed; e.g. Edin and Holmlund (1991) and van Ours (2002). However, the received literature has little to say about potential locking-in effects after program completion. One strand of the literature studies the relationship between migration rates and program activity across local labor markets; e.g. Westerlund (1998) and Fredriksson (1999). The general result is that, if anything, higher program activity reduces migration. Another approach is to examine whether the individual mobility decision is affected by program activity in the region of residence; according to Widerstedt (1998) there is no significant relationship between the individual out-migration propensity and program activity.

Our data are much richer than the data sets commonly employed to study the relationship between employment, mobility and ALMPs. We know whether an individual has participated in a program or not. Therefore, we

[1] Notice that we take no stance on whether the locking-in effect good or bad for efficiency. Efficiency may be reduced if it exacerbates the misallocation of labor. If there is too much (or wasteful) mobility, a locking-in effect may be beneficial. Notice, though, that if there is wasteful mobility, then the reduction of mobility can be achieved at less cost by raising unemployment benefits; see Diamond (1981) on the last point.

can estimate individual treatment effects of program participation rather than using the correlation between migration rates and regional program activity. We distinguish between two types of programs: job creation programs and training programs. Job-creation programs are essentially measures that provide temporary employment in the home region. Training programs, on the other hand, offer re-training and, presumably, individuals acquire qualifications that are in general demand on the labor market. Therefore, it seems reasonable to expect that locking-in effects may be a more serious problem for job creation programs.

Our results can briefly be summarized as follows. Both programs reduce the outflow to employment. Similarly, program participation implies that contracted mobility declines. Relatively speaking, these effects are quite substantial in the longer run. We do not find much evidence suggesting that the type of program is important.

The remainder of the paper is outlined as follows. The next section sketches a simple analytical framework that we use as a guide for specification and interpretation. Section 3 discusses empirical and econometric issues confronting the evaluation. Section 4 describes the data. In section 5, we report the results. Section 6 concludes.

2 An analytical framework

This section describes the job search problem for heterogeneous individuals who are non-employed. We characterize the self-selection into labor market programs and use the model as a guide for the empirical analysis.

2.1 General set-up

Non-employed individuals can be in two states: (open) unemployment and labor market programs, indexed by $j = 0, 1$ respectively. They search in two regions – at (h)ome and (a)broad, indexed by $i = h, a$. Participation in labor market programs ($j = 1$) is assumed to affect the job-finding probability in two ways relative to open unemployment ($j = 0$). First, it shifts the overall probability of finding a job. Second, it shifts the relative efficiency of search in the home region. We think it is plausible that participation in programs shifts the relative search efficiency in favor of the home region; after all, program participation generally means more contacts with local employment officers and they specialize in job placements locally (at least to some extent).

Non-employed individuals choose search intensity in the two regions optimally. The rates at which jobs are located depend on the state and the number of units of search allocated to a particular region. A permanent move takes place if they locate a job outside the home region; thus, there is only contracted mobility. The job is kept until individuals "die" and permanently exit the labor market. The event of death happens at rate δ, and in such case they receive zero utility. Under these conditions, and assuming no discounting, risk neutrality, as well as infinite horizons, we can write the expected present value associated with finding a job at home as: $V_h^w = (w_h/\delta)$, and the expected present value of finding a job abroad as $V_a^w = (w_a/\delta) - m$.[2] The mobility cost (m) differs across individuals and is distributed according to $F(m)$, defined on the support $m \in (0, \overline{m})$. The value of employment as such (w_i/δ) does not depend on whether the individual has entered from unemployment or programs, or whether the individual had to move in order to get it.[3]

Let us describe the environment more precisely by writing down the asset values associated with open unemployment (V^0) and program participation (V^1). We measure search in efficiency units and let e_i^j denote the effort needed to produce s_i^j efficiency units of search. Effort is (strictly) increasing and convex in the efficiency units of search

$$\delta V^0 = b - \sum_i e_i^0 + \alpha^0 \sum_i [s_i^0 (V_i^w - V^0)] + \gamma \left[\max\left(V^0, V^1\right) - V^0\right] \tag{1}$$

$$\delta V^1 = b - \sum_i e_i^1 + \alpha^1 \sum_i [s_i^1 (V_i^w - V^1)] \tag{2}$$

The direct utility obtained in each state is given by $b - \sum_i e_i^j$, where b denotes unemployment income and utility is decreasing in effort. The effort functions depend on the state, which captures the fact that programs may shift the relative efficiency of searching at home. Job offers arrive at rate $\alpha^j s_i^j$ and in that event the job searcher enjoys a gain of $(V_i^w - V^j)$ in present value terms. The overall job finding rate may differ by state as indicated by α^j. Offers to participate in programs arrive at rate γ. Having received an offer the

[2] We impose a single regional wage rather than a regional wage offer distribution. This simplification is without loss since search intensity and reservation wages usually do the same job in the partial equilibrium search model.

[3] Although we do not model wage setting explicitly, these assumptions can be rationalized in a wage bargaining set-up if: programs do not affect labor productivity, there is continuous renegotiations, and the disagreement point is always the state of unemployment (or "death"). Note that, with continuous bargaining, any mobility costs incurred are sunk. The assumption that labor market programs do not affect productivity is of course a simplification, but we want to avoid this complication since we will only examine employment and migration in our empirical work.

unemployed decides on whether to participate or not; there are no sanctions imposed on those who reject the offer.

Individuals choose search intensity in each state by maximizing V^j. The first order condition for search is

$$\frac{\partial e_i^j}{\partial s_i^j} = \alpha^j (V_i^w - V^j) \tag{3}$$

for $i = h, a$ and $j = 0, 1$. To avoid corner solutions to the search problem we assume that $(w_a/\delta - \overline{m} - V^j(\overline{m})) > 0$, implying that all individuals will search at home as well as abroad. Total search as well as search along each specific channel of course depends on m. In particular, search at home increases and search abroad decreases with m. To see this, note that (1) and (2) implies $\partial V^j/\partial m \in (-1, 0)$. Thus the marginal return to search at home $(\alpha^j(V_h^w - V^j))$ is increasing in m, while the marginal return to search abroad is decreasing in m.

2.2 Self-selection into programs

As can be seen from equation (1) there will be some individuals who accepts and some who rejects an offer to participate in a program; thus self-selection into the program is an issue. Here we characterize self-selection into programs. In order to simplify the exposition, we make some assumptions about, *inter alia*, the functional form of the effort functions. We assume that

$$e_i^0 = \frac{(s_i^0)^2}{2} \quad i = h, a; \; e_h^1 = \frac{(s_h^1)^2}{2(1+\kappa)}; \text{ and } e_a^1 = \frac{(s_a^1)^2}{2(1-\kappa)} \tag{4}$$

where κ is the extent of "home bias" associated with participating in active labor market programs. If $\kappa > 0$, it is less costly in terms of effort to produce a given amount of search intensity when searching in the home market. Given (4) the conditions for optimal search have the following convenient forms

$$s_i^0 = \alpha^0(V_i^w - V^0), \quad i = h, a \tag{5}$$

$$s_h^1 = \alpha^1(1+\kappa)(V_h^w - V^1) \text{ and } s_a^1 = \alpha^1(1-\kappa)(V_a^w - V^1) \tag{6}$$

Moreover, we let the job arrival rate in programs be proportional to the job arrival rate in open unemployment:

$$(\alpha^1/\alpha^0) = 1 + \eta \tag{7}$$

Invoking (4) and (7), the program is completely characterized by two parameters: κ and η. If $\kappa = \eta = 0$, programs and unemployment are identical states.

Now, let us characterize the self-selection into programs. Define the individual with moving cost $\widehat{m}$ as the individual who is just indifferent between the program and open unemployment, i.e., $\widehat{m}$ is defined by $V^1(\widehat{m}) = V^0(\widehat{m})$. The type of self-selection into programs will be determined by the sign of

$$\Omega(\widehat{m}) \equiv \left[\frac{\partial V^1(\widehat{m})}{\partial m} - \frac{\partial V^0(\widehat{m})}{\partial m}\right] \tag{8}$$

If this is positive, it means that $V^1(m) \geq V^0(m)$ for $m \geq \widehat{m}$, that is, those with comparatively high moving costs will enter the program; if $\Omega(\widehat{m}) < 0$, then $V^1(m) < V^0(m)$ for $m \geq \widehat{m}$. The derivative of the value of unemployment with respect to the mobility cost for an individual who do not accept program offers is given by

$$\frac{\partial V^0}{\partial m} = -\frac{\alpha^0 s_a^0}{\delta + \alpha^0(s_h^0 + s_a^0)} \tag{9}$$

while the derivative of the value of program participation with respect to the mobility cost equals

$$\frac{\partial V^1}{\partial m} = -\frac{\alpha^1 s_a^1}{\delta + \alpha^1(s_h^1 + s_a^1)} \tag{10}$$

According to (9) and (10) the crucial aspect is whether the program increases or reduces the job-finding rate abroad relative to open unemployment. Some manipulations of (8) using (5), (6), (9)-(10), $V^1(\widehat{m}) = V^0(\widehat{m})$, and (7) yield

$$\text{sign}\left[\Omega(\widehat{m})\right] = \text{sign}\left[\delta\left(1 - (1+\eta)^2(1-\kappa)\right) + \frac{2\kappa\alpha^1 s_h^1}{1+\kappa}\right] \tag{11}$$

In general the sign of $\Omega(\widehat{m})$ is ambiguous, depending *inter alia* on the sign of η and κ. We summarize three distinct cases in the following proposition.[4]

Proposition 1 *Case* (i) : *If* $\eta \leq 0$ *and* $\kappa > 0$, *those with* $m \geq \widehat{m}$ *will opt for the program. Case* (ii) : *If* $\eta > 0$ *and* $\kappa \leq 0$, *those with* $m < \widehat{m}$ *will opt for the program. Case* (iii) : *If* $\eta = \kappa = 0$, *there is no selection into programs.*

The interpretation of these conditions is straightforward. Case (iii) is trivial: if $\eta = \kappa = 0$, there is no difference between participating in a labor market program and being openly unemployed; hence there will be no self-selection into programs. If $\eta = 0$, the type of selection is solely determined by the sign of κ. If there is a home bias, $\kappa > 0$, then those with higher moving

[4] These cases do not exhaust all possibilities. It is readily verified that $\kappa, \eta > 0$, but $\kappa/(1-\kappa) > \eta(2+\eta)$ implies $\Omega(\widehat{m}) > 0$.

costs will choose the program since for given search effort they will find a job more easily in the home region. If $\kappa = 0$, the sign of η determines the type of selection. If $\eta > 0$, those with comparatively low moving costs enter the program since the probability that they will pay the moving cost increases along with the improvement of general employment prospects. Cases (i) and (ii) are the interesting ones that we will consider in more detail in what follows.

2.3 Evaluation parameters

For purposes of evaluating labor market programs, we are interested in the outflow to employment at home and to employment in other regions than the home region. The framework outlined above implies a variable coefficients framework, i.e. the response to treatment will vary with m. In the empirical analysis we will focus on estimating a variant of treatment on the treated.[5] In this setting, treatment on the treated (TT_i, $i = h, a$) for those participating in a labor market program ($D = 1$) equals

$$TT_i = \alpha^0 \int_{D=1} [(1+\eta)s_i^1(m) - s_i^0(m)]dF(m) \tag{12}$$

In principle we need two observations for each individual to calculate (12). The classical problem is, of course, that we cannot observe $s_i^1(m)$ and $s_i^0(m)$ at the same time for each individual.

What we can readily observe in the data is the average outflow of individuals who have participated in the program and those who have not. Let Δ_i denote the difference in the average outflow. Then the simple difference equals $\Delta_i = \alpha^1 \int_{D=1} s_i^1(m)dF(m) - \alpha^0 \int_{D=0} s_i^0(m)dF(m)$. The following proposition gives the sign of the bias of Δ_i for estimating TT_i (B_i).[6]

Proposition 2 *Case* (i) : $B_h > 0$ *and* $B_a < 0$. *Case* (ii) : $B_h < 0$ *and* $B_a > 0$.

Proof. The bias equals

$$B_i = \alpha^0 \left[\int_{D=1} s_i^0(m)dF(m) - \int_0^{\overline{m}} s_i^0(m)dF(m) \right]$$

[5] Since this is a variable coefficients framework the average treatment effect (ATE) will differ from treatment on the treated (TT). It is fairly straightforward to verify that $ATE > TT$ independently of the type of selection into the program.

[6] It may also be of interest to examine the bias of the naive estimator of the overall outflow to employment. If aggregate search intensity $(s_h^0 + s_a^0)$ is decreasing in m then this bias has the same sign as B_a. Aggregate search intensity will be decreasing in m if $s_h^0 \geq s_a^0$ which holds if $w_h \geq (w_a - \delta m)$.

In case (i) $D = 1$ if $m \in (\widehat{m}, \overline{m})$. Since $\partial s_h^j / \partial m > 0$ and $\partial s_a^j / \partial m < 0$ (for all m) the term in square brackets is positive when considering the outflow to employment in the home region and negative when considering the outflow to employment outside the home region. In case (ii) $D = 1$ if $m \in (0, \widehat{m})$ and the opposite holds. ■

It is not obvious how to define the locking-in effect. One candidate is just to compare the employment hazard to other regions than the home region, e.g., to calculate TT_a as defined in (12). Another alternative is to define the locking-in effect relative to the overall job-finding rate in each state, that is

$$TT_a' = \int_{D=1} \left[\sigma_a^1(m) - \sigma_a^0(m)\right] dF(m) \tag{13}$$

where $\sigma_a^j = s_a^p(m)/(s_h^p(m) + s_a^p(m))$. This measure of the locking-in effect thus amounts to comparing search allocation across the two states for each individual. What is the sign of the bias of the naive estimator of TT_a': $\Delta_a' = \int_{D=1} \sigma_a^1(m) dF(m) - \int_{D=1} \sigma_a^0(m) dF(m)$? The following proposition gives the result

Proposition 3 *Case* (i) : $B_a' < 0$. *Case* (ii) : $B_a' > 0$.

Proof. In this case we can write the bias term as

$$B_a' = \left[\int_{D=1} \sigma_i^0(m) dF(m) - \int_0^{\overline{m}} \sigma_i^0(m) dF(m)\right]$$

In case (i) $D = 1$ if $m \in (\widehat{m}, \overline{m})$. Since $\mathrm{sign}\frac{\partial \sigma_a^j}{\partial m} = \mathrm{sign}\{\frac{\partial s_a^j}{\partial m} s_h^j - \frac{\partial s_h^j}{\partial m} s_a^j\} < 0$ (for all m) the bias is negative. In case (ii) $D = 1$ if $m \in (0, \widehat{m})$ and the opposite holds. ■

Thus, the bias for the candidate evaluation parameter is equal in sign as the one we considered earlier. In the empirical application we will consider both evaluation parameters. If the results differ depending on whether we consider, e.g., TT_a or TT_a' it may say something about what the program does to treated individuals. In particular, the direct influence of the overall job finding rate is eliminated when considering search allocation so TT_a' will reflect the potential for home bias associated with program participation.

3 Empirical and econometric issues

The evaluation problem considered in this paper has at least three facets; by and large they stem from the fact that we have to rely on observational data to examine the issues at hand. First, as high-lighted by the previous section, there is the problem of self-selection. Second, there may be duration dependence such that elapsed duration before entering the program will affect the chances of finding a job. Third, programs may start at any point in time during the unemployment spell. The solution to the first problem requires a conditional (or mean) independence assumption. The complications arising from duration dependence requires some care but has a fairly straightforward solution. The third problem puts restrictions on what we can consistently estimate. In a companion paper (see Fredriksson and Johansson, 2003) we have discussed the third issue at length. In what follows we discuss these three problems in isolation; the final subsection outlines our estimation approach.

3.1 Conditional independence and matching

Our model suggests that it is difficult to come up with an instrument that influences the selection into programs but does not affect search. The reason is that search intensity is proportional to $(V_i^w - V^j)$ and the selection into programs is determined by a comparison of V^1 and V^0. So, anything that influences search intensity will in general also determine the selection into programs. Instead we rely on the argument that we can observe and, hence, condition on many factors influenced by unobserved heterogeneity. So in terms of the model we effectively make a selection on observables assumption, i.e., we assume that we can write the mobility cost as $m = m(\mathbf{x})$. Since we have an unusually rich data set containing indicators of previous mobility, unemployment and income histories, household composition etc. this is potentially a viable strategy. The conditioning on observed covariates can be done via (the equivalent of) regression or matching (see e.g. Rosenbaum and Rubin, 1983; Rosenbaum, 1995; Heckman et al., 1998; Dehejia and Wahba, 1999). In this paper we will take a matching approach.

The fact that we have multiple treatments (job creation programs and training programs) makes the matching problem slightly non-standard. Here we will discuss the conditional (or mean) independence assumption required for matching in this setting. We make two simplifying assumptions for expositional convenience. First, we assume that there is only one outcome: unemployment duration (T). Second, we assume that the program starts at a fixed point in time. We will relax the second assumption later.

Let us introduce some notation. The treatment states are denoted as follows: 0 denotes open unemployment; 1 participation in a job creation program; and 2 participation in a training program. Define the potential unemployment duration associated with each of the three states as T_0, T_1, and T_2, with T_{ik} denoting the outcome for individual i if i were to receive treatment k. Further, let $D = \{0, 1, 2\}$ denote the actual treatment, so that $D_i = k$ if individual i receives treatment k. Since each individual receives only one of the treatments, the remaining two potential outcomes are unobserved counterfactuals.

In the evaluation we are interested in the pair-wise comparisons of the average effect of treatment k relative to treatment k' conditional on assignment to treatment k, for all three combinations of k and k'. Given that the program starts at a fixed point in time, the object of evaluation is

$$E(T_k - T_{k'}|D = k) = E(T_k|D = k) - E(T_{k'}|D = k), \;\; k, k' = 0, 1, 2 \qquad (14)$$

In our application this is, for instance, the average effect of participating in a training program ($k = 2$) for an individual registering as unemployed compared to the hypothetical state in which (s)he stays openly unemployed ($k' = 0$) or participates in job creation ($k' = 1$). The first term, the average duration following treatment k for individuals who have participated in k, is observed (if there is no censoring). This is not the case for the counterfactual $E(T_{k'}|D = k)$, i.e., the expected duration participants in k would have experienced had they taken k' is not observed. Hence, we need to invoke identifying assumptions to overcome this fundamental missing data problem.

Since we are interested in pair-wise comparisons, we require conditional independence for the sub-populations receiving either treatment k or treatment k'. One such identifying assumption (see Imbens, 2000) is that, conditional on $\mathbf{X}$, T_k and $T_{k'}$ are statistically independent of the assignment:[7]

$$(T_{k'}, T_k) \perp\!\!\!\perp D|\mathbf{X} = \mathbf{x}, \forall \mathbf{x} \in \Xi \;\; D \in (k, k') \qquad (15)$$

where $\Xi \subseteq \mathbb{R}^P$ defines the set of $\mathbf{X}$ for which the treatment effect is defined. If we let

$$D_k = \begin{cases} 1 & \text{if } D = k \\ 0 & \text{otherwise} \end{cases}$$

the independence assumption (15) implies that the unobserved counterfactuals can be identified as

$$E(T_{k'}|D = k) = E_X[E(T_{k\prime}|D_k = 1, \mathbf{x})|D_k = 1] = E_X[E(T|D_{k'} = 1, \mathbf{x})|D_k = 1]$$

Here the inner expectation is identified using the independence assumption and the outer expectation is taken with respect to the distribution of $\mathbf{X}$ for the participants in treatment k. The outer expectation highlights that there must be sufficient overlap in the distribution of $\mathbf{X}$ in order to adjust for

[7] Imbens (2000) refers to this type of identifying assumption as weak unconfoundedness.

differences in $\mathbf{x}$ among the participants in k and k'. If there are regions where the support of $\mathbf{x}$ does not overlap, matching has to be performed over the common support; the estimated treatment effect is then the mean treatment effect for those treated within the common support.[8]

The independence assumption (15) is stated in terms of a potentially large set of covariates ($\mathbf{x}$). An important practical result, derived by Rosenbaum and Rubin (1983) for the single treatment case, is that it is sufficient to condition on a scalar function of the covariates – the propensity score – to adjust for differences in observed characteristics. This result generalizes to the case of several treatments. Let

$$e_{kk'}(\mathbf{x}) = \Pr(D_k = 1|\mathbf{x}, D_k = 1 \vee D_{k'} = 1)$$

be the conditional probability to enter k given a choice between k and k'. Then the scalar $e_{kk'}(\mathbf{x})$ is a "balancing score" for the separate comparison of the two sub-populations. Under the independence assumption (15), the counterfactual can be estimated as

$$E(T_{k'}|D_k = 1) = E(E(T|D_{k'} = 1, e_{kk'}(\mathbf{x}))|D_k = 1),$$

where E is the expectation with respect to $e_{kk'}(\mathbf{x})$. Thus by controlling for systematic differences between the two sub-populations based on $e_{kk'}(\mathbf{x})$, the average outcome experienced by the matched pool of participants in k' identifies the average counterfactual outcome for participants in k had they participated in k' instead

3.2 Duration dependence

Suppose now that there is duration dependence. Continue to assume that the program starts a specific time point $\bar{t}$, but suppose that there is some variation in the date of unemployment entry (t_{0i}). For illustrative purposes, assume that this is the only form of heterogeneity, such that we can suppress the covariates ($\mathbf{x}$), and that there is only one prototypical program.

The fact that there is variation in the date of unemployment entry implies that the duration prior to program *start* (t^s) will vary over individuals since $t_i^s = \bar{t} - t_{0i}$. The question then is: What is a valid control group for a treated individual with prior duration t_i^s, given that there is duration dependence? Since the outflow to jobs will be different for individuals with durations less than t_i^s for reasons unrelated to the program as well as heterogeneity,

[8]Notice that if the treatment effect varies among individuals, restricting the interpretation to the common subset may change the interpretation of the parameter being estimated.

it is clear that one should remove these individuals from the control group. Moreover, one would generally like to condition on the date of unemployment entry as the state of the cycle at this point may have implications for future outcomes. Thus the comparison sample for individual i with prior duration t_i^s consists of those unemployed individuals with the same date of registration as i and unemployment duration (T) satisfying $T > t_i^s$. All that this means is that the control group should consist of individuals who were at risk of starting the program at $\bar{t}$. Thus it is fairly straightforward to take care of the complications arising from duration dependence.

Typically we are interested in the difference in the hazard to employment between the treated and the controls after the start of the program. But the treatment effect potentially varies by duration prior to program start if there is duration dependence. Therefore, to construct an average effect we calculate weighted averages of these treatment effect using the distribution of prior durations for the treated as weights.

3.3 Random program starts

The assumption that the program starts at a fixed time point is clearly unrealistic in most situations. In most cases, the timing of a program start is best thought of as the outcome of a stochastic process involving inter alia the arrival rate of job offers and program participation offers.

In the case of only one treatment, one would in general like to estimate, e.g., treatment on the treated

$$E(T_1|D = 1) - E(T_0|D = 1) \tag{16}$$

This is simply the one-dimensional analogue of (14). Note that $D = D(t^s)$, where the duration prior to program start (t^s) is stochastic even if we consider only individuals with the same date of unemployment entry.

To estimate (16) one is tempted to define a control group that was never treated for each treated individual. But finding this control group involves conditioning on the future since programs may start at any point in time. Defining the control group in this way implies conditioning on the outcome variable as those who do not enter the program in the future to a large extent consist of those who had the luck of finding a job. Thus the CIA required for estimating (16) will not be valid.

The above observation suggests that the object of evaluation has to be more modest. Without additional assumptions it is only possible to estimate a treatment on the treated for those treated up to a certain time point, $\bar{t}$, i.e.

$$E(T_1(\bar{t})|D(\bar{t}) = 1) - E(T_0(\bar{t})|D(\bar{t}) = 1) \tag{17}$$

where $T_1(\bar{t})$ is the potential unemployment duration if treated at $\bar{t}$ and $T_0(\bar{t})$ is the potential duration of unemployment if not treated at $\bar{t}$. The implied definition of the control group is one that includes individuals who may take part in the program in the future. As such it may be difficult to interpret this estimand. But notice that it is a relevant policy parameter in the environment we are considering. It answers the question: What is the average effect of treating an individual at $\bar{t}$ relative to not doing so? Choosing the alternative – i.e. no treatment at $\bar{t}$ – does not rule out future treatment since the program continuous to operate.

3.4 Estimation

Now we have set the stage for describing our estimation approach. We are interested in the hazards to employment in the home region and employment outside the home region. We treat these states as absorbing in a competing risks framework.

In the previous sub-section we discussed the evaluation parameters in the terms of the difference in unemployment duration. For our purposes it is more convenient to base the estimates on the survival functions. This is no restriction since

$$E(T_j) = \int_0^\infty S_j(t)dt,$$

where $S_j(t) = \exp(-\int_0^t \lambda^j(\tau)d\tau)$ is the potential survival function if treated with $j = 0, 1$, and $\lambda^j(t)$, the hazard to employment, is analogously defined. In other words the survival function integrates to mean unemployment duration. If the data contained completed spells we can thus estimate

$$\begin{aligned} & E_{T^s|D=1}\left[E(T_1(t)|D(t) = 1) - E(T_0(t)|D(t) = 1)\right] \\ = \; & E_{T^s|D=1}\left\{\int_0^\infty \left[S_1(t|D(t) = 1) - S_0(t|D(t) = 1)\right] dt\right\} \end{aligned}$$

Here $S_1(t|D(t) = 1)$ is the survival function for those treated up to t, and $S_0(t|D(t) = 1)$ is the counterfactual survival function for this group of individuals. We take the expectation over the inflow distribution of the treated, $E_{T^s|D=1}$, in order to calculate an average effect. We should emphasize that this average is not equal to treatment on the treated.

In practical applications the data are almost always right-censored. It is then appropriate to base the estimates of the effects on the survival function up to a censoring point ($\bar{T}$). Thus, our evaluation focuses on the parameter

$$\Delta(t) = S_1(t|D(t) = 1) - S_0(t|D(t) = 1), t \in (0, \overline{T})$$

If the sample is homogenous – both in terms of covariates and the date of unemployment entry – one option is to calculate the difference in survival propensities using a Kaplan-Meier (KM) estimator (or product limit estimator). We apply the KM estimator to estimate $S_1(t|D(t)=1)$ for those treated at t or earlier. The comparison group of individuals still unemployed but not treated at t forms the KM estimator of $S_0(t|D(t)=1)$.

For an individual who has been treated at $t^s \leq t$ the empirical hazard at time t is given by

$$\lambda(t, D(t)=1) = \frac{n^1(t)}{R^1(t)} = \frac{1}{R^1(t)} \sum_{i=1}^{R^1(t)} y_i(t)$$

where $y_i(t) = 1$ if individual i that starts a program in $t^s \leq t$ leaves unemployment at t and $R^1(t)$ is the number of individuals still at risk among individuals who entered treatment at $t^s \leq t$. Hence, $n^1(t) = \sum_{i=1}^{R^1(t)} y_i(t)$ is the number of individuals in the risk set leaving in t. For the comparison group we calculate

$$\lambda(t, D(t)=0) = \frac{n^0(t)}{R^0(t)}$$

Here $R^0(t)$ is the set of individuals that has not joined the program at t and are at risk of being employed in t; $n^0(t)$ is the number of individuals in the risk set leaving in t. $\lambda(t, D(t)=0)$ is an unbiased estimator of the hazard rate to employment for a randomly chosen individual who have not received treatment at t.

The potential survival functions conditioning on $D(t)=1$, $S_j(t|D(t)=1)$ $j=0,1$ are estimated as

$$S(t|D(t)=j) = \prod_{s=l}^{t} (1-\lambda(s, D(s)=j)), \ t=l,...,\overline{T}, \ j=0,1$$

Treatment on the treated for those treated prior to t can then be calculated as the difference between the two survival functions, i.e.

$$\widehat{\Delta}(t) = S(t|D(t)=1) - S(t|D(t)=0), \quad t=l,...,\overline{T}. \tag{18}$$

Finally, we calculate the variance of the survival functions as (cf. Lancaster, 1990)

$$\mathrm{Var}(S(t|D(t)=j) = S(t|D(t)=j)^2 \sum_{t=l+1}^{s} \frac{n^j(t)}{(R^j(t)-n^j(t))R^j(t)}, \ j=0,1 \tag{19}$$

So far we have been silent about controlling for heterogeneity. Now let us adapt our estimation approach to incorporate this complication. As indicated above, we introduce a matching approach based on the conditional independence assumption (15).

For treated individuals we simply use the estimator above, i.e., $\lambda(t, D(t) = 1) = n^1(t)/R^1(t)$. The problem is finding a matched comparison group for individuals treated at $t = \bar{t}$. To illustrate the matching procedure let i index treated individuals and c index individuals in the comparison group.[9] We use one-to-one matching based on propensity scores $\omega(m)$, $m = i, c$. To this end we apply a logit maximum likelihood estimator to estimate $\omega(m)$. The unique match (for each $\bar{t}$) is found by minimizing the distance between the estimated propensity scores:

$$c_{i\bar{t}} = \arg\min_{c \in N(\bar{t})} |\widehat{\omega}(i) - \widehat{\boldsymbol{\omega}}(c)|, \tag{20}$$

where $\widehat{\boldsymbol{\omega}}(c)$ is the $(N(\bar{t}) \times 1)$ vector of estimated propensity scores at time $\bar{t}$. After finding a match for a randomly drawn individual i, the process starts over again until $n_{cs}(\bar{t})$ comparable individuals is found in the comparison sample. Here $n_{cs}(\bar{t})$ is the number of individuals on the common support.

With a matched sample of controls we can e.g. calculate the adjusted hazard

$$\lambda(t, D(t) = 0) = \frac{1}{R^1(t)} \sum_{i=1}^{R^1(t)} y_{c_{i\bar{t}}}(t),$$

where the index $c_{i\bar{t}}$ is defined by (20). With this estimate in hand we have all the components necessary to calculate (18) adjusted for heterogeneity.

4 Data

Our empirical analysis is based on the data base LINDA; see Edin and Fredriksson (2000). LINDA contains a panel of around three percent of the Swedish population; the data are also cross-sectionally representative. LINDA is a collection of register data including the income registers, the censuses, and the unemployment register. The unemployment register begins in August 1991; the censuses are available every fifth year from $1960 - 90$, while the income registers are available annually starting in 1968. The important registers for our purposes are the unemployment and income registers. The latter register contains very detailed residential information; from this

[9] Notice that since the treatment indicator is time-varying the set of individuals in the treated and comparison group changes over time.

information we can meaningfully construct local labor markets and analyze mobility between local labor markets.

Using the unemployment registers, we construct a flow sample from the spells of unemployment and program participation starting during 1993. We include individuals aged 25 to 50 at the time. Moreover, we exclude individuals suffering from a work related handicap and individuals who participated in a vocational rehabilitation program. Temporary employment, job change, and part-time unemployment are not considered spells of unemployment, even though individuals in these states can register at the employment offices.

To these data we match individual earnings, mobility, and unemployment histories. We trace the sampled individuals back to 1987 for earnings and mobility, and to August 1991 in the case of unemployment; we also know whether the individual was in the unemployment register when it started.

We follow the individuals in our sample until the end of 1997. So treatment can, in principle, take place between 1993 and 1997 and analogously for the outcomes of interest.

The resulting data set has 11,462 individuals. For 400 individuals we lack information on some of the key characteristics; these individuals are deleted from the sample.[10] This leaves us with 11,062 observations. In the appendix we provide exact definitions of the variables used in the analysis; see Table 5. Descriptive statistics are given in Table 1. For variables that are time varying, this information pertains to the time of registration at the unemployment office. Notice that, in addition to the mobility, unemployment, and earnings histories, we have access to information on the economic status of the household and program activity at the local Public Employment Service (PES) office.

Our classification of individuals as participants in a job creation program (JC) or participants in a training program (TP) is based on the first program they participate in after unemployment entry. In the descriptive part of the analysis, we also use the individuals who never took part in any of these programs. We refer to this group as non-treated (NT) but note that this is not a valid comparison group for estimating causal treatment effects for reasons outlined in section 3. The total sample includes 1,063 (9.6 percent) individuals who were classified as TP-participants and 1,857 (16.8 percent) individuals who were classified as JC-participants.

We are primarily interested in the time it takes to find employment at home or abroad. An individual is defined as having found employment if

[10]For 119 individuals we lack information on social assistance receipt and we could not identify the Public Employment Service (PES) office handling the individual in 374 cases.

Table 1: Descriptive statistics

Variable	Mean	Std.Dev.	Min.	Max.
JC	0.17	0.37	0.00	1.00
TP	0.10	0.29	0.00	1.00
Employment	0.71	0.45	0.00	1.00
Mobility and employment	0.02	0.15	0.00	1.00
Duration (months)	27.76	21.11	1.00	60.00
Female	0.45	0.50	0.00	1.00
Age/10	3.42	0.73	2.50	5.00
Immigrant	0.18	0.38	0.00	1.00
High school education	0.53	0.50	0.00	1.00
University education	0.22	0.41	0.00	1.00
No UI eligibility	0.19	0.39	0.00	1.00
Cash assistance	0.06	0.25	0.00	1.00
Single	0.51	0.50	0.00	1.00
# kids > 0	0.43	0.49	0.00	1.00
House owner	0.36	0.48	0.00	1.00
(Household earnings)/10^5	0.63	0.95	0.00	8.66
Social assistance receipt	0.14	0.34	0.00	1.00
Individual histories				
# moves prior to -93 ≥ 1	0.18	0.39	0.00	1.00
No unemployment info.	0.40	0.49	0.00	1.00
(Days in open unempl.)/100	0.74	1.09	0.00	5.18
(Days in TP)/100	1.78	6.41	0.00	5.18
(Days in JC)/100	0.53	2.97	0.00	4.44
(Earnings -90)/10^5	1.09	0.70	0.00	10.18
(Earnings -91)/10^5	1.12	0.77	0.00	13.49
(Earnings -92)/10^5	1.04	0.79	0.00	15.22
Local characteristics				
Fraction in TP at PES	0.09	0.06	0.00	1.00
Fraction in JC at PES	0.18	0.07	0.00	1.00
# vacancies	0.54	0.67	0.00	2.08
# unemployed	3.43	3.60	0.02	10.64

(s)he left the register for: employment, temporary employment, and part time employment.[11] However, there is also a substantial amount of attrition in the unemployment register. It is reasonable to assume that some of this attrition is due to employment. Indeed, Bring and Carling (2000) show that the misclassification using the register can be severe. In a follow up study of 200 droup-outs they asked the question: "Were you employed (or becoming employed) at the time of attrition?". 44.7 percent (of the effective sample of 168 individuals) respondent "yes" to this question. A plausible assumption is that the classification error is larger among the individuals who move. Therefore, we apply the following strategy in classifying an individual as employed given that (s)he is a "drop-out". We begin by calculating the monthly earnings (starting from the time of classification as a drop-out) for the drop-outs. Then, separately for JC, TP and NT, we classify the top 44.7 percent in terms of earnings as being employed at the time of being registered as a drop-out.[12]

With the definition of employment in hand, we also know the exact date when the individual obtained a job from the unemployment register. Notice, though, that we have no information pertaining to the location of the job at that point in time. However, there is continuous time information on the local labor market where the individual resides as long as (s)he is registered at the unemployment office.[13] Also, there is information on the residence of an individual at the end of each year in the income registers. We combine these two pieces of information with a few assumption in order to create monthly job and mobility data. The following example describes the procedure; the principle is that we use the residence information that is closest in time to the employment event.

Suppose that an individual obtains a job sometime during a year y. If this individual returns to the unemployment register before the end of that year, we classify him as having obtained a job outside the home region if the PES office where (s)he registers is located in a different local labor market relative to the original residence according to the unemployment register. If the individual does not return to the unemployment register, we use the

[11] We have also used a more generous definition of employment including the individuals that remain in the register despite having found employment (remember that an individual can remain in the register, searching for a new job, despite being, e.g., part-time employed). Using the more generous definition does not change the results.

[12] Notice that we have performed the analysis either treating all "drop-outs" as attrition or treating all drop-outs as employed. The results are insensitive with respect to these two alternative classifications of the drop-outs.

[13] A local labor market (or travel to work area) is defined on the basis of observed commuting behavior in the Population Census of 1990. The classification we are using divides Sweden into 111 local labor markets.

Table 2: Descriptive statistics by treatment status

Variable	mean			t-ratio		
	JC	TP	NT	JC/TP	JC/NT	TP/NT
Employment	0.29	0.30	0.81	-0.14	-45.68	-35.42
Mobility and employment	0.01	0.01	0.03	-0.76	-6.57	-4.11
Duration (months)	46.62	46.67	22.82	-0.09	53.50	43.59
Female	0.54	0.53	0.42	0.72	9.21	6.39
Age/10	3.40	3.44	3.42	-1.44	-1.24	0.70
Immigrant	0.22	0.31	0.15	-4.82	7.25	10.79
High school education	0.55	0.53	0.53	0.72	1.25	0.14
University education	0.18	0.13	0.24	3.37	-5.88	-9.21
No UI eligibility	0.14	0.25	0.20	-7.40	-6.56	3.92
Cash assistance	0.04	0.09	0.07	-4.72	-4.13	2.72
Single	0.52	0.49	0.51	1.90	1.30	-1.21
# kids > 0	0.47	0.50	0.42	-1.98	3.85	5.36
House owner	0.30	0.29	0.38	0.60	-6.27	-5.72
(Household earnings)/10^5	0.57	0.59	0.65	-0.55	-3.48	-2.15
Social assistance receipt	0.18	0.25	0.11	-3.97	7.43	9.88
Individual histories						
# moves prior to -93 ≥ 1	0.20	0.22	0.18	-1.65	1.84	3.33
No unemployment info.	0.27	0.37	0.44	-5.30	-13.89	-4.24
(Days in open unempl.)/100	1.05	0.82	0.66	4.91	12.36	4.21
(Days in TP)/100	2.67	2.91	1.44	-0.82	6.39	5.99
(Days in JC)/100	1.14	0.51	0.39	4.60	7.04	1.23
(Earnings -90)/10^5	0.94	0.89	1.16	1.63	-13.66	-11.75
(Earnings -91)/10^5	0.93	0.92	1.19	0.58	-14.35	-11.33
(Earnings -92)/10^5	0.82	0.86	1.12	-1.13	-16.39	-10.43
Local characteristics						
Fraction in TP at PES	0.09	0.10	0.09	-4.67	3.46	6.76
Fraction in JC at PES	0.19	0.19	0.18	-0.44	2.62	2.54
# vacancies	0.52	0.47	0.55	2.09	-1.94	-4.09
# unemployed	3.30	3.06	3.51	1.90	-2.24	-4.13
# observations	1,857	1,063	8,142			

information in the income register pertaining to the end of year y and classify him analogously if the local labor market differs from the original one according to the unemployment register. The remainder of the outflows to employment are classified as being to jobs in the home region.

The strategy outlined above, in principle, yields daily duration data for unemployment and residence. However, in the empirical analysis we aggregate time to monthly intervals. Descriptive statistics by treatment status are reported in Table 2.

We note that prior mobility is somewhat greater among those in TP; 22 percent of those in TP have moved prior to 1993 as compared with 20 and 18 percent for those in JC and NT. Unemployment insurance (UI) eligibility is lower for those in TP than those in NT and JC. This is probably related to the fact that proportion of immigrants is greatest among TP participants. There are more females in the programs than in the group of non-treated. Also, program participants have a longer history of previous unemployment and program participation. This is especially true for those in JC. The last four rows shows that local characteristics are important for treatment status. For instance, the probability of taking part in a training program is higher if the individual is registered at a PES office which has a greater share of its "clients" in such a program relative to other offices.

The final pieces of descriptive facts that we want to show pertains to employment and employment outside the home region. We first note that contracted mobility is a fairly rare event. Only two percent has made a move to a job outside the home region. Moreover, exits to employment in general and contracted mobility are more prevalent among the non-treated. Figures 1 and 2 show that this pattern is true for all durations. However, these may be a consequence of how we have defined non-treated sample and can neither be taken as evidence that the programs are not useful in providing work nor that the programs reduce mobility. In the following section, we will analyze these effects in more detail.

5 The evidence

This section reports the estimation results. We begin with the discrete choice regression model. These estimates form the basis for the estimated propensity scores that we use to adjust for heterogeneity. The next two sections give estimates of the treatment effects. We focus on three outcomes: the overall outflow to employment, the outflow to employment in another region, and the outflow to employment elsewhere given that a job has been found: the last outcome is our empirical counterpart to search allocation; see eq. (13). Throughout we report the effect of treatment for those who have been treated prior to a fixed time point. These parameters are more restricted in scope than the common treatment of the treated estimand. However, they do have a causal interpretation given conditional independence.

5.1 The logit regression model

To be as flexible as possible, we run separate logits for each starting point and each program. Table 3 gives an example of the results from such re-

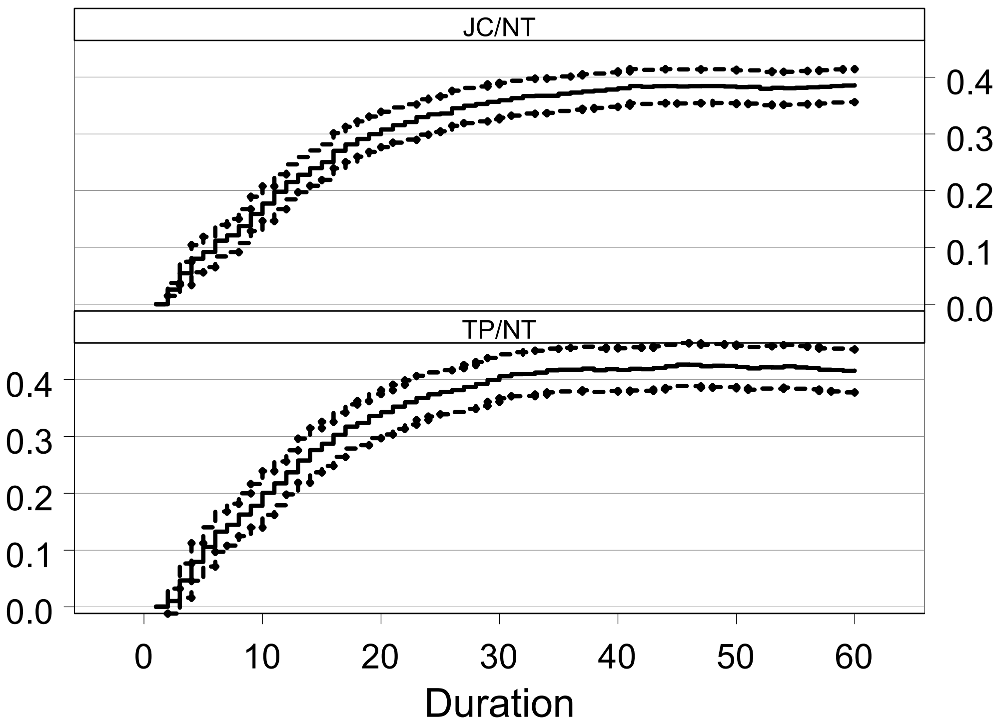

Figure 1: Difference in survival rates. Risk: employment. (Dotted lines are 95 % confidence intervals)

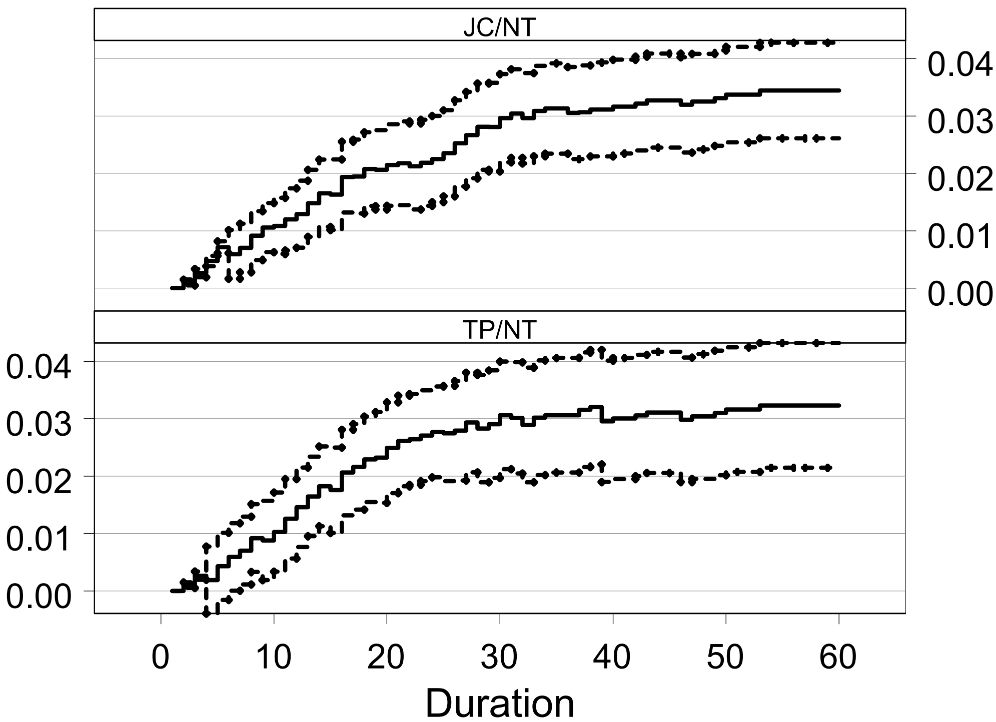

Figure 2: Difference in survival rates. Risk: employment and mobility. (Dotted lines are 95 % confidence intervals)

Table 3: The probability of entering a program during the first month after unemployment entry

	Unconditional				Conditional on job			
	JC		TP		JC		TP	
Variable	Est.	t	Est.	t	Est.	t	Est.	t
Female	0.09	0.53	0.05	0.20	0.14	0.64	0.07	0.23
Age/10	0.55	0.50	-0.12	-0.08	0.24	0.17	3.00	1.47
(Age^2/10)	-0.05	-0.32	0.05	0.25	-0.03	-0.13	-0.39	-1.40
Immigrant	-0.15	-0.68	0.28	1.06	-0.24	-0.74	0.02	0.06
High school ed.	0.00	0.01	-0.09	-0.37	0.03	0.11	0.21	0.67
University ed.	-0.29	-1.21	-0.40	-1.26	-0.61	-1.78	0.17	0.45
No UI eligibility	-0.67	-2.57	0.77	3.01	-0.47	-1.41	0.34	0.89
Cash assistance	-1.59	-2.68	0.53	1.37	-2.05	-2.02	0.86	2.08
Single	-0.22	-1.07	0.15	0.52	0.01	0.05	-0.77	-2.21
# kids > 0	-0.08	-0.45	0.55	2.14	0.14	0.59	-0.50	-1.55
House owner	0.14	0.84	-0.05	-0.18	0.17	0.82	-0.68	-2.20
(Household earnings)/10^5	-0.02	-0.09	0.58	2.23	0.15	0.53	0.18	0.47
Social assistance receipt	-0.20	-1.69	-0.05	-0.33	-0.14	-0.91	0.03	0.17
# moves prior to -93 > 1	-0.10	-0.51	-0.02	-0.08	-0.26	-0.93	0.33	1.08
No unemployment info.	-2.56	-6.00	-2.16	-4.88	-1.97	-4.46	-1.07	-2.91
(Days in open unempl.)/100	0.18	2.75	0.28	3.14	0.23	2.62	0.18	1.38
(Days in TP)/100	0.00	-0.21	0.00	0.16	0.02	1.46	0.00	-0.04
(Days in JC)/100	0.06	5.02	-0.10	-1.88	0.06	2.92	0.02	0.60
(Earnings -90)/10^5	0.03	0.17	-0.20	-0.82	-0.16	-0.71	-0.27	-0.90
(Earnings -91)/10^5	0.14	0.93	0.37	1.50	0.40	2.49	0.33	1.11
(Earnings -92)/10^5	-0.53	-3.21	-0.15	-0.65	-0.44	-2.17	-0.03	-0.09
Fraction in TP at PES	-0.02	-1.09	0.07	5.98	-0.02	-0.84	0.07	4.26
Fraction in JC at PES	0.04	4.18	0.01	0.62	0.04	3.03	0.02	1.09
# vacancies	0.05	0.09	-0.02	-0.02	-0.78	-1.02	0.80	1.06
# unemployed	-0.10	-1.23	-0.08	-0.65	0.00	0.01	-0.21	-1.54

The models also include dummy variables for region (at the county level) and registration month at the PES.

gressions (a maximum likelihood estimator is used).[14] The results pertain to the probability of entering a program during the first month after unemployment entry.[15] Columns headed "Unconditional" refer to the entire sample, while columns headed "Conditional on job" refer to a sample restricted to those finding employment. We use the latter estimates when analyzing search allocation.

It is reassuring to see that the unemployment histories and the variables

[14] Parameter estimates for other entry periods can be obtained from the authors upon request.

[15] All local characteristics are time-varying. This is also true for the social assistance indicator and the earnings of other household members. However, we introduce the latter two variables lagged once to avoid simultaneity bias.

measuring the activity at the local offices do a good job in predicting early program entry. The coefficients suggest, for instance, that individuals who have long previous spells of open unemployed are more likely to enter programs. Also there is something of an Ashenfelter dip in the data. The earnings of program participants are lower in the year preceding unemployment entry. It seems like the PES offices specialize in providing particular types of programs. Being registered at offices where the fraction placed in job creation (training) programs is high increases the individual probability of entering a job creation (training) program. The other (monthly varying) local variables have no effect on program participation. The lack of significance is probably due to the fact that the equations include county fixed effects.

UI eligibility at the start of the spell is a good predictor of program entry. Interestingly, the effects are the opposite for the two programs. On the one hand, being eligible for UI increases the probability of taking part in job creation programs; on the other hand, it reduces the probability of entering labor market training. Thus, it seems that training programs are used for retraining previous workers with obsolete skills to a limited extent.

Estimates such as those in Table 3 form the basis of the matching procedure. When constructing the matched comparison sample, we condition on the month of unemployment entry, in addition to matching on the estimated propensity scores. To give a sense about the quality of the matching procedure it is customary to report the absolute standardized bias (ASB) pre and post matching. Table 4 gives an example. It refers to the characteristics of JC participants in relation to the comparison group. The first column (from the left) of ASB statistics refer to the standardized difference in means while the second shows the standardized difference for *matched pairs*. Since there is a reduction in the ASB statistics matching takes care of some observed heterogeneity. Of course this will be true for any conditioning set as long as the characteristics included in that set are related to program participation. However, we have varied the conditioning set without significant reductions in the median absolute standardized bias ($MASB$).[16]

It is perhaps more interesting to investigate if match quality changes by duration until program entry. This issue is examined in Figure 3, which plots the $MASB$ at each time point. Match quality is reduced somewhat over the time period. The variance increases more than the level, however. This is to

[16] Matching reduces $MASB$ from 13.68 to 3.28 for training programs in the unconditional sample. For JC in the conditional sample there is a reduction from 13.38 to 2.80; for TP in the conditional sample the reduction is from 14.94 to 7.26. The common support requirement reduces the sample by: 7.6 % for TP in the unconditional sample; 10.4 % for JC in the conditional sample; and 12.6 % for TP in the conditional sample.

Table 4: Bias reduction due to propensity score matching. JC vs. comparison group, unconditional sample.

					Matched sample		
	JC	Comp.			JC	Comp.	
Variable	Mean	Mean	t-ratio	ASB^a	Mean	Mean	ASB^b
Female	0.54	0.43	8.46	21.67	0.54	0.53	1.36
Age/10	3.40	3.43	-1.47	3.76	3.40	3.44	5.29
Immigrant	0.22	0.16	5.83	15.51	0.21	0.22	3.32
High school ed.	0.55	0.54	1.19	3.05	0.54	0.55	0.79
University ed.	0.18	0.23	-4.78	11.86	0.22	0.20	3.36
No UI eligibility	0.14	0.20	-7.17	17.37	0.13	0.13	0.60
Cash assistance	0.04	0.07	-4.78	11.31	0.04	0.05	1.95
Single	0.53	0.50	1.65	4.23	0.53	0.52	2.70
# kids > 0	0.46	0.43	2.95	7.56	0.47	0.47	0.00
House owner	0.31	0.37	-5.76	14.49	0.31	0.33	3.93
(Household earnings)/10^5	0.57	0.65	-3.32	8.23	0.58	0.60	2.93
Social assistance receipt	0.18	0.12	5.85	15.70	0.20	0.20	1.89
# moves prior to -93 ≥ 1	0.19	0.18	1.38	3.56	0.19	0.18	1.31
No unemployment info.	0.27	0.42	-13.47	33.17	0.27	0.26	2.05
(Days in open unempl.)/100	1.06	0.68	11.88	32.19	1.06	1.10	3.14
(Days in TP)/100	2.69	1.63	5.48	15.09	2.62	2.87	3.66
(Days in JC)/100	1.15	0.41	6.92	20.42	0.95	1.35	11.02
(Earnings -90)/10^5	0.94	1.14	-12.23	29.76	0.95	0.97	2.72
(Earnings -91)/10^5	0.94	1.17	-13.03	31.59	0.95	0.95	0.16
(Earnings -92)/10^5	0.83	1.10	-15.16	36.58	0.83	0.85	2.49
Fraction in TP at PES	9.06	8.79	1.92	4.87	7.85	7.91	1.02
Fraction in JC at PES	18.58	18.17	2.23	5.70	22.21	22.07	1.90
# vacancies	0.52	0.54	-1.14	2.89	0.88	0.84	5.70
# unemployed	3.30	3.43	-1.46	3.72	3.42	3.39	0.78
# observations		1,857				1,841	
$MASB^c$				13.17			2.27

$^a ASB(k) = 100 \cdot \left| \frac{1}{n} \left[\sum_{i=1}^{n} x_i^k \right] - \frac{1}{N} \sum_{c=1}^{N} x_c^k \right| / \sqrt{(s^2(x_i^k) + s^2(x_c^k))/2}$

$^b ASB(k) = 100 \cdot \frac{1}{n} \sum_{i=1}^{n} \left| x_i^k - x_{c_i}^k \right| / \sqrt{(s^2(x_i^k) + s^2(x_c^k))/2}$

$^c MASB = \text{median}(\mathbf{ASB}),\ \mathbf{ASB} = (ASB(1), ..., ASB(\text{rows in table}))'$

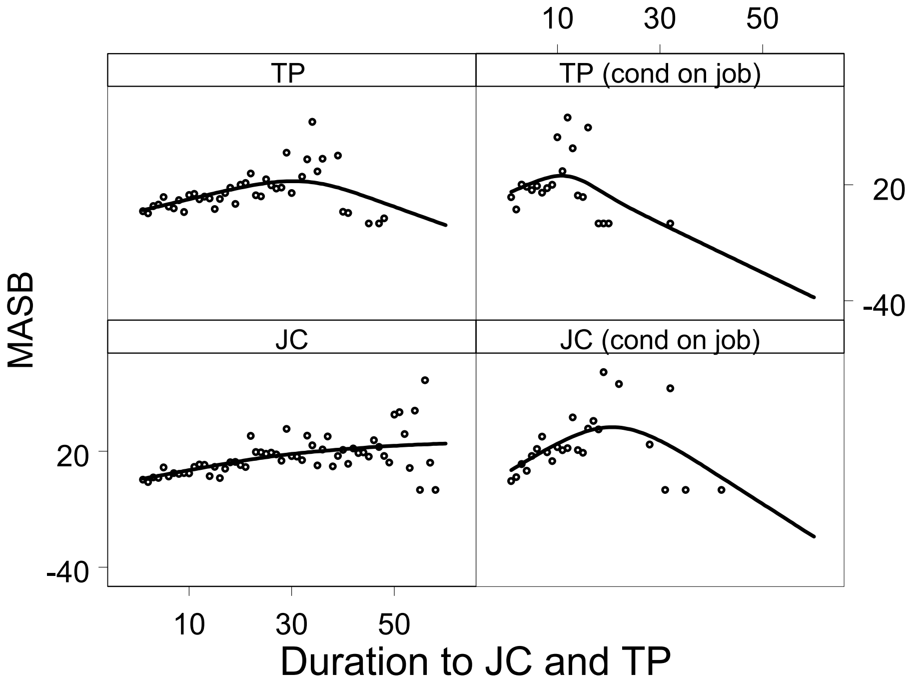

Figure 3: Match quality by duration until program start. (MASB=median absolute standardized biased. The line is a smoothing spline with 3 df.)

be expected since about three quarters of the training programs starts within the first year of an unemployment spell; see Figure 4. Problems associated with the apparent reduction in match quality should not be large, since the *MASB* is roughly constant for the first 12 months for training programs and for the first 24 months for job creation programs (accommodating 90 percent of the inflow into *JC*).

According to Figure 4, most programs start early on in the unemployment spell. The estimated treatment effects will mostly reflect this fact, i.e., programs starting early in the spell will be most influential in the estimate.

5.2 Employment

The first set of causal treatment effects is presented in Figure 5. The outcome of interest is the outflow to employment irrespective of location. The figure shows the difference in survival rates between treated individuals and matched controls. The upper panel refers to job creation programs and the

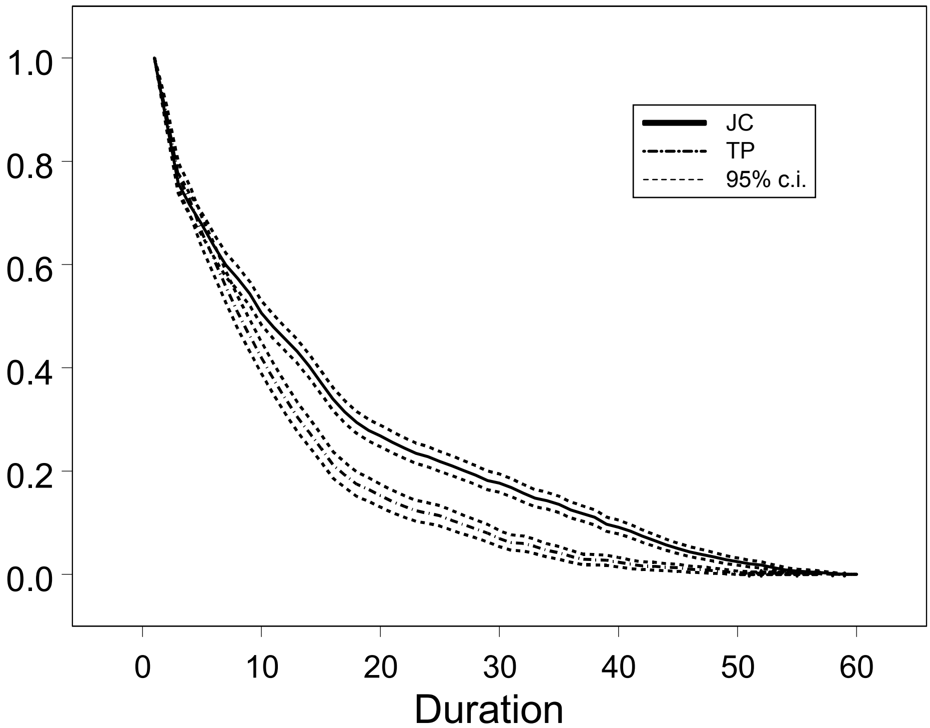

Figure 4: Distribution of program starts.

lower panel to training programs. The dotted lines in the figures are confidence intervals.[17]

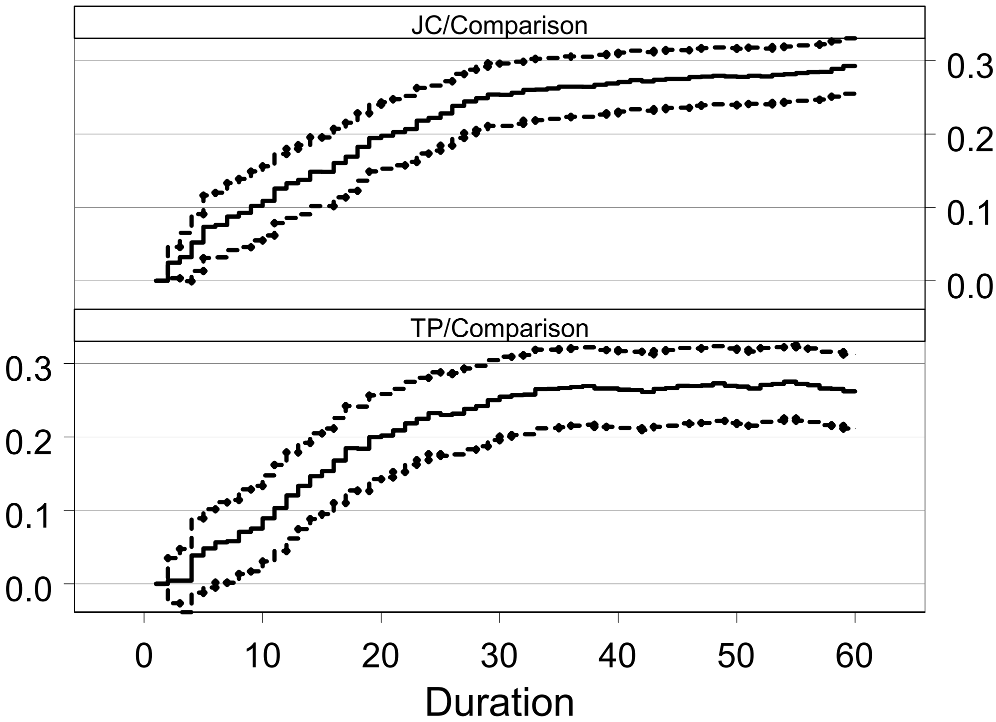

Figure 5: Causal treatment effects. Difference in survival rates. Risk: employment. (Dotted lines are 95 % confidence intervals)

The estimates suggest that participation in a job creation program reduces the cumulative outflow to employment by 29 percentage points. The corresponding number for training programs is a reduction by 26 percentage points. Relative to the cumulative outflow in the two comparison groups (0.73 for the JC comparison group and 0.68 for the TP comparison group) these estimates are remarkably similar. The relative decline associated with JC participation is 40 percent, while the decline associated with TP participation amounts to 39 percent. Thus, the size of the estimates have to be

[17] The confidence intervals are calculated as $\pm 2\sqrt{\mathrm{Var}S(t|D(t)=1)+\mathrm{Var}S(t|D(t)=0)}$, where $\mathrm{Var}S(s|D(t)=1)$ is calculated according to equation (19).

considered large. But notice that they are only about three quarters of the size suggested by the naive comparisons in Figure 1.

The slope of the two curves implies that the two programs have an effect on the job hazard mainly during the first two years after program entry. This may indicate that active search is vital in the initial period after unemployment entry. A somewhat speculative interpretation is that the programs fail because they reduce search activity initially which, in turn, impedes a successful reentry on the regular market.

From a relative comparison of the two programs by duration it appears that job creation programs are plagued by direct locking-in effects to a greater extent that training programs. The negative effect on the job hazard is immediately visible in the graph pertaining to JC; for TP, however, there is literally no effect during the first three months.

As we have emphasized repeatedly, the evidence presented in Figure 5 represent causal effects for those treated prior to a certain time point. The wider question is whether we can infer something about treatment on the treated as conventionally defined. If the treatment effects are of equal sign at all durations, then the analysis in Fredriksson and Johansson (2003) suggests that the long run effect (at $\overline{T} = 60$) in Figure 5 is lower in absolute size than the conventional effect of treatment on the treated.

5.3 Mobility

Now, let us turn to the program effects on contracted mobility. Figure 6 shows the differences in survival rates when the risk is employment outside the home region. In the longer run, both programs reduce the outflow to employment elsewhere. The cumulative outflow is reduced by 3.2 percentage points for job creation programs and by 2.6 points for training programs. Relative to cumulative contracted mobility in the comparison groups (0.049 for the JC comparison group and 0.044 for the TP comparison group) these estimates imply a reduction of 67 and 59 percent for JC and TP respectively.

Again the shape of the graphs are fairly similar for the two programs. There is one difference though: the effect on the contracted mobility hazard seems to disappear after 1.5 years in the case of training programs; this is not the case for job creation programs. Thus job creation programs appear to be plagued by negative long run effects to a greater extent than training programs. (An analogous pattern is visible in Figure 5). Having said this, we should emphasize that there are no statistical differences (at the 5 % level) between the two programs.

The evidence thus suggests that the programs reduce contracted mobility. Can we say anything about the mechanisms delivering this result? In

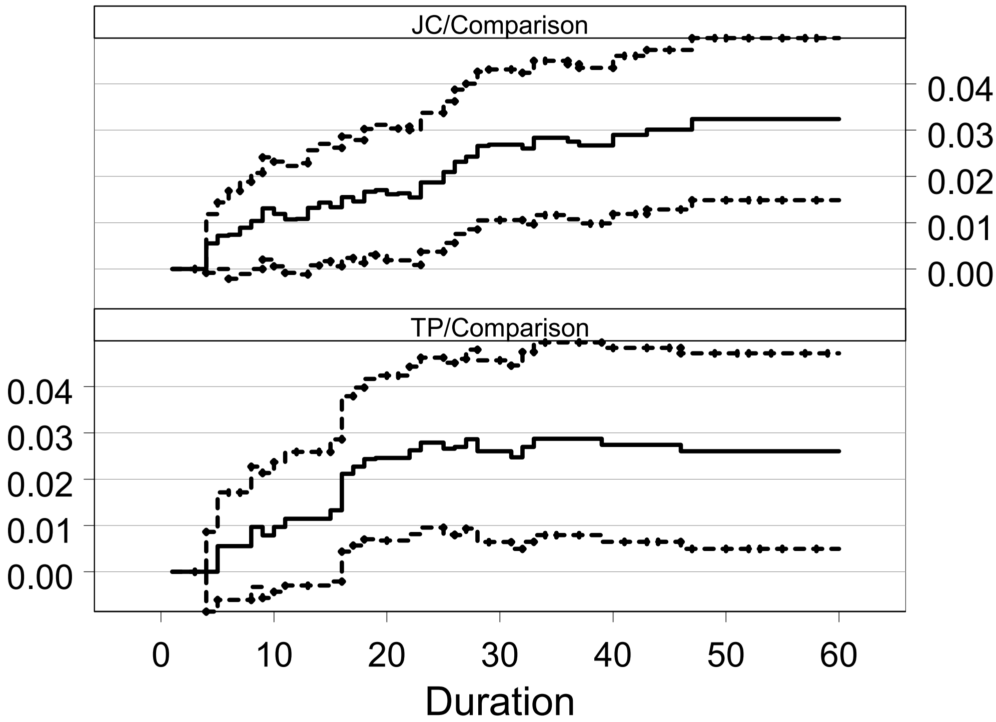

Figure 6: Causal treatment effects. Difference in survival rates. Risk: employment and mobility. (Dotted lines are 95 % confidence intervals)

the analytical framework of section 2 we assumed that the program effects may work along two dimensions: it may change the overall job finding probability and there may be a home bias in search allocation because program participation implies more contacts with local PES officers. So, is the result in Figure 6 due to the fact that programs reduce employment prospects in general or is it due to participants allocating less search outside the home region. Figure 7 presents some evidence pertaining to this issue. It presents differences in survival rates for the sub-sample who got a job. The risk, again, is employment outside the home region. The evidence suggests that there is an equal-sized reduction in outflows to jobs at home and outside the home region after taking part in either a job creation or training program. Thus, the evidence does not suggest that there is a home bias associated with program participation. Program participation reduces the outflow to jobs outside the home region because it lowers the overall contact frequency.

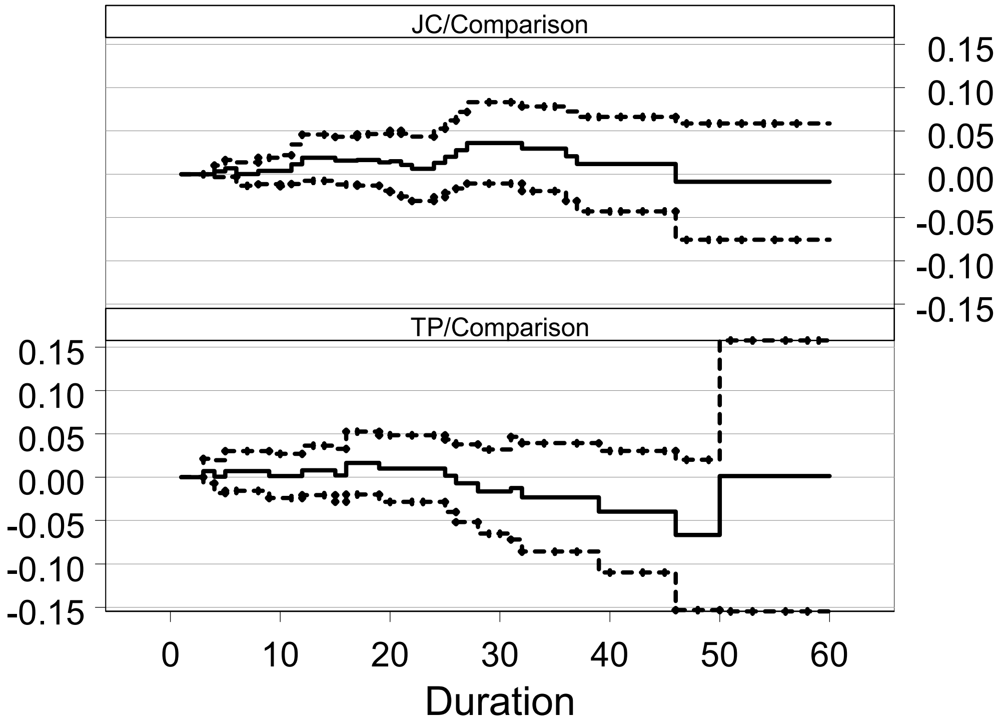

Figure 7: Causal treatment effects. Difference in survival rates. Risk: employment and mobility conditional on job. (Dotted lines are 95 % confidence intervals)

Another interesting issue related to our analytical framework is whether we can characterize the selection into programs. Ostensibly one can get at the answer by comparing, e.g., Figures 2 and 6. However, this is not right since Figure 6 features propensity adjustment *and* a different comparison population. The simplest way to characterize the selection into programs is to calculate a propensity-adjusted version of Figure 2.[18] Propensity adjustment reduces the long-run decline in contracted mobility from 3.4 to 3.2 percentage points for JC and from 3.2 to 2.5 percentage points for TP. Taking the analytical framework literally, this suggests that less mobile individuals are

[18] This graph is available upon request.

more likely to enter labor market programs. But the more general interpretation is that the probability of entering a program is higher for the least employable (in the observed sense).

6 Conclusion

In this paper we have analyzed whether participation in active labor market programs improve employment prospects and increase mobility in the longer run. We have conducted the analysis using unique micro data with, at least to some extent, novel estimation techniques.

The answers to the above questions turn out to be negative. Moreover, the picture is even more dismal than that. Participation in either of the two prototype programs – job creation and training programs – reduces the long-run probability of finding employment, irrespective of where it is located. These negative effects must be considered substantial on any metric. Relative to the overall outflow to employment in the comparison group, program participation reduces the outflow to employment by around 40 percent. The bulk of this effect occurs during the first two years after program entry.

With respect to geographical mobility, we find that both programs reduce the outflow to jobs outside the home region. Relatively speaking this effect is greater than the reduction of the outflow to employment in general. Still, the decline in contracted mobility appears to be driven by the fact that program participation reduce employment prospects in general. We find no evidence suggesting that labor market programs change the allocation of search across regions.

The negative effects of program participation on the overall outflow to employment are consistent with recent results pertaining to the effects of program participation in Sweden; see the extensive review in Calmfors et al. (2002). Larsson (2000), for instance, finds that training programs reduce employment and earnings for youths substantially. Moreover, Regnér (1997) concludes that the return to participation in training programs is in most cases negative.

The time period we are considering was a rather extreme period on the Swedish labor market. Unemployment rose dramatically over just a few years in the beginning of the 1990s. Concomitantly, program activity rose to unprecedented levels. Our results contributes to the weight of evidence suggesting that active labor market programs were not well-functioning at the time.

References

Bring, J. and K. Carling (2000), Attrition and Misclassification of Drop-outs in the Analysis of Unemployment Duration, *Journal of Official Statistics* **16**, 321–30.

Calmfors, L. (1994), Active Labour Market Policy and Unemployment: A Framework for the Analysis of Crucial Design Features, *OECD Economic Studies* **22**, 7-47.

Calmfors, L., A. Forslund, and M. Hemström (2002), Does Active Labour Market Policy Work? Lessons from the Swedish Experiences, IIES Seminar Paper 700, Institute for International Economic Studies, Stockholm University.

Dehejia, R.H. and S. Wahba (1999), Causal Effects in Nonexperimental Studies: Evaluating the Evaluation of Training Programs, *Journal of American Statistical Association* **94**, 1053–62.

Diamond, P.A. (1981), Mobility Costs, Frictional Unemployment, and Efficiency; *Journal of Political Economy* **89**, 798-812.

Edin, P-A. and P. Fredriksson (2000), LINDA – Longitudinal INdividual DAta for Sweden, Working Paper 2000:19, Department of Economics, Uppsala University.

Edin, P-A. and B. Holmlund (1991), Unemployment, Vacancies and Labour Market Programmes: Swedish Evidence, in F.P. Schioppa (Ed.) *Mismatch and Labour Mobility*, Cambridge University Press.

Fredriksson, P. (1999), The Dynamics of Regional Labor Markets and Active Labor Market Policy, *Oxford Economic Papers* **51**, 623-48.

Fredriksson, P. and P. Johansson (2003), Program Evaluation and On-going Program Starts, Working Paper 2003:1, Department of Economics, Uppsala University.

Heckman, J.J., H. Ichimura and P. Todd (1998), Matching as an Econometric Evaluation Estimator, *Review of Economic Studies*, **65**, $261 - 94$.

Imbens, G.W. (2000), The Role of the Propensity Score in Estimating Dose-response Functions, *Biometrika* **87**, 706-710.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data,* Cambridge: Cambridge University Press.

Larsson, L. (2000), Evaluation of Swedish Youth Labour Market Programmes, Working Paper 2000:6, Department of Economics, Uppsala University (forthcoming *Journal of Human Resources*).

OECD (2001), *Employment Outlook*, Paris.

Regnér, H. (1997), Training at the Job and Training for a New Job: Two Swedish Studies, Ph.D. Thesis, Swedish Institute of Social Research, Stockholm University.

Rosenbaum, P.R. (1995). *Observational Studies* (*Springer Series in Statistics)*, Springer Verlag. New York.

Rosenbaum, P.R. and D.B. Rubin, (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effect, *Biometrika*, **70**, 41 – 55.

van Ours, J.C. (2000), The Locking-in Effect of Subsidized Jobs, CEPR Discussion Paper 3489.

Westerlund, O. (1998), Internal Migration in Sweden: The Effects of Mobility Grants and Regional Labour Market Conditions, *Labour* **12**, 363-88.

Widerstedt, B. (1998), Moving or Staying? Job Mobility as a Sorting Process, Ph.D. Thesis, Department of Economics, University of Umeå.

Appendix: More details on sampling

Our basic strategy when sampling spells was to consider all new spells starting in 1993. We restricted the data by excluding individuals who participated in a vocational rehabilitation program or were classified as having a work related handicap in any period from the start of the unemployment register.

During a spell in the unemployment register, individuals are classified as belonging search categories (*cat*). With this information we excluded spells that we did not consider to be a spell of unemployment in the usual sense (i.e. that the individual is searching full time and can take a job immediately). On the basis on this consideration we excluded those who were part-time unemployed ($cat = 21$), temporarily employed ($cat = 31$), and on-the-job searchers ($cat = 41$) during their entire spell in the register.[19]

[19] For completeness we report the code which identifies these categories, although they may not have much meaning to the reader.

Individuals who were classified as being difficult/impossible to help find a job or a slot in an active labor market program, ALMP, ($cat = 14$ or $cat = 91$) during their entire spell have been deleted.

A period of part-time unemployment ($cat = 21$), temporary employment ($cat = 31$), and on-the-job search ($cat = 41$) ended the spell if the period lasted more than seven days.

If the individual started a spell with being classified as difficult/impossible to "place" in a job or an ALMP ($cat = 14$ or $cat = 91$) we deleted this initial period (i.e. the spell started when the individual moved to another state). The logic for doing this is that getting this classification initially and then moving to another search category often indicates that the individual is about to finish their education.

We also made a consistency check on the data. For each individual, the spell data were sorted by start date and end date. If the start date of a search category was prior to the previous end date, the observation was given the previous end date as a starting date. If the spell (so generated) implied a negative duration the observation was deleted. Also, spells with negative dates were deleted.

All in all, these transformations and exclusions resulted in the data containing 11,462 individuals starting their unemployment spell in 1993.

Table 5: Definitions of varibles used in the analysis.

Variable	Definition
Female	
Age	
Immigrant	=1 if born outside Sweden
High school education	=1 if attained upper secondary school
University education	=1 if undergraduate or postgraduate education
No UI eligibility	=1 if not eligible for UI or Cash assistance at start of spell
Cash assistance	=1 if eligible for Cash assistance at start of spell
Single	=1 if not married nor cohabiting
# kids > 0	=1 if at least one kid < 18 years of age
House owner	=1 if owning a house
Household earnings	the earnings of other household members
Social assistance receipt	=1 if member of social assistance receiving household
# moves prior to -93 ≥ 1	=1 if at least one move prior to 1993
No unemployment info.	=1 if not in unempl. register 1 Aug. 1991 – 31 Dec. 1992
Days in open unempl.	days in open unemployment 1 Aug. 1991 – 31 Dec. 1992
Days in TP	days in training program 1 Aug. 1991 – 31 Dec. 1992
Days in JC	days in job creation program 1 Aug. 1991 – 31 Dec. 1992
Earnings -90	individual earnings in 1990
Earnings -91	individual earnings in 1991
Earnings -92	individual earnings in 1992
Fraction in TP at PES	share of unempl. in training programs at the local PES
Fraction in JC at PES	share of unempl. in job creation programs at the local PES
# vacancies	number of vacancies in the local labor market
# unemployed	number of unemployed in the local labor market