Introduction
This thesis consists of three separate essays in labour market policy evaluation using both natural and classical experimental evaluation designs. In this introduction, I first present the research questions of each essay before describing the implementation and the results in more detail. The last section reports practical experiences from pursuing four randomised experiments presented in essay 2 and 3. Since the previous, and only, Swedish experiment in this field was conducted in 1975, I extract some lessons of perhaps some use in future randomised demonstrations.

The first essay explores an often neglected parameter of the unemployment insurance (UI) system that regulates the required attachment with the labour market in order to qualify for UI: the entrance requirement (ER). The ER specifies the minimum amount of work necessary to be entitled for UI benefits. Most research on UI has focused on the impact of changes in the replacement ratio or the length of the benefit period on unemployment duration. However, significantly less is known about the UI regulation’s influence on the employment relationship. To what extent does it affect people’s timing of job separations?

Both firms and employees have incentives to account for the UI system in their decisions. The prototypical case is the seasonal industry where workers combine spells of work and unemployment in circular flows. Since the firms know that the UI system will attenuate the workers’ separation costs, they employ workers for short periods to meet short-term needs. In this case, benefit receipt primarily acts to redistribute income and leisure for actors “playing the system” – and not as an insurance.

In Sweden during the 1990s, the ER was tightened on two occasions, requiring additional attachment to the labour market in order to qualify for benefits. To what extent do the agents of the labour market adjust to the new requirements and how does the change in the ER affect average employment duration? These issues are dealt with in the first essay of this dissertation.
The second paper considers the new opportunities offered by the Internet for the public employment services (PES). Since 1995, several on-line placement services have been introduced in Sweden. Today, the PES Vacancy bank, where employers advertise their job vacancies, has 600,000 unique visitors every month. At the employment offices, several self-service instruments have been introduced. For instance, in a computerised training programme, the unemployed can learn the basics of effective job search. Also, for a few years now, the newly unemployed who register at the employment office is responsible for formulating the individual “action plan”.\footnote{The individual “action plan” consists of, among other things, the specified job-search requirements, the current geographical search area (which depends on the length of the unemployment period), and suitable active measures.} This is all part of a new labour market policy strategy where resources are reallocated from newly unemployed to long-term unemployed and other disadvantaged groups.

The new policy has been subject to much criticism, not only among the unemployed but also among employment officers and union representatives. They question the proposition that the self-service instruments can be as efficient as personal meetings with the employment officers. With less personal contact early in the unemployment spell, identifying the unemployed in most need of assistance becomes more difficult. Extensive periods without intervention then risks worsening job chances even further. Another objection is that people are not as skilled using computers as many may think. The self-registration has caused trouble where much effort must be spent on correcting and complementing data in the information system. This has to some extent annulled the intended time-saving from implementing the self-service instruments.

Current developments involve a higher degree of interactivity between job seekers and employment officers, which means that further dimensions of traditional employment services are being added to the Internet services. In 2002, the Swedish Labour Market Board (SLMB) carried out a nation-wide demonstration programme for pursuing job-search activities on the Internet. The
idea was to investigate whether or not these types of services were feasible to perform on the Internet, and whether or not the services should be a permanent feature of the employment services. The demonstration was preceded by an application procedure where voluntary job seekers could sign up for participation. What is the clients’ interest in this type of services, and what are the job-chance enhancement prospects for the unemployed?

In the third and final essay, I investigate the occurrence of anticipatory effects of active placement efforts at the employment offices in Sweden. Being one of the countries with the highest expenditures on active labour market policies, Swedish research on individual programme effects has explicitly focused on the impact during and after the programme services. The results of these studies have often shown only small and even adverse effects of programme participation (see Calmfors et al. 2001).

Besides upgrading skills and activating long-term unemployed to improve employability, policymakers often motivate the usage of active measures as a tool for testing work motivation. A common perception among employment officers is that referrals to different types of compulsory programme activities help to remove those having little problem finding employment, thus reducing the extent of moral hazard behaviour. This is because participation is expected to be time consuming and thus reduce time for both leisure and job search. If the activities are anticipated, this would lower the value of being unemployed before start and the exit rate from unemployment is expected to increase. If, on the other hand, the expected returns of participation – in terms of improved job chances and/or the distribution of wage offers during and after the services – overshadow the negative aspects, the value of unemployment increases and the escape rate drops before start.

Empirically, the former version is supported. A few non-experimental Swedish studies find evidence of increased exit rates from unemployment in the weeks prior to the start of labour market programmes (see for instance Carling et al., 1996). None of these studies, however, explicitly set out to study the “motivational” aspects of the active labour market policies. Outside Sweden,
experimental studies from the U.S (see for instance Black et al. 2003) and the U.K (Dolton & O’Neill, 1996), find significantly increased off-unemployment hazard rates prior to attending different job-search and re-employment services. The present study explicitly investigates the behavioural adjustment of being referred to active placement efforts. Using experimental data from three different regions of Sweden, I try to answer the questions; how common are ”pre-programme” effects, under what circumstances do they occur, and can they in turn motivate the large spending on active measures?

The three essays

Essay I. Effects of Changes in the Unemployment Insurance Eligibility Requirements on Job Duration – Swedish Evidence

This paper contributes to the very sparse empirical literature on the relationship between the entrance requirement (henceforth: ER) and the timing of job separations. Exploiting changes in the ER taking place in 1994 and 1997, I study behavioural responses in the timing of job exits in 1992, 1996 and 1998. I also study the effect of the last ER change on average employment duration. To my knowledge only four Canadian studies, between 1994 and 1998, have investigated the effect of the ER on employment duration. All find evidence of significant increases of the employment-unemployment hazard when the ER was satisfied.

My analyses are restricted to employment spells of persons who prior to employment were unemployed. To establish the length of employment spells, I use unemployment register data and information on job duration until returning to unemployment. A flexible piece-wise constant hazard model captures weekly job exits corresponding to the time of fulfilling the ER.

Studying each year separately, I find in contrast to the Canadian studies no clear evidence of adjustments to the ER in terms of distinct mass termination of job spells at, or around, the week of fulfilment. I propose several possible explanations for this finding, for instance the lack of data on exact employment duration and the possibility of fulfilling the ER through several, instead of just one, composite
periods of employment. Instead performing across-year analyses with years with different ER, I find evidence of adjustments in job turnover to the entrance requirement in all three years. Analysing the effect in one industry (farmers) and one region (Norrbotten), both characterised by relatively high seasonality in the production process, suggests that changes in the ER primarily affects sectors with relatively large recurrent flows between jobs and unemployment (indicating a high degree of awareness of the UI system). Using the estimated UI parameters, I estimate an approximate 3-week prolonging of the average employment duration due to the 5-week extension of the ER between 1996 and 1998. On Canadian data, Green & Riddell (1997) estimated a 1.5-week increase in the average duration of employment in regions with high unemployment as a result of a temporary extension of the ER from 10 to 14 weeks.

Essay II. Job-search Assistance Using the Internet – Experiences from a Randomised Experiment in Sweden

The contribution of this paper is threefold. First, I investigate the employment outcome from a demonstration programme offering voluntary job-search assistance on the Internet. Second, by randomly assigning job seekers to different services, I explore to what extent the experimental evaluation design succeeds in circumventing common difficulties in experimental evaluation. Third, taking advantage of the experimental design and the non-biased experimental impact estimate, I assess to what extent standard non-experimental evaluation methods succeed in replicating consistent impact estimates.

In North America, practical and analytical experiences from conducting social experiments have been accumulated for almost 40 years. In Europe, however, with a few exceptions in Sweden, Britain, Norway (2), Denmark and the Netherlands, experiments have not been an alternative evaluating social programmes. Therefore, documenting the experiences and learning the most from every opportunity is important.
I find that the experimental design successfully avoids many typical problems inherent in social experiments, for instance randomisation bias, bureaucratic behaviour and Hawthorne effects. However, the sample size is small (636). Furthermore, a large proportion of the experiment group members either never entered the programme (47%), or dropped out early. Both “no-shows” and dropouts are common in experimental evaluation. However, they pose a problem only in situations where the effect of actual participation is the parameter of most interest. In this case, with no activity-level restriction, I estimate the effect of the “intent-to-treat”, which explicitly considers the choice of not participating at all.

With the low activity level, the intent-to-treat impact estimate fails to reject the hypothesis of a zero programme impact. Furthermore, using the random assignment as an instrument in an IV model, studying the effect of various levels of actual treatment, no significant effects are found.

Finally, the comparison of methods shows that standard non-experimental evaluation techniques successfully reproduce the experimental intent-to-treat impact estimate. The results may, however, in part be due to the non-adjusted outcome similarity between the control and the (constructed) comparison group members. Also, with an experimental impact estimate of low precision only large deviations would have generated another conclusion. Further analyses of the ability of these methods to account for selection into other types of programmes are necessary in future research.

**Essay III. Are there Pre-Programme Effects of Swedish Active Labour Market Policies - Evidence from Three Randomised Experiments**

The third and final essay investigates the incentive effects of being referred to compulsory active placement efforts at the employment office. Using unique experimental data from three separate demonstration programmes in 2004, the exit rate from unemployment between referral to and start of the programme services is compared between UI eligible experiment and control group members.
The results from the three studies are mixed. In two of the demonstrations, in Uppsala and Östergötland, no effect of the referrals was found. In the third experiment, in Jämtland, I find evidence of a 38 per cent increase in the off-unemployment hazard rate preceding services involving a combination of compulsory job-search assistance activities and increased monitoring of job-search efforts. By offering two different treatment packages with random assignment to each treatment, I find that the positive effect derives from the job-search assistance activities. The effect of referrals to recurrent interviews for monitoring the job search is significantly lower and non-significantly different from the exits of the control group. This finding is possibly the result of the job-search assistance activities being arranged in groups, which for some unemployed may be stigmatising, as opposed to the in-person interviews. The estimated effect of being referred to the job-search assistance activities corresponds to a two-week reduction of the ongoing UI spells. The positive effect is not the result of more temporary interruptions of the unemployment spells among the experiment group members due to, for instance, less attractive job matches.

I propose two possible explanations for the diverse results. First, whereas the Jämtland demonstration invited a broad group of unemployed to participate, the other two targeted on specific groups of unemployed. These groups could on average have relatively less scope for finding a way out of unemployment. Second, Lindbeck et al. (2004) argue that the large local variations in sick leave in Sweden are related in major part to a “sick leave culture” based on local-specific attitudes towards sick leave. An interesting fact is that the county of Jämtland reported the highest sick leave in the country in 2003. If there exists a sick leave culture in Jämtland, it would be easy to imagine a similar tradition within the UI system, which to some extent might explain the results of the Jämtland demonstration.

A final result of this paper suggests that using the unemployment register as the sole informant when analysing programme effects on unemployment duration could cause substantially biased results. Comparing outcomes when using the
unemployment register and those when instead using the more reliable UI payment register data, the programme effect of the activities in two of the experiments are severely upward biased.

**Some lessons from conducting the experiments**
Since the experiments presented in this thesis are the first in this policy field in Sweden in 30 years, I take the opportunity to report some experiences from conducting them.

*Planning the experiment*
In the experiments, especially those reported in the last essay, the planning period was somewhat short. Additional time spent with the case workers in the scheme groups would have helped me, as an evaluator, to learn more about the administrative work at the employment offices, for instance regarding the enrolment routines. Extra preparation time would also have allowed the scheme group members to more fully grasp the intention behind the evaluation requirements. The significance of the fact that this “learning process” (on both sides) to some extent reached into the active period of the experiments should, however, not be overrated. Working routines as well as service elements are inevitably fine-tuned over time as skills increases among the scheme workers.

With only limited preparation time before evaluation, however, a recommendation is to first test the outlined strategies in a brief trial period before start. This way, obvious flaws can immediately be corrected and the services offered at the start of the evaluation period and the ones offered in the end can be made as similar as possible.

*Organisation*
The organisation behind the demonstrations reported in the last essay consisted of three levels. At the top, supervising the programmes, was a *project group* at the
SLMB. This project group consisted of four persons, one assigned to each scheme and as such appointed chairman of the reference group attached to each scheme.

The reference group, which represented the middle level of the scheme organisation, also contained the assigned SLMB evaluator, representatives from the county labour boards and the operative team manager of the particular project (the project team manager). In some cases, the local industry and/or the local trade unions were also represented. The role of the reference group was to continuously follow the demonstration activities and to make strategic decisions on comprehensive matters concerning the course of the project.2

Finally, at the operative level, a project team of 3-5 case workers including the project team manager carried out the services at the local employment offices. Members of the project team were selected by the county labour board.

The reference groups were thus responsible for operating the demonstrations and making decisions along the way. However, the reference group met only every 4-6 weeks and transitional decisions were sometimes necessary to settle immediate and unforeseen issues. For instance, when the extent of the register deficiencies (described in more detail in the third paper) became apparent shortly into the evaluation period, the evaluator’s recommendation to “clean” the register before randomisation met some resistance.3 There was confusion about “who is making the decisions here”. Clearly, a more well-defined, as well as a more flexible, decision-making structure was necessary to deal with these situations, particularly in view of the lack of experience in this field of evaluation among those involved.

---

2 These matters could for instance concern the amount of time and resources allocated to contacts with the unemployed and the employers respectively, or the rate at which new job seekers should be added to the experiment (and control) group.

3 The register deficiencies primarily consist of late registration of events. For instance, job seekers who find jobs or leave the work force sometimes omit to inform the employment office. Using register data to identify the target populations thus involves the risk of wrongly identifying employed persons as unemployed.
Bureaucratic resistance

Bureaucratic resistance, i.e., a lack of cooperativeness among those administrating the experiments, is probably the greatest threat to a successful experiment. Experience shows that routines implemented without the necessary approval risk being poorly executed (see Björklund & Regnér, 1996). Due to the risk of resistance, or sometimes for reasons of cost, designing the best possible experiment could involve specifying arrangements not optimal from a theoretical point of view. Finding compromises, balancing the expected returns in terms of a higher quality evaluation against the risk of lower acceptance among those carrying it out is an important task for an evaluator. The following are two examples of concessions made in the experiments conducted.

First, as regards the register deficiencies, an outreach procedure where target group members were contacted before randomisation would have been the most efficient way to reduce the occurrence of no-shows. This way, current status could have been checked and the register information corrected before start. However, such a routine would have involved extensive administrative procedures before each new enrolment and would have been costly, not only through a lower acceptance for the experiment, but also in terms of a lower flexibility in the enrolment procedure with fewer observations in consequence.

Second, the control group members being directed towards the employment office’s regular services make interpretation of the experimental estimator more difficult since the effect from treatment relative to non-treatment is no longer identifiable. When controls receive good substitutes for a programme, the outcome difference could be zero although the effect from treatment relative to no treatment is positive. The problem of substitution is endemic in both experimental and non-experimental evaluation. However, obstructing the controls from receiving alternative training, in this case their normal services, would undoubtedly have raised ethical objections, not only among the employment officers.
Compliance

In contrast to the substitution problem, compliance is specifically related to experimental evaluation. The problem of compliance occurs when people assigned to participate either do not show up (“no-shows”) or drop out of the programme evaluated. This dilutes the experimental estimator because the difference in treatment between the experiment and control group is reduced.

Certainly, unlike researchers in fields such as medicine and biology, a researcher conducting a social experiment cannot fully control the level of treatment among either experiment or control group members. If the choice of participation is related to the entitlement of compensation from the social insurance system, for instance UI, at least there are strong incentives to comply with the programme assigned. Still, compliance problems can arise for several reasons. For example, if register data are used to identify the population to which the programme is targeted, observations can be lost due to register deficiencies. Second, if there is a waiting period between randomisation and programme start (or rather between informing the experiment group members of the randomisation outcome and programme start), some observations are almost inevitably lost due to people moving or changing status (for instance from unemployment to employment). In the more unusual case where the programme is voluntary, i.e. when rejecting treatment does not involve UI sanctions, non-compliance should be regarded as an outcome of the programme. Here, the policy parameter of most interest is the effect of the “intent-to-treat” rather than the effect of participation. From an evaluation perspective, the lack of compliance in this case constitutes no “compliance problem”.

In the second essay – evaluating the effects of offering job seekers Internet services – the experiment and control group members were not informed of the randomisation outcome until the day of programme start. With no waiting period, those leaving job-seeker status between randomisation and programme start could be eliminated because this outflow can be assumed to be unrelated to the randomisation outcome. However, with the programme being voluntary, a large
amount of the experiment group members chose either to not show up (47%) or to drop out after receiving only a small treatment dose. Although the policy-relevant intent-to-treat parameter was still retrieved, this estimator was not very informative to the group of prospective participants as to whether or not to apply for participation. For this group, the effect of actual treatment, or the “treatment-on-the-treated” parameter, had been of more interest.

The last paper explicitly deals with the compliance problem generated by the waiting period. More specifically, the incentive effects of being referred to active placement efforts were investigated. In order to isolate the effect of this particular type of compliance problem, I must account for other types of compliance problems in the analysis. Since the programmes of interest were mandatory, this was not an issue here. However, using flawed data when attempting to identify target populations caused some problems. In particular, employed persons were wrongly identified as unemployed. However, by using alternative data (UI-payment data) in performing the analyses, the impact of register deficiencies on the results was neutralised.

Concluding remarks
Reporting some experiences from conducting the experiments presented in two of the essays, I have especially pointed to some flaws in the implementations such as too short preparation period, ambiguous decision-making policies, and various kinds of compliance issues. In this respect, the lesson from these experiments is similar to the experiences from earlier experiments, showing that the design and the interpretation of experiments, in practice, are more difficult than it may first seem.

However, the benefits from experiments as opposed to non-experimental methods are undeniable. First of all, correctly performed, a simple comparison of mean outcomes generates a consistent estimate of the programme impact on those participating. Second, the experimental design with a randomly assigned experiment and control group is easy for policymakers to understand. Finally,
experiments provide valuable opportunities to assess the performance of various standard non-experimental methods.

With this in mind, the perhaps most useful finding is that despite the flaws, parameters providing a basis for policy decisions have been retrieved. In that respect, these experiments serve as good examples for future experiments.

References


