

IZA DP No. 5596

The Relative Efficiency of Active Labour Market Policies: Evidence from a Social Experiment and Non-Parametric Methods

Johan Vikström
Michael Rosholm
Michael Svarer

March 2011

The Relative Efficiency of Active Labour Market Policies: Evidence from a Social Experiment and Non-Parametric Methods

Johan Vikström

IFAU-Uppsala and UCLS

Michael Rosholm

Aarhus University and IZA

Michael Svarer

Aarhus University, CAM and IZA

Discussion Paper No. 5596

March 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Relative Efficiency of Active Labour Market Policies: Evidence from a Social Experiment and Non-Parametric Methods^{*}

We re-analyze the effects of a Danish active labour market program social experiment that included a range of sub-treatments, including monitoring, job search assistance and training. Previous studies have shown that the overall effect of the experiment is positive. We apply newly developed non-parametric methods to determine which of the individual policies that explains the positive effect. The use of non-parametric methods to separate sub-treatment effects is important from a methodological point of view, since the alternative, namely parametric/distributional assumptions, is in conflict with the concept of experimental evidence. Our results are highly relevant in a policy perspective, as optimal labour market policy design requires knowledge on the effectiveness of specific policy measures.

JEL Classification: C14, C41, C93

Keywords: active labour market policy, treatment effect, non-parametric bounds

Corresponding author:

Michael Rosholm
Department of Economics and CIM
Aarhus School of Business, Aarhus University
Hermodsvej 22
DK-8230 Aabyhøj
Denmark
E-mail: rom@asb.dk

^{*} We would like to thank Jon Gemus, Caroline Runeson, Erik Grönqvist and seminar participants at IFAU-Uppsala and the ELE Workshop in Bornholm helpful comments. We are grateful to the National Labour Market Board of Denmark for making the data available and for the assistance they provided in preparing the data and resolving various issues. Rosholm acknowledges support from the Strategic Program for Welfare research and from the Social Science Research Council (grant no. 275-05-0273).

1 Introduction

During the course of a single unemployment spell, the public employment service (PES) often assigns an unemployed worker to a number of different labour market policies, such as monitoring and sanctions (e.g. in the form of meetings with case workers), job search assistance, and different activation programmes. Efficient construction of such combinations of policies requires knowledge of both the combined effect of certain policy packages and on the effect of the individual policy instruments. Using a Danish social experiment consisting of a combination of intensive monitoring, job search assistance, and active labour market programmes, we provide information on the effects of the individual policy components using only non-parametric methods. Even in a social experiment, inference on specific policies implemented some time after randomisation is contaminated by dynamic selection bias, and often parametric methods have been applied to circumvent this problem. However, parametric methods are fundamentally in conflict with the spirit of the experimental approach, which is to provide the required evidence non-parametrically. Hence, the provision of non-parametric evidence of specific policy instruments is the main contribution of our paper. This is of great policy relevance, as it is crucial information to policy makers for designing labour market policy packages that meet the aim of reducing structural unemployment rates in a cost-minimizing way.

The experiment analyzed in this paper was conducted in two counties in Denmark during the winter of 2005-6. All individuals in the two counties who became unemployed during this period were randomized into either a treatment or a control group based on their birthday. The treated experienced a dramatic intensification of labour market policies in terms of early and mandatory participation in a job search course, very frequent meetings with employment officers, and program participation after a few months in open unemployment. The experimental protocol dictated that these policies should be applied at specific unemployment durations. We analyze the results from this experiment, and study both the overall effect of

the full treatment package, and the effects of the individual sub-treatments: job search assistance, monitoring, and (activation) program participation.

The overall effect of the experiment requires identification of the overall difference in the transition rate, whereas the effect of the individual policies requires identification of the effect on the conditional transition rate out of unemployment during the specific week(s) when the specific policy is applied. Since the policies were applied at pre-specified unemployment durations, individual policy effects can with some certainty be represented by the evolution of the treatment effect over the unemployment duration. As an illustration, say that one is interested in the effect of the experimental job search assistance program. One would then like to compare the conditional transition rates out of unemployment for the treatment and control group during the week(s) when the experimental protocol dictates that job search assistance is taking place (and possibly the following weeks).

The distinction between the overall effect and the effects of the individual policies gives rise to interesting identification issues. The overall effect could be estimated by considering the effect of the experiment on the cdf of the distribution of the time to transition or its inverse, the quantile function. Graversen and van Ours (2008a) provide Kaplan-Meier estimates on the effect on the cdf (survival function) for the Danish experiment considered in this paper. They find that the experiment had large positive effects in terms of shorter average unemployment duration. From the effect on the cdf we cannot, however, obtain the effect on the conditional transition probabilities. This is because of the so-called dynamic selection problem. Random assignment of treatment only ensures comparability of the treatment and control groups at the time of randomisation. At later times treated units with characteristics that have a positive/negative interaction effect with treatment on the transition probability leave the initial state first/last, so that these characteristics are under/over represented among the treated relative to the controls, and this confounds the effect of the treatment. For instance, up until the time of the job search course, some individuals have left the treatment

group and some have left the control group, such that the two groups are no longer identical, and the randomisation imposed at the time of inflow to unemployment is no longer present. In order to identify the effect of the individual policies, one would have to correct for this dynamic selection.

One approach to solve the dynamic selection problem is to rely on semi-parametric models. As shown by Elbers and Ridder (1982) the Mixed proportional hazard (MPH) model is semi-parametrically identified, so that if the multiplicative structure is maintained, identification does not rely on arbitrary functional form or distributional assumptions. The parametric MPH model is used in Graversen and van Ours (2008b) and Rosholm (2008) to estimate the effects of the individual sub-treatments used in the Danish experiment. Indeed, both studies conclude that the meetings with caseworkers (or the perception that you had to meet with a caseworker) were the most effective among the specific policy instruments. However, the estimated effects of specific policy instruments are only valid if the parametric assumptions underlying the model are correct. One may argue that in the case of a social experiment, where the aim is to provide non-parametric evidence, parametric assumptions are especially troublesome. Ideally, we would prefer to rely solely on the experimental variation to estimate the treatment effects of interest.

In this study, we avoid parametric assumptions, by using newly developed non-parametric methods by Ridder and Vikström (2011). They derive non-parametric bounds on the effect on the conditional transition rate that are solely based on the random assignment and not on arbitrary parametric assumptions. The bounds provide the causal effect on the conditional transition rate in each time period, i.e. the evolution of the effect of the experiment over unemployment duration. This gives us information on the effect of the individual labour market policies. Ridder and Vikström (2011) also provide bounds under additional weak assumptions that give more tightly identified bounds. We show that these bounds are highly informative on the performance of the individual policies.

We find interesting results. We find significant effects of intensified job search

assistance, frequent meetings with caseworkers, and pre-programme or threat effects of program participation. The bounds on the activation program effect are wide since the dynamic selection at that stage is large. We have also tested for heterogeneous treatment effects by age and sex. Considering the rather limited experimental evidence on the performance of different labour market policies these results based on a social experiment provide important insights.¹ In this respect disentangling the effects of the individual policies are important for designing the most efficient labour market policy package. Our results are also interesting since 'the Danish model' for the labour market has received much attention lately, because of its perceived effectiveness in maintaining high participation rates and low unemployment rates. This successful performance is often attributed to the 'Flexicurity' model, with flexible hiring and firing rules and regulations combined with generous income transfer systems and very active labour market policies. In the 1980s, when unemployment rates were persistently high in Denmark, the first two components of the Flexicurity model - flexibility and security - were already in operation, but active labour market policies were only in its infant stages and not nearly as intensive as they are today. Therefore, many have seen these intensive policies as a pivotal component in the Flexicurity model (see e.g. Andersen and Svarer, 2007). Our results provide non-parametric evidence to indicate that this might indeed be the case.

The remainder of the paper is organized as follows; sections 2 and 3 discuss the Danish labour market and the social experiment. Section 4 presents the data and provide descriptive statistics. In section 5 we discuss the problem of dynamic selection and the methods we use to handle it. Section 6 presents the results and section 7 concludes.

¹The surveys in Kluge (2009) and Card et al. (2009) show that experimental evidence on training programmes is scarce. Job search assistance and monitoring experiments have previously been analyzed by e.g. Meyer (1995), Ashenfelter et al. (2005), Dolton and O'Neill (2002) and Van den Berg and van der Klauuw (2006). Experimental evidence on pre-programme effects are found in Black et al. (2003) and Hägglund (2006).

2 The Danish labour market

The Danish labour market is highly organized in the sense that almost 80% of workers are union members and members of an unemployment insurance fund, giving them access to fairly generous UI benefits should they become unemployed. The remainder of workers are eligible for social welfare in case of unemployment provided that they have no wealth to draw upon. Wage negotiation was decentralized in the late 1980s and early 1990s, such that today mostly broad issues such as minimum wage levels (and increases therein), hiring and firing rules, holidays, safety at work, etc., are centrally negotiated. This has resulted in a system where there are quite low costs of labour turnover (hiring and firing) but fairly generous income transfers to compensate the workers for income losses during periods without work. Such a system is conducive to high rates of labour turnover, and indeed nearly one quarter of the labour force switches jobs during any year (see Frederiksen and Westergaard-Nielsen, 2007). In such a flexible labour market, it is important to have labour market policies in place that support the ability of workers to make fluent transitions between different jobs. That is, the skill enhancing and job search assistance components of labour market policies are quite important. The generous income transfer schemes (including UI benefits) necessitate a labour market policy which provides the unemployed with sufficient incentives to search (given that replacement rates are quite large), thus emphasizing the importance of frequent monitoring (of job search activities) and sanction policies, as well as the importance of certain workfare arrangements to tax away some of the leisure otherwise available to unemployed workers.

In many ways the Danish system for the labour market relies heavily upon the effectiveness of its labour market policies, and therefore experiments such as the one analyzed here make up a crucial element in the collection of empirical evidence on how to best design these policies.

3 The Experiment

The social experiment considered in this paper was carried out in the counties of Southern Jutland and Storstrøm. Workers eligible for unemployment insurance (UI) benefits, who became unemployed during the period from November 1, 2005 to February 28, 2006, were randomly assigned into a treatment and a control group. Those born on the 1st to the 15th were assigned to the treatment group, while those born on the 16th to the 31st were assigned to the control group.²

Those in the treatment group received a letter after approximately 1.5 weeks in open unemployment telling them about the new policy regime. They were not told explicitly of the randomized experiment, rather they were told that they were taking part in a 'pilot study' regarding a new labour market policy regime. They were also told about the contents of the new labour market policy regime. We label the period before the letter as the *pre-treatment* regime, and the period after the letter the *new policy letter* regime.

After 5-6 weeks of unemployment, those in the treatment group were supposed to participate in a two-week Job Search Assistance (JSA) programme, intended to determine the capabilities of the unemployed worker and helping her to search for jobs. This is labeled the *JSA programme* regime. Thereafter, they should meet frequently with a case worker in order to ensure that they are searching actively and in order to assist them in their job search.³ There were three types of meetings; 1) job-plan meetings, during which an action plan for getting back to work is formulated, 2) contact-meetings, in which the progress and job search strategies of the unemployed worker are discussed, and finally 3) job assignment meetings, during which the case worker informs the worker about an open vacancy.⁴ We label this the *frequent meetings* regime.

²It was not possible to escape treatment by leaving unemployment for a short period, since persons in the treatment group who return to unemployment during the period of the experiment will enter the experiment at the stage where they left it.

³In the country of Southern Jutland, meetings would take place every fortnight, and in the county of Storstrøm meetings would take place each week.

⁴Besides these meetings job interviews organized at a firm with a vacancy is registered as a meeting, even if the case worker typically does not participate during the interview.

Those who have not found employment after four months of unemployment had to participate in an unspecified programme of at least three months' duration. Case workers had some discretion in how to allocate workers to programmes. There were 5 types of programmes; the JSA programmes described above, private sector temporary employment subsidy jobs, temporary employment within the public sector, classroom training programmes ranging from IT courses to courses in brick-laying, to truck drivers license courses, and finally vocational training programmes in firms.

These programmes may have different effects. Full-time programme participation implies a great loss of utility from leisure. Some of the unemployed workers may therefore perceive programme participation as a 'threat', which gives them incentives to increase their job search efforts prior to enrolment into the programme. We therefore label the 3 weeks prior to program participation the *pre-programme* regime. While enrolled, the unemployed have less time to search for employment, in many cases leading to well documented lock-in effects. If the programme is successfully in providing the unemployed with new abilities, we expect a positive effect once programme participation is completed. In accordance, we label the first three weeks during which the unemployed is supposed to be in a programme as the *start programme* regime, and the rest of the period in programmes as the *later programme* regime. Finally, we label the period after programmes as the *post-treatment* regime, since this also marks the end of the experiment and the return of the treatment group to the ordinary labour market policy trajectory followed by the control group.

Persons in the control group were subjected to the ordinary labor market policy implying that they have to attend meetings with case workers once every three months, and after one year of unemployment they are required to participate in an unspecified program of unspecified duration. They had a right to participate in a programme of their own choice for up to six weeks during the first year in open unemployment. The time span between the end of one programme and the start

of the next after one year of unemployment should be shorter than six months.

The experiment thus involves eight different treatment regimes, summarized in table 1. Although the treatments are quite precisely defined, as in all social experiments there is still some scope for discretion by region and by case worker. We also have to take into account that the control group outcomes are not the result of non-treatment but rather of a different, less intensive treatment. The exact meeting intensities and programme participation intensities among the treatment and control groups are analyzed in detail in Rosholm (2008). Here we present an overview of the most important findings. Three points are important. First of all, the activity level in terms of meetings and programme participation is considerably higher in Storstrøm county than it is in Southern Jutland county. The general pattern is that in Storstrøm county the experiment has been conducted more in line with the experimental protocol, while in Southern Jutland, enforcement has been more lax. There are, in essence, two different experiments, one in each region, and they should therefore be analyzed separately. Second, none of the counties come close to the prescribed treatment intensities in the treatment group. For example, in Storstrøm county the unemployed on average met with a case worker only every 2.5 weeks on average, where they should have done so each week. Third, even if the activity intensities are lower than prescribed by the treatment protocol, the increases in JSA, meetings, and programme participation starts at the prescribed unemployment durations, and there are large differences in treatment intensities between the treatment and control groups in both counties. This implies that we can use the experimental protocol treatment regimes, but we have to keep in mind that not everyone in the treatment group receives all treatments even if they stay unemployed. We will therefore be estimating the intention-to-treat (ITT) effects of the experiment.

4 Data, sample selection, and descriptive statistics

The data used for this study come from administrative databases used by the case workers to register their dealings with the unemployed workers, from the central register on the labour market (CRAM) that registers UI benefit payments, and various other administrative registers. All data are collected by The National Labour Market Board and made available to the research community. The data contain weekly information on the type of transfer received and the activities undertaken in each week (meetings, programmes, etc.) for each worker.

The total inflow into open unemployment in the period between week 43 of 2005 and week 8 of 2006 was 5180 individuals. We remove individuals that are not assigned correctly to the treatment and control groups based on an inspection of their birthday as well as individuals who die or emigrate during the observation period. The removal of the latter is done since we do not know the exact dates of the events. We also remove individuals who do not receive any UI benefits during the first six weeks of their unemployment spell. This is done because the receipt of UI benefits defines the unemployment spell. The reason for choosing a six-week window is that we want to include persons who have just left education or who have quit their job themselves, and these events involve a five-week sanction period, in which the person may not receive UI benefits.⁵

The resulting sample has 4513 individuals. In the sample selection process we lose 15% in Southern Jutland county, and 12% in Storstrøm county. This is mainly due to exclusion of non UI recipients. The fraction of removed treatments and controls in Southern Jutland (Storstrøm) are 14.% and 15.7%, respectively (10.5% and 11.2%), i.e. no remarkable differences. We therefore see no reason for questioning the validity of the experiment.

⁵Some are unemployed due to 'bad weather' and due to 'work-sharing arrangements', and since these categories of unemployed workers are not intended to take part in the experiment, they are also removed from the sample.

For each individual we determine the length of the unemployment spell that triggered their participation in the experiment. The unemployment spell is considered to have ended when the individual has not received UI benefits or some related benefit for more than four consecutive weeks. We divide the unemployment duration into eight intervals corresponding to the eight different treatment regimes discussed above. Besides the information used to construct unemployment spells, the data contain some further information that may be used to construct control variables, or, as in this paper, to estimate treatments effects for different sub-populations. We have information on starting week, gender, age, ethnicity, unemployment insurance fund membership, and fraction of time spent on public income transfers during different time spans before becoming unemployed.

Table 2 reports descriptive statistics for the variables we exploit in this study, age and gender, by region and treatment status. As expected, due to randomization we find no significant differences between the treatment and control group for these, nor for any of the possible control variables (see Rosholm, 2008). There are, however, minor regional differences. Thus, the sample selection process does not seem to have invalidated the experimental nature of the data.

5 Econometric specification

5.1 The identification problem

We are interested in estimating: (*i*) the combined effect of the Danish experiment policy package and (*ii*) the effect of the individual monitoring, job search assistance and programme policies. The overall effect of the experiment requires identification of the overall difference in the transition rate or in other words the effect on the average unemployment duration, whereas the effect of the individual policies requires identification of the effect on the conditional transition rate at the specific week when the policy is applied.

To be precise, what we are aiming to identify is the time pattern of the overall

effect of the experiment, which could be interpreted in terms of treatment regime effects. It will provide insightful information about the effects of the individual policies. In a strict sense there are, however, three potential limitations with this approach. First, for all treatment regimes after the first we can only hope to estimate the effect given that the individual already have been subject to earlier treatments. Second, any important delayed effects from previous treatments make the interpretation of the later treatment regime effects more difficult. Third, the experiment only provide information on the effects of the individual policies given at a particular week. However, given the quite limited experimental evidence on the effects of individual ALMP policies we are still confident that the time pattern of the overall effect provide very interesting insights into the effects of the individual ALMP policies.

More importantly, the distinction between the combined effect and the effect of the individual policies gives rise to interesting identification issues. In this subsection we describe and discuss these identification issues in detail, using the familiar Mixed Proportional Hazards (MPH) parametric model.⁶ In the next subsection we then describe the non-parametric methods we use to address these identification problems.

In the case of a binary treatment the MPH model specifies the individual hazard or transition rate $\theta(t, d, V)$ as

$$\theta(t, d, V) = \lambda(t)\gamma(t)^dV,$$

with t the time spent in the origin state, $\lambda(t)$, the baseline hazard, d , the treatment indicator taking the value one if treated and zero otherwise, and V , a scalar nonnegative unobservable that captures population heterogeneity in the transition rate.

The nonnegative function $\gamma(t)$ captures the effect of the experiment at time

⁶The MPH model is a duration model for continuous time, whereas our outcome is discrete with an exit in one of the eight treatment regimes. However, since the MPH model is so widely used (see Van den Berg, 2001) we still use it as illustration of the identification problem.

point t . If applied to the experiment analyzed in this paper one could think about the experiment as one overall treatment, and $\gamma(t)$ as capturing the effects of the different sub-treatments. For instance, during week 5-9, i.e. in the JSA treatment regime, $\gamma(t)$ captures the effect of the JSA programme.

Define the time to transition T^d . The MPH model implies that T^d the hazard/transition rate given the treatment history is

$$\theta(t|d) = \lambda(t)\gamma(t)^d \mathbb{E}_V [V|T^d \geq t].$$

Consider making a naive comparison of the transition rates for the non-treated

$$\theta(t|d = 0) = \lambda(t)\mathbb{E}_V[V|T^0 \geq t]$$

and the treated

$$\theta(t|d = 1) = \lambda(t)\gamma(t)\mathbb{E}_V[V|T^1 \geq t]$$

Then, because most likely

$$\mathbb{E}_V[V|T^1 \geq t] \neq \mathbb{E}_V[V|T^0 \geq t],$$

the effect of the intervention on the transition rate is confounded by its differential effect on the distribution of the unobservable among the treated and controls. The intuition behind this result is that the difference between the treated and control transition rates is proportional in V , and this difference determines the survival probability. If $\gamma(t) > 0$ for all values of V , the survival probability is smaller for the treated than for the controls and the difference is largest for large values of V . Therefore the average V among the survivors will be smaller for the treated than for the controls and this implies that the comparison of the average transition rates of the treated and controls is confounded by this difference, which is labeled dynamic selection. This dynamic selection or survivor bias is not just a feature of the MPH model. It is present in any population where the treatment and the

individual characteristics interact to increase or decrease the transition probability.

Parametric and semi-parametric models for the transition can be used to correct for the survivor bias in the average treatment effect. In a fully specified MPH model, we specify a distribution for V , so that we can estimate $\mathbb{E}_V [V|T^0 \geq t]$ and $\mathbb{E}_V [V|T^1 \geq t]$ to obtain the correction factor. However, such an approach requires us to maintain the multiplicative specification with a proportional unobservable. As argued by e.g. Van den Berg (2001), economic models for the hazard rate are typically not multiplicative. That is one reason why avoid functional form assumptions.

5.2 Non-parametric bounds

The previous section showed that if we do not want to specify a model for the transition rate we need to find another way to maintain the comparability of the treatment and control groups over the spell. The approach that we take in this paper is to consider average transition rates, where the average is taken in the same population for both treated and controls. We follow Ridder and Vikström (2011) (RV henceforth) and focus on additive average treatment effects, where the average is taken over the subpopulation of individuals who would have survived until a certain time point in both the treatment and control group. The individuals in this population have the same survival experience and any difference between the transition rates must be due to the effect of the treatment.

Our outcome of interest is the transition out of unemployment in either one of the eight discrete treatment regime intervals. Formally, let the potential outcome Y_k^d be an indicator of a transition during treatment regime k if treatment is assigned ($d=1$) or if no treatment ($d=0$) is assigned. Since there is only randomization at time of inflow to unemployment, the effect of the intervention on the transition rate in treatment regime k for the hypothetical subpopulation of individuals who would have survived until k under both treatment and no-treatment

is

$$\begin{aligned} \text{ATES}_k = & \\ & \mathbb{E}[Y_k^1 | Y_{k-1}^1 = 0, \dots, Y_1^1 = 0, Y_{k-1}^0 = 0, \dots, Y_1^0 = 0] - \\ & \mathbb{E}[Y_k^0 | Y_{k-1}^0 = 0, \dots, Y_1^0 = 0, Y_{k-1}^1 = 0, \dots, Y_1^1 = 0] \end{aligned}$$

Following Rubin (2000), Abbring and van den Berg (2001) and RV, this can be called the *Average Treatment Effect on the Survivors*. Note that there is dynamic selection and the ATES_k controls for that by comparing the transition rates for individuals with a common survival experience.

RV derive bounds on ATES_k under sequential randomization. Sequential randomization means that in each time period, a subset of the survivors up until this period are randomized into treatment. The treatment assignment in the Danish experiment is a special case of sequential randomization, where the probability to be selected for treatment is zero in all time periods after the first.

RV show that only certain ATES_k effects are point identified. From the expression above, this is directly apparent. In the first treatment regime period there are only survivors, so that the randomization ensures that ATES_1 is point identified. For $k > 1$ this no longer holds, except in the special case when everyone survives the first $k - 1$ periods. This means that ATES_k for $k > 1$ cannot be point identified.

Let D_k be the indicator that treatment is assigned in period k , and Y_k the observed outcome as an indicator of a transition in period k . Note that in our case $D_k = 0$ for $k > 1$. Using the results in RV⁷ and imposing this restriction we have

⁷The only additional assumption we have to make is independent censoring. Define C as the censoring time. Under independent censoring we e.g. have $\Pr(Y_2 = 1 | D_1 = 1, Y_1 = 0) = \Pr(C = 2) \Pr(Y_2 = 1 | C = 2, D_1 = 1, Y_1 = 0) + \Pr(C > 2) \Pr(Y_2 = 1 | C > 2, D_1 = 1, Y_1 = 0) = \Pr(Y_2 = 1 | C > 2, D_1 = 1, Y_1 = 0)$, so that the average observed transition for the uncensored observations could be used instead of the average transition rate for all individuals. Note that the censoring rate is zero at week 20 and 1.4 % at week 30 (see Table 2). Given this and the fact that the censoring rate is quite similar across the treatment and control groups we think that this assumption is unproblematic.

that the outcomes in the second treatment regime ($k = 2$) is bounded by⁸

$$\begin{aligned}
& \max \left\{ \frac{[\Pr(Y_2 = 1|D_1 = 1, Y_1 = 0) - 1] \Pr(Y_1 = 0|D_1 = 1)}{\Pr(Y_1 = 0|D_1 = 1) + \Pr(Y_1 = 0|D_1 = 0) - 1} + 1, 0 \right\} - \quad (1) \\
& \min \left\{ \frac{\Pr(Y_2 = 1|D_1 = 0, Y_1 = 0) \Pr(Y_1 = 0|D_1 = 0)}{\Pr(Y_1 = 0|D_1 = 1) + \Pr(Y_1 = 0|D_1 = 0) - 1}, 1 \right\} \leq ATES_2 \leq \\
& \min \left\{ \frac{\Pr(Y_2 = 1|D_1 = 1, Y_1 = 0) \Pr(Y_1 = 0|D_1 = 1)}{\Pr(Y_1 = 0|D_1 = 1) + \Pr(Y_1 = 0|D_1 = 0) - 1}, 1 \right\} - \\
& \max \left\{ \frac{[\Pr(Y_2 = 1|D_1 = 0, Y_1 = 0) - 1] \Pr(Y_1 = 0|D_1 = 0)}{\Pr(Y_1 = 0|D_1 = 1) + \Pr(Y_1 = 0|D_1 = 0) - 1} + 1, 0 \right\}.
\end{aligned}$$

Note that these bounds do not impose any assumptions beyond random assignment. The bounds are sharp in the sense that there exist feasible joint distributions of the potential outcomes which are consistent with the upper bound and the lower bound. It gives an upper and lower value on the sub-treatment effects that is consistent with the available data and expressed in terms of transition probabilities that are directly observable in the data. The above bounds are for the second period, but bounds for later time periods follow the same pattern. The interested reader is referred to RV for derivations and detailed expressions. The intuition behind the bounds is that even if the data does not contain enough information to point identify the treatment effect of interest, the data often places an upper bound on how much the dynamic selection can confound the treatment effect of interest.

Besides the general bounds that impose no assumptions beyond random assignment, RV also explore the identifying power of additional weak assumptions. They show that weak assumptions may potentially tighten the bounds considerably. We apply these bounds, but postpone the discussion of these assumptions until section 6.3.

⁸This holds if $\Pr(Y_1 = 0|D_1 = 1) + \Pr(Y_1 = 0|D_1 = 0) - 1 \geq 0$, since otherwise $\mathbb{E}[Y_2^{11}|Y_1^1 = 0, Y_1^0 = 0]$, $\mathbb{E}[Y_2^{00}|Y_1^1 = 0, Y_1^0 = 0]$ and $ATES_2$ are not defined.

6 Results

6.1 Overall effect of the experiment

We first present results on the overall effect of the experiment. Table 2 presents the mean unemployment duration for the control group and treatment group in the two counties. Due to censoring of longer unemployment durations, we also report the truncated mean duration (truncated after 280 days of unemployment). These two measures produce very similar results. They show that the intensified policies provided by the experiment on average reduce the time in unemployment by about two weeks. It is the same result as in Graversen and Van Ours (2008a, 2008b) and Rosholm (2008).

The fact that intensified policies can reduce the time in unemployment is an important result. However, from a policy perspective it is crucial to investigate which one of the different policies were (most) effective, since the least effective may possibly be removed, and since they vary considerably in their costs, with the meetings being by far the cheapest. In the next subsection we address this more difficult question using non-parametric methods.

6.2 Which policies work?

Tables 3 and 4 present the upper and the lower bound on $ATES_t$ for the different treatment regimes for Storstrøm and Southern Jutland county, respectively.⁹ The upper left panel of figures 1 and 2 display the same bounds. As implied by the identification results, the treatment effect in the first period is point identified. As expected we find no significant differences between the control groups and treatment groups in this pre-treatment regime period; it is before the treated were informed about the new policy regime. For the other treatment regimes, the bounds on the average treatment effects are quite wide. For all of the following treatment regimes we can neither rule out a positive nor a negative effect of the

⁹Note that we do not report confidence intervals. In principle these can be constructed using the approach in Hahn and Ridder (2010).

experiment. However, note that for some of the earlier treatment regimes the bounds are nevertheless informative about the average treatment effect. Still, we must conclude that without further assumptions the experiment cannot inform us which of the individual policies were the most effective. Fortunately, RV provide bounds under a couple of additional weak assumptions. In the next section, we discuss these assumptions and present the bounds they deliver.

6.3 Weak assumption bounds

RV present results for three weak assumptions: monotone treatment response, common shocks, and positively correlated outcomes. Here we provide an intuitive discussion of these assumptions and refer to RV for an exact specification of the assumptions.

The assumption of a positively correlated outcomes (PCO) implies that some unemployed individuals have a higher probability of finding employment compared to other unemployed no matter whether they are in the control group or the treatment group. This assumption is fulfilled if the ranking of the individuals in terms of the characteristics that determines job offers, such as experience and job search effort, stays the same during the entire job search period. More precisely, in terms of the ALMP experiment the assumption means that the ALMP treatment should not cause some individuals that quickly found a job without treatment to suddenly slowly find a job, and vice versa. We see no important reasons to why there should be such flipping behavior, especially since we are focusing on relatively short unemployment durations (less than 40 weeks).

The common shocks (CS) assumption is related to the way duration models are formulated. There are two pieces of information that determines if an individual leaves unemployment in a certain time period: the individual transition rate and a person-specific shock that determines if the transition is realized or not for that person. As illustration, one individual may be likely to leave in a certain period, but draws a negative shock and by chance finds no employment, and another

individual may be unlikely to leave in a certain time period but is lucky and draws a positive shock and thus finds employment anyway. One reason for why the general bounds are wide is that they allow for different shocks under treatment and under no treatment. In the unemployment example in this paper this means that the treatment not only systematically alters the transition rate out of unemployment but also affects the random events that determines if a transition is realized or not. For the Danish ALMP experiment for unemployed it is difficult to see why this should be the case. The shocks assumption on the other hand means that the individual is exposed to exactly the same shock under both treatment and no treatment. We see no reason for doubting the validity of this assumption.

The monotone treatment response (MTR) assumption has been explored by e.g. Manski (1997) and Manski and Pepper (2000). In a transition framework one could formulate several different kinds of MTR assumptions. In this paper we make the same assumption as in RV. That is that the treatment effect either is positive or negative for all individuals in all time periods. It means that we do not assume a specific direction for the treatment effects, only that the direction is the same for all individuals. For that reason we do not find the assumption to be very restrictive. Also, note that it is considerably weaker than the assumption of constant treatment effect made in the MPH model.

The results exploring these assumptions are presented in tables 3 and 4, and are displayed in figure 1 and 2. These bounds are for the additive effect of the treatment regime on the transition probability during the treatment regime. Hence, the bounds are not transformed to a 'per week' effect on the transition probability.

It is immediately apparent that these weak assumptions tighten the bounds considerably. In fact for the most restrictive specification using all three assumptions we are able to bound the effect away from zero (in the positive direction) for many treatment regimes. Consider the effects of the experiment in Storstrøm county, reported in table 3. In the six first treatment regimes we find positive effects of the experiment on the transition rate to unemployment, and in the 3rd

to the 6th treatment regime, the lower bounds are considerably larger than zero. These regimes are the JSA programmes, frequent meetings, pre-programme and start programme.

Hence, in Storstrøm county, where the treatment protocol was followed most strictly, we conclude that the overall positive effects of the experiment is a combined effect of several sub-treatments. Hence, a range of different policies contribute to the decrease in the time spent unemployed. Namely, including job search courses (carrots) and frequent meetings, which may both have a monitoring (stick) component and a job search assistance (carrot) component. The positive effects shortly before and in the early phases of programme participation further suggests the presence of a threat effect of future and present programme participation. We should also point out that the bounds for the later programme regime and the post treatment regime are quite wide even for this more restrictive specification. In fact, we cannot rule out very large positive nor small negative effects after the programme has ended. This is because at this later stage the dynamic selection implied is potentially severe.

The results from the experiment in Southern Jutland are slightly less clear cut. For the most restrictive specification, we find that the lower bounds for JSA programmes and the frequent meetings are larger than zero, indicating positive effects on the transition rate out of unemployment. For the subsequent treatment regimes the lower bounds are all positive but very close to zero. The slightly more uncertain bounds for Southern Jutland county is most likely a consequence of the less strict adherence to the treatment protocol.

The overall conclusion that stands out from this analysis of the bounds of the treatment regimes is that JSA courses and frequent meetings with case workers certainly have positive effects on transition probabilities out of unemployment, while pre-programme and early programme treatment regimes (threat effects of programmes) have positive effects in one region and possibly has so in the other region as well. Later effects are more difficult to assess due to the extent of dynamic

selection at these stages.

6.4 Heterogeneous effects

We now test for important heterogeneous effects by gender and by age (less than 40 and 40 or above). We provide bounds under the combined assumption of positively correlated outcomes, monotone treatment responses and common shocks.

The results for the two counties are presented in table 5 and 6. There are no clear patterns that stand out for both regions. Still, there are interesting results, nevertheless. In Storstrøm county, there are large (coefficient larger than 0.05) positive effects of the new policy letter, of the frequent meetings and of the pre-programme regime for women, while only the JSA programme gave large positive effects for men. In Southern Jutland county, there are only large positive effects of the JSA programme for women, while for men, there are large positive effects of the JSA programme, the frequent meetings, and the pre-programme regime, i.e. almost exactly the opposite from Storstrøm. The differences in the treatments in the two regions is first that in Storstrøm, meetings were every week, while in Southern Jutland they only took place one every fortnight. This may seem innocuous, but a recent social experiment investigating specific treatments in separate experiments find precisely that women react to meetings while men react to programme threats, see Rosholm et al. (2010). The results presented here are not in disagreement with such an interpretation.

Splitting the samples by age, we find that in Storstrøm, there are large positive effects of JSA programmes, frequent meetings, and pre-programme regimes for those aged below 40, likely the more mobile part of the labour force. These treatments have no strong effects for older workers. For Southern Jutland county, the results are less clear cut in this dimension.

Summing up, there is a tendency that meetings are more likely to have positive effects for women, while a policy mix relying more heavily on the programme components tend to produce larger effects for men. Moreover, if anything, effects

appear to be slightly larger for the younger age group.

7 Conclusion and policy discussion

In this paper, we have used non-parametric techniques to make inference on treatment effects of specific policies applied sequentially to unemployed workers in a social experiment. The dynamic selection problem, which is normally dealt with by resorting to distributional assumptions and parametric or semi-parametric models, is dealt with by providing bounds on the treatment effects. Adding a few fairly weak assumptions, the bounds narrow considerably, allowing us to infer that some of the specific treatments had positive effects.

Specifically, the social experiment was one where active labour market policies were applied very early during the unemployment spell and at an intensive rate. We find that the overall policy package reduced unemployment duration by about two weeks, and that job search assistance, frequent meetings, and threat effects associated with perceived future programme participation have significant positive effects on the transition probability out of unemployment. The Danish Economic Council (2007) perform a cost-benefit analysis of the experiment and show that the overall package led to a net benefit of €2000 per fresh unemployment spell, mostly due to the increase in production associated with the shortening of unemployment spells.

Given that a meeting is the cheapest policy instrument (a meeting of an average duration of half an hour has marginal costs of approximately €23 per week for weekly meetings) and that meetings appear to have consistently positive effects, particularly for women, that begs the case for frequent meetings between case workers and clients for newly unemployed workers, especially women.

Job search courses also showed positive effects, but as they are more expensive (on average €200 in operating costs per participant per week, excluding the UI benefits they will get whether they participate or not), the case for job search

courses is considerable weaker. Moreover, the policy of frequent meetings has not been subjected to a general equilibrium analysis, and the assumption of no impact on other unemployed workers might be problematic. See, however, Gautier et al. (2010) for an equilibrium analysis on the same data suggesting that there are no general equilibrium effects of the experiment.

Longer lasting programme participation gives rise to a threat effects but is also quite costly, since the average cost is around €200 per participant per week, excluding UI benefits. Nevertheless, if sufficiently many unemployed workers are pushed into employment before programme start or in the early phases of the programme, such policies might be worthwhile.

References

- Andersen, T. and Svarer, M. (2007), ‘Flexicurity-labour market performance in denmark’, *CESifo Economic Studies* **53**, 389–429.
- Ashenfelter, O., Ashmore, D. and Deschênes, O. (2005), ‘Do Unemployment Insurance Recipients Actively Seek Work? Evidence from Randomized Trials in Four U.S. States’, *Journal of Econometrics* **125**, 53–75.
- Black, D., Galdo, J. and Smith, J. (2003), ‘Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach’, *American Economic Review* **93**, 1313–1327.
- Card, D., Kluve, J. and Weber, A. (2009), Active labor market policy evaluations: A Meta-Analysis. IZA Discussion Paper No. 4002.
- Chernozhukov, V., Hong, H. and Tamer, E. (2007), ‘Estimation and Confidence Regions for Parameters Sets in Econometric Models’, *Econometrica* **75**, 1243–1284.
- Danish Economic Council (2007), The Danish Economy, spring 2007 (in Danish), Secretariat of the Danish Economic Council. (n.d).*
- Dolton, P. and O'Neill, D. (2002), ‘The Long-Run Effects of Unemployment Monitoring and Work Search Programmes’, *Journal of Labor Economics* **20**, 381–403.
- Elbers, C. and Ridder, G. (1982), ‘True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model’, *The Review of Economic Studies* **49**, 403–409.
- Frederiksen, A. and Westergaard-Nielsen, N. (2007), ‘Where did they go?’, *Labour Economics* **14**, 811–828.

- Gautier, P., van der Klaauw, B., Rosholm, M. and Svarer, M. (2010), General Equilibrium Effects of Intensive Active Labour Market Policies; Evidence from a Social Experiment. Draft.
- Graversen, B. and Van Ours, J. (2008a), ‘Activating Unemployed Workers Works: Experimental Evidence from Denmark’, *Economic Letters* **100**, 308–310.
- Graversen, B. and Van Ours, J. (2008b), ‘How to Help Unemployed Find Jobs Quickly: Experimental Evidence from a Mandatory Activation Programme’, *Journal of Public Economics* **92**, 2020–2035.
- Hagglund, P. (2006), Are there pre-programme effects of swedish active labour market policies?: Evidence from three randomised experiments. IFAU Working Paper 2006:2.
- Hahn, J. and Ridder, G. (2010), Partial identification and confidence intervals. Working Paper.
- Imbens, G. and Manski, C. (2004), ‘Confidence Intervals for Partially Identified Parameters’, *Econometrica* **72**, 1845–1857.
- Kluve, J. (2006), The Effectiveness of European Active Labor Market Policy. IZA Discussion Paper No. 2018.
- Manski, C. (1997), ‘Monotone Treatment Response’, *Econometrica* **65**, 1311–1334.
- Manski, C. and Pepper, J. (2000), ‘Monotone Instrumental Variables: With an Application to the Returns to Schooling’, *Econometrica* **68**, 115–136.
- Meyer, B. (1995), ‘Lessons from the U.S Unemployment Insurance Experiments’, *Journal of Economic Literature* **33**, 91–131.
- Ridder, G. and Vikström, J. (2011), Bounds on Treatment Effects on Transitions. IFAU Working paper 2011:6.

- Rosholm, M. (2008), Experimental Evidence on the Nature of the Danish Employment Miracle. Working Paper 08-14 University of Aarhus, Aarhus School of Business, Department of Economics.
- Rosholm, M., Svarer, M. and Staghoj, J. (2010), Four Social Experiments on Intensive Active Labour Market Policies. Draft.
- Van den Berg, G. (2001), Duration Models: Specification, Identification and Multiple Durations, *in* H. J. and Leamer E., ed., 'Handbook of Econometrics', Vol. 5, Amsterdam, North-Holland, pp. 3381–3460.
- Van den Berg, G. and van der Klaauw, B. (2006), 'Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment', *International Economic Review* **47**, 895–936.

Tables and Figures

Table 1: Treatment regimes

Treatment regime:	Weeks	
1	1-2	Pre-treatment
2	3-4	New policy letter
3	5-9	JSA programme
4	10-15	Frequent meetings
5	16-18	Pre-programme
6	19-22	Start programme
7	23-30	Later programme
8	31-39	Post-treatment

Table 2: Descriptive statistics

	Storstrøm county		Southern Jutland county	
	Treatment	Control	Treatment	Control
<i>Observations</i>	1174	1219	1055	1065
<i>Unemployment duration</i>				
Unemployment duration	14.80	16.72	15.63	17.52
Truncated duration	14.60	16.52	15.44	17.26
<i>Raw transition rate</i>				
Period 1	0.087	0.095	0.085	0.092
Period 2	0.105	0.128	0.106	0.095
Period 3	0.233	0.289	0.217	0.290
Period 4	0.283	0.326	0.257	0.318
Period 5	0.174	0.242	0.175	0.204
Period 6	0.189	0.243	0.188	0.196
Period 7	0.278	0.242	0.259	0.275
Period 8	0.184	0.121	0.117	0.167
<i>Censoring</i>				
within 20 weeks	0.000	0.000	0.000	0.000
within 30 weeks	0.008	0.023	0.011	0.017
within 40 weeks	0.097	0.129	0.102	0.140
Total	0.144	0.180	0.152	0.217
<i>Gender</i>				
Male	0.62	0.60	0.54	0.54
Female	0.38	0.40	0.46	0.46
<i>Age group</i>				
Less than 24	0.09	0.09	0.11	0.13
25-29	0.13	0.12	0.14	0.14
30-39	0.24	0.27	0.26	0.24
40-49	0.26	0.24	0.26	0.25
50-59	0.25	0.25	0.20	0.22
60 and above	0.03	0.03	0.03	0.04

Table 3: Bounds on treatment regime effects for Storstrøm.

Treatment regime:	Weeks	NO [1]		MTR+CS [2]	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.008	0.008	0.008	0.008
2. New policy letter	3-4	-0.082	0.140	0.022	0.030
3. JSA programme	5-9	-0.240	0.376	0.047	0.082
4. Frequent meetings	10-15	-0.942	0.974	0.011	0.126
5. Pre-programme	16-18			0.035	0.222
6. Start programme	19-22			-0.002	0.243
7. Later programme	23-30			-0.143	0.241
8. Post-treatment	31-39			-0.121	0.121

Treatment regime:	Weeks	PCO [3]		MTR+CS+PCO [4]	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.008	0.008	0.008	0.008
2. New policy letter	3-4	-0.070	0.126	0.023	0.030
3. JSA programme	5-9	-0.128	0.287	0.055	0.082
4. Frequent meetings	10-15	-0.283	0.326	0.043	0.126
5. Pre-programme	16-18	-0.174	0.242	0.068	0.222
6. Start programme	19-22	-0.189	0.243	0.053	0.243
7. Later programme	23-30	-0.278	0.242	-0.036	0.241
8. Post-treatment	31-39	-0.184	0.121	-0.063	0.121

Notes: Number of observations are 2393. NO indicates the bounds under random assignment. MTR stands for assumption monotone treatment response, CS common shocks, and PCO positively correlated outcomes.

Table 4: Bounds on treatment regime effects for Southern Jutland.

Treatment regime:	Weeks	NO [1]		MTR+CS [2]	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.007	0.007	0.007	0.007
2. New policy letter	3-4	-0.115	0.099	-0.012	-0.003
3. JSA programme	5-9	-0.189	0.372	0.070	0.074
4. Frequent meetings	10-15	-0.736	0.828	0.035	0.134
5. Pre-programme	16-18			-0.005	0.191
6. Start programme	19-22			-0.037	0.196
7. Later programme	23-30			-0.049	0.204
8. Post-treatment	31-39			0.017	0.167

Treatment regime:	Weeks	PCO [3]		MTR+CS+PCO [4]	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.007	0.007	0.007	0.007
2. New policy letter	3-4	-0.104	0.089	-0.011	-0.003
3. JSA programme	5-9	-0.103	0.264	0.070	0.073
4. Frequent meetings	10-15	-0.257	0.318	0.061	0.134
5. Pre-programme	16-18	-0.175	0.204	0.029	0.191
6. Start programme	19-22	-0.188	0.196	0.008	0.196
7. Later programme	23-30	-0.259	0.275	0.017	0.204
8. Post-treatment	31-39	-0.117	0.167	0.050	0.167

Notes: Number of observations are 2120. NO indicates the bounds under random assignment. MTR stands for assumption monotone treatment response, CS common shocks, and PCO positively correlated outcomes.

Table 5: Heterogenous treatment regime effects for Storstrøm.

Treatment regime:	Weeks	Men [1] (N=934)		Women [2] (N=1459)	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.006	0.006	0.007	0.007
2. New policy letter	3-4	-0.011	-0.005	0.042	0.049
3. JSA programme	5-9	0.068	0.072	0.041	0.084
4. Frequent meetings	10-15	0.014	0.082	0.065	0.147
5. Pre-programme	16-18	0.035	0.128	0.105	0.299
6. Start programme	19-22	0.040	0.170	0.078	0.312
7. Later programme	23-30	-0.023	0.133	-0.052	0.255
8. Post-treatment	31-39	-0.067	0.076	-0.053	0.147

Treatment regime:	Weeks	Young < 40 [3] (N=1200)		OLD > 40 [4] (N=1193)	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.006	0.006	0.010	0.010
2. New policy letter	3-4	-0.001	0.005	0.047	0.057
3. JSA programme	5-9	0.070	0.074	0.042	0.095
4. Frequent meetings	10-15	0.065	0.145	0.025	0.115
5. Pre-programme	16-18	0.091	0.241	0.047	0.180
6. Start programme	19-22	0.030	0.250	0.074	0.232
7. Later programme	23-30	-0.083	0.212	0.001	0.255
8. Post-treatment	31-39	-0.153	0.054	0.003	0.144

Notes: N is the number of observations. Bounds under monotone treatment response, common shocks and positively correlated outcomes.

Table 6: Heterogenous treatment regime effects for Southern Jutland.

Treatment regime:	Weeks	Men [1] (N=973)		Women [2] (N=1147)	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.000	0.000	0.014	0.014
2. New policy letter	3-4	0.013	0.013	-0.031	-0.018
3. JSA programme	5-9	0.057	0.069	0.074	0.087
4. Frequent meetings	10-15	0.093	0.163	0.025	0.102
5. Pre-programme	16-18	0.062	0.183	-0.018	0.100
6. Start programme	19-22	0.006	0.152	0.002	0.099
7. Later programme	23-30	0.026	0.248	-0.008	0.081
8. Post-treatment	31-39	0.054	0.143	0.028	0.126

Treatment regime:	Weeks	Young < 40 [3] (N=1132)		OLD > 40 [4] (N=988)	
		LB	UB	LB	UB
1. Pre-treatment	1-2	0.014	0.014	-0.001	-0.001
2. New policy letter	3-4	-0.012	0.002	-0.011	-0.010
3. JSA programme	5-9	0.082	0.084	0.053	0.062
4. Frequent meetings	10-15	0.050	0.139	0.071	0.125
5. Pre-programme	16-18	0.040	0.205	0.018	0.166
6. Start programme	19-22	0.043	0.217	-0.027	0.139
7. Later programme	23-30	-0.042	0.199	0.074	0.202
8. Post-treatment	31-39	0.016	0.157	0.080	0.176

Notes: N is the number of observations. Bounds under monotone treatment response, common shocks and positively correlated outcomes.

Figure 1: Bounds on conditional transition probabilities for Storstrøm. Different assumptions.

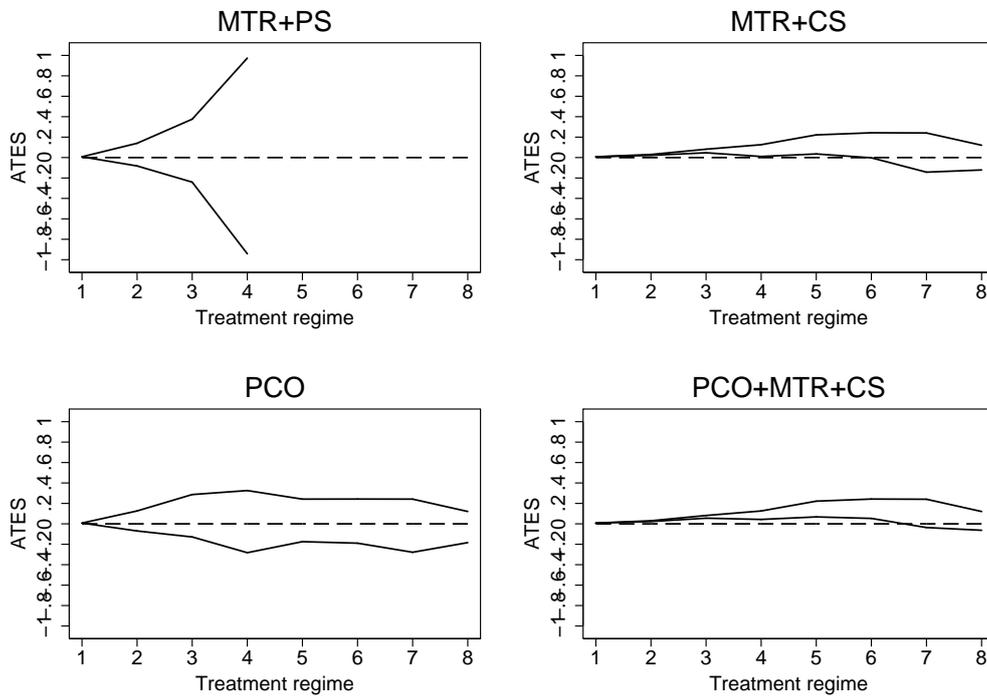


Figure 2: Bounds on conditional transition probabilities for Southern Jutland. Different assumptions.

