Chapter 1

Assessing Welfare Effects of ALMPs: Combining a Structural Model and Experimental Data

Jonas Maibom
Aarhus University, CAFE

Abstract[1]

Utility costs associated with participation in Active Labour Market Programs (ALMPs) have been suggested in the literature as a way to interpret the well established existence of threat or ex ante effects associated with programme participation. This paper combines data from a randomized experiment with a structural economic model to estimate the utility costs and potential productive effects from programme participation. The model generates a link between observed behaviour such as job finding rates into structural parameters such as utility costs, while the experiment generates exogenous variation in programme participation and ensures that results are not driven by unobserved confounders. The estimates of the model are used to calculate the compensating variation, i.e. the monetary compensation which leaves individuals indifferent between belonging to the treatment or the control group. This enables an analysis of whether the programmes represent a worthwhile social investment by comparing the gains to costs including those borne by the participating individuals. Thereby some empirical quantification of a long lasting discussion in the literature that analyses the optimal design of labour market policies is provided. The estimates of the structural model are exploited to analyse the heterogeneity in the compensating variation in relation to future prospects and the timing of treatment in an envir-

[1]Acknowledgement: This paper benefited from numerous comments and discussions at seminars and conferences. I gratefully acknowledge comments and suggestions from C. Ferrall, J. Lise, T.M. Andersen, J. Bagger, M. Svarer and R. Vejlin. I thank the Danish Labour Market Board for making the data available and the CAFE grant for enabling part of this research. Results were generated using Ox (www.doornik.com) and FiveO (a part of niqlow: jdi.econ.queensu.ca/niqlow).

onment which is characterized by duration dependence in unemployment and rich heterogeneity across individuals. The results suggest that traditional Cost-Benefit calculations which do not take the individual costs into account largely overstate the gain from having these programmes. The costs are substantial and are important to quantify in order to assess whether the current mix between programmes and UI is optimal.

1.1 Introduction

In this paper I estimate how individuals value participation in Active Labour Market Programs (ALMPs) which serve as a conditionality for receiving unemployment insurance (UI). Any potential costs associated with programme participation are a crucial input in an analysis of whether such conditionalities in transfer recipiency constitute a worthwhile social investment or whether potential suboptimal individual behaviour is better controlled by e.g. reducing benefits. Two kinds of ALMPs are analysed: meetings and short activation programmes at the job centre. By combining data from a randomized experiment and a structural economic model, I estimate the utility costs and calculate the resulting compensating variation (CV), i.e. the monetary compensation which equalize expected utility across the treatment and control group at inflow into the experiment. The CV takes potential productive effects from programme participation (increases in job offer rates), the value of alternative choices and future prospects for participants into account. The estimates, and detailed data on other costs and gains from the programmes,[2] allows for an assessment of whether the programmes under investigation constitute a worthwhile social investments. Thereby some empirical evidence on how ALMPs affect individual behaviour and how this should affect our use of these programmes is provided.

While there exists a large literature evaluating the effectiveness, in terms of e.g. job finding, of various kinds of ALMPs (for reviews see Card et al. (2010), Kluve (2010)) there are very few papers focusing on the mechanism behind any generated impact and in particular the importance of utility costs. The empirical literature (see Black et al. (2003a) and Hagglund (2011)) has documented the presence of so-called threat or ex ante effects which suggest that individuals view programme participation as costly. As these programmes 'tax' leisure time by replacing it with time in the job centre (public employment service); unpleasant or uninspiring work, increased effort, monitoring or stigma are all potential explanations for the existence of such costs.[3]

The existence of costs implies that an evaluation of programme impacts through an analysis of its impact on e.g. employment, is only partial in nature. Impacts arise at a cost, and in order to assess whether a programme is actually beneficial one needs to contrast the benefits generated by the programme with its costs - here both actual programme costs (staff at the job centre)

[2]The costs considered are i) costs associated with running the programmes, ii) the compensating variation associated with the existence of the experiment, iii) costs associated with an increase in production (lost leisure). Gains considered are i) value of increased production, ii) saved income transfers.

[3]Both pecuniary and non-pecuniary costs are therefore explanations for why the non-market (job centre) wage could differ from the market wage and they are essentially explanations of compensating wage differentials (effort in the job centre is unpleasant and thus the "payment" is higher). In this paper I do not try to distinguish these different explanations, but instead estimate the total impact on agents.

and individual costs. Since information about individual costs are generally not available this ultimately introduces an imbalance between what programme evaluators evaluate as beneficial, and what society (or a social planner) would.[4] The imbalance stems from the fact that the unemployed respond to costs which are not included in the evaluation of the programme. This favours programmes which generate e.g. the largest reductions in unemployment duration regardless of how participants value participation in the programmes. The importance of this imbalance depends on the magnitude of utility costs, and therefore further knowledge is of central interest. The question is also interesting from a theoretical point of view as the magnitude of costs is important for whether conditionalities in transfer recipiency *can* actually constitute a worthwhile social investment (see next section).

Any quantification of costs borne by the individual requires some link to a behavioural model, the surrounding environment and an accurate description of the incentives faced by potential participants over time. This link generates a translation of observed behaviour into decision theoretic parameters such as utility costs. In the case of ALMPs this quantification is further challenged by the fact that participation is (ultimately) a conditionality for receiving UI benefits. Non-participation is therefore associated with a substantial loss of income due to sanctions or suspension from benefits for a period of time. Therefore a direct expression of preferences for the programme through choices of potential participants, choosing whether to participate or not, are not present - unemployed workers will choose to participate although participation is associated with costs that outweigh direct benefits from participation.[5] Costs must therefore be determined indirectly through behaviour such as job finding rates or wages in future employment. This requires a full economic model of behaviour and an accurate description of behaviour in the absence of the programme in order to identify the change in behaviour induced by the programme and thus the individual costs.

In order to quantify costs this paper develops a dynamic discrete choice model of job search and estimates it exploiting data from a Danish randomized experiment. The structural framework provides a mapping from observed behaviour into the determinants of decision making at the individual level. The experiment improves identification of unobserved costs for two reasons. First the experiment generates exogenous variation in programme participation, which ensures that differences in behaviour between control and treatment group can be prescribed to the impact of the programme. Secondly as the experiment is a finitely lived and time-varying intervention which generates useful variation in the incentives faced by individuals and improve identification of the central parameters.[6] To exploit the experimental variation the model contains a thorough

[4] In the words of Heckman et al. (1999): *".. By doing this, however, these evaluations value labour supply in the market sector at the market wage, but value labour supply in the non-market sector at a zero wage. By contrast, individuals value labour supply in the non-market sector at their reservation wage."*

[5] A literature starting with Moffitt (1983) identifies the stigma/utility cost associated with receiving welfare comparing take-ups and non-take-ups (extensive margin). This paper use variation in the intensive margin (the intensity of the conditionality) and compare the behaviour of individuals in intensive regimes with similar individuals in less intensive regimes. Variation in the intensive margin is generated by a social experiment and is thus exogenous, which is useful in the identification of the utility cost.

[6] From a methodological point of view this paper is therefore a part of a growing literature combining economic models and empirical strategies with high internal validity (here experiments). Two different approaches can roughly be distinguished by whether the experimental variation is used as a source of validation (a test of the

description of how the treatment changes over time which is particularly important to take into account when estimating the costs of programme participation in a dynamic setting. It allows agents to take into account that incentives for the treatment group change as they progress through the experiment - every week is one week closer to the expiration of the intensified treatment and thus the future cost associated with programme participation declines.

The model is used to calculate the compensating variation (CV) associated with the experiment. The CV takes into account that individuals can influence their likelihood of remaining unemployed, and thus their chances of participation in the programmes. The CV is therefore different from utility costs that reflect the immediate cost associated with inevitable programme participation. For instance the CV will be lower for individuals "capable" of leaving unemployment fast compared to individuals with worse employment prospects. Similarly the CV associated with interventions at inflow into unemployment is higher than in the case where programmes start later in the unemployment spell because the former makes future participation more likely. A final important aspect which influences the size of the CV is the presence of risk aversion in the model. This increases the monetary compensation due to decreasing marginal utility of wealth (decreasing "efficiency" of the initial monetary compensation) and the inter-temporal separation between the paid compensation and future programme participation.[7] Naturally a quantification of these aspects - and an assessment of their relative importance - are important inputs in the discussion and future design of optimal labour market policies. The aim of the model is to generate an environment with several sources of heterogeneity between unemployed agents and in the cost associated with programme participation. The heterogeneity in the environment implies that the impact of ALMPs differs across individuals depending on their current state and future prospects. This way heterogeneous treatment effects are endogenous to the model and the resulting CV - which serves as a crucial input in a subsequent welfare calculation - will also vary across agents.

In the model agents face two discrete choices: while unemployed they choose a level of search intensity and if a job offer is present they choose whether to accept the job offer or not. The social environment is stationary and ergodic. Employed individuals stochastically accumulate skills each period while employed. Their job separation rate depends on their level of skills. If they loose their job their stock of skills may depreciate. Unemployed workers face job offer rates which depend on their search activity and their duration in unemployment. Wages depend on a draw of firm productivity and the level of skills. While unemployed individuals receive UI and in return have to participate in meetings/activation programmes - participation in these programmes is potentially costly but may increase job offer rates.

From a methodological point of view the model follows in the lines of a novel framework de-

behavioural model) or identification of parameters of the model. See Wolpin and Todd (2006), Attanasio et al. (2012), Ferrall (2012) and Lamadon et al. (2004) for examples of different approaches.

[7] There are no asset markets or savings in the model, the existence of these would allow the agents to smooth consumption across states and thus potentially decrease the impact of this later channel. In an environment without risk aversion the accumulated utility costs would represent an upper bound of the welfare costs associated with the experiment, but since risk aversion is an important justification for the existence of UI this is incorporated into the model.

veloped in Ferrall (2002, 2012).[8] This framework extends the classical work by e.g. Rust (1987) into a setting with unobserved non-IID time-varying state-variables, unanticipated (or zero probability) choices, corrections for endogenous sampling (initial conditions) and the inclusion of a finitely lived experiment. In order to improve the identification of unobserved state variables the framework is extended in this paper. In particular, moments which are only indirectly linked to state variables, and therefore not directly computable from the distribution over states in a given period, are added to the set of moments which are used in estimation. The extension includes introducing an "inner Markov chain" to the solution algorithm outlined in Ferrall (2012), which calculates the distribution of e.g. employment duration over time although employment duration is not a state variable in the model. The modification shows how further moments can be added to the model without increasing the state space or having to simulate the model. The extension improves the estimation of the transition probabilities for unobserved state-variables as it increases the number of predictions of the model which can be compared to corresponding data moments, for instance moments describing the distribution of employment durations are informative about the interaction between skills and job separations.

The estimates suggest that the cost associated with programme participation is non-negligible, in particular unemployed would be willing to decrease UI benefits in a given week with up to 50 % in order to escape ALMP participation. The size of the utility costs are just below the lowest possible sanction individuals may receive if they do not participate in ALMPs. The average CV associated with the experiment is up to 20 times larger than the monetary costs associated with programme participation. The analysis shows that the CV varies with future prospects, in particular it is smaller for individuals where alternative choices are more valuable - for instance in the case of high skilled versus low skilled workers. The high average CV is partly driven by individuals with low employment prospects who need larger compensation.

Using detailed information on the benefits and costs associated with the experiment under investigation the paper presents a welfare analysis which includes the costs associated with the loss of leisure in relation to both increases in employment rates *and* due to an increase in participation in ALMPs. The size of the compensating variation implies that the gain from the most favourable intervention (meetings at the job centre) is reduced by 50 % while the other intervention (early activation) is associated with only a small increase in welfare. The welfare analysis thereby illustrates the importance of including more aspects that just direct programme costs in an assessment of the optimal level of ALMPs in the labour market.

This paper proceeds as follows: the next section contains some background and a review of the related literature. Next the experiment and the available data are presented. The following section contains some key features of the data which the model will try to incorporate. Then the

[8]Ferrall (2012) studies the Self-Sufficiency-Project in a structural model and develops a framework which incorporates the non-stationarities implied by the design of the experiment. He use the model to study how the SSP affect incentives for low wage workers and whether the policy enables them to escape the "poverty trap". The model includes a waiting period and a qualifying period where potential participants must obtain work to qualify for a wage subsidy. The analysis illustrates that these non-stationarities are crucial in interpreting the experimental impact. Furthermore the paper shows how a well-defined structural model which incorporates these non-stationarities substantially improves out of sample predictions, the overall fit of the model and thus any policy recommendations.

model and the empirical implementation are presented. The final sections contains results and a conclusion.

1.2 Background and Related Literature

Policy makers have become increasingly focused on adverse selection and moral hazard in relation to UI as the empirical relevance of such phenomena has been documented in the literature (see e.g. review by Chetty and Finkelstein (2013)). Several countries, and especially Northern European countries (see e.g. Andersen and Svarer (2007)), have introduced programmes targeting UI recipients such as meetings, job search assistance and workfare/activation programmes in an attempt to re-align incentives, reduce moral hazard and improve market functioning. By some this is referred to as 'active social insurance' (Roed (2012)) to underline that UI is not only a passive transfer of income, but instead participation in these programmes serve as a conditionality for receiving benefits.[9] ALMPs can have two very different aims: i) improve the qualification level of the unemployed through e.g. counselling or training and thus improve future job possibilities, or ii) they serve as mechanisms for ensuring that the unemployed are actually available and searching to get out of unemployment. The latter objective is often mentioned as an important component as the empirical literature has found limited relevance of the first aim - especially in the case of traditional training programmes (see for instance Heckman et al. (1999) and Kluve (2010)).

In this paper ii) is rationalized as a utility cost associated with programme participation while i) enters through an increase in job offer arrival rates immediately after programme participation. The cost might consist of several policy invariant parameters such as stigma or disutility associated with participation (see e.g. Moffitt (1983)), loss of leisure and an increase in effort in order to attend meetings at the job centre.[10] Although ALMPs might be successful in reducing moral hazard in the market by increasing e.g. search activity, any costs associated with programme participation challenges whether these programmes actually make individuals better off - or whether they would instead prefer lower benefits. These costs imply that some individuals are worse off than before the introduction of ALMPs, this is in fact why some search more to leave unemployment before being activated, while at the same time the market is now more efficient. The overall implications for welfare are therefore less clear.

There is very little empirical work trying to quantify costs or assess the welfare implications of programme participation. Greenberg and Robins (2008) provide estimates of the value of lost leisure for participants in the Self-Sufficiency-Project in Canada. This enables them to quantify

[9] One example of this is the Danish labour market model were UI is generally generous and the level of employment protection is quite low. The sustainability of such a system could be challenged by high structural unemployment rates, e.g. due to low incentives for workers to leave unemployment. Therefore ALMPs are considered a crucial part of the model and participation in such programmes is considered both a right and a duty (see e.g. Andersen and Svarer (2007)).

[10] Search activity could also change due to the fear/risk of getting a sanction for non-compliance with the search requirements (see below). In the model presented below search activity will change as a response to utility costs, there is no risk of getting a sanction in the model.

the gain in consumer surplus instead of the raw income gain associated with the wage subsidy.[11] The authors find that when the loss in non-market time is taken into account, the net benefits from that policy is substantially reduced and sometimes even negative. Their analysis thereby provides further empirical justification for why knowledge of how participants value their time in different settings should be of central interest in the literature and in the evaluation of programmes.

The analysis in Greenberg and Robins (2008) is different in a number of dimensions compared to the current paper. First, as participation is voluntary, participants prefer participation and the size of the subsidy is used as a reflection of the value generated by the programme. A similar expression for the value of the programme does not exist for the experiment presented below - here participation is an obligation. The value of the programme will therefore have to be determined indirectly through changes in behaviour such as job finding rates. Second, while Greenberg and Robins (2008) use the size of the subsidy as an expression of the value created by the programme, in the current paper exogenous variation in treatment status is exploited to compare behaviour between treated and non-treated. This source of variation, and the structural model, allows for a quantification of a broader concept of costs including fixed costs associated with actual participation. The model generates a mapping of different channels of behaviour into decision parameters and therefore exploit differences in behaviour along other channels than wages only to learn about the size of costs.

Below other related theoretical and methodological literature is discussed.

Other related theoretical and empirical literature

The theoretical literature has analysed how and whether conditionalities such as workfare *can* in fact improve welfare in a setting where society has a preference for redistribution. In summary there exists normative work on whether, and under which conditions, conditionalities in benefit recipiency are welfare improving. The theoretical literature has studied two different margins of behaviour. Both along the extensive margin,[12] i.e. the selection of individuals into unemployment, and along the intensive margin,[13]i.e. behaviour while in unemployment (e.g. job search),

[11] Using a matching procedure they identify the group of compliers in the experiment (the part of the treatment group which enters employment caused by the subsidy). For this group of workers they use the earned wage in employment, w^* (including the subsidy) and the same wage without the subsidy, w^n. The two observations and economic theory can be used to bound the individual labour supply curve. The analysis exploits the fact that the individual reservation wage for starting to work must be above w^n - as the compliers do not work at inflow into the experiment - and thus by adding assumptions about the value of w_R and the curvature of the labour supply curve the authors can calculate the part of the gain in income which is offset by increased effort.

[12] Besley and Coate (1992) show that while conditionalities (costly unproductive activities) improve market functioning and redistribute income to 'needy' individuals, this does not imply that agents are better of in terms of utility. In particular the work requirement implies a cost of leisure which is high enough to offset the increase in benefits. Kreiner and Tranaes (2005) show that in an environment with voluntary and involuntary unemployment, workfare can be an effective screening device for UI and lead to a Pareto improvement in the economy. The main difference to the setting in Besley and Coate (1992) is that the screening problem is now focused on individuals who differ in their preference for leisure and not in terms of productivity. Other papers that analyse settings where conditionalities can be welfare improving are Cuff (2000) and Beaudry et al. (2009).

[13] Andersen and Svarer (2014) focus on the effects of workfare on moral hazard in job search in a search and matching model. To study behaviour along the intensive margin their framework is dynamic, and their analysis shows that the threat of future participation in workfare increases the search effort of the unemployed before

behaviour may change with the introduction of workfare. The literature shows that workfare can be welfare improving in some settings but it depends on the environment, the nature of costs and the margin on which behaviour is studied.

A number of other papers have analysed how labour market programmes affect individual behaviour in the labour market in a theoretical and empirical framework (see also Cohen-Goldner and Eckstein (2010); Albrecht et al. (2009a); Lamadon et al. (2004)). Adda et al. (2007) develop a structural dynamic model of labour supply to study the impacts of the Swedish labour market programmes. The study differs in a number of ways from the current one, most importantly programme participation is voluntary in their setting and without costs. The model is used to solve the self-selection problem into programme participation and analyse programme impacts on earnings and job offers.[14] Van Den Berg and Van Der Klaauw (2006) analyse how counselling and monitoring programmes affect the transition rate into employment in a Dutch setting. They show theoretically that incomplete monitoring of job search can have adverse effects as individuals substitute search towards formal (and measurable) search channels and away from informal search. They compare the predictions to results from a social experiment which includes a survey about search channels and find some evidence of substitution effects. The paper is focused on the impact of closer monitoring on different search channels and the existence of individual costs beyond the costs of searching or the implications for welfare are not analysed. The impacts on employment from the studied intervention are small and the authors explain this by inefficient targeting of the programme and a low intensity of treatment. They argue that a too excessive focus on the monitoring of job search activity is inefficient and that alternative policies such as 'leisure taxes' may be more efficient.

Summary of literature and relation to model

The presentation above have shown that while there exists some empirical and theoretical work on how labour market programmes affects both participants and non-participants there exists very little work focused on the existence of individual costs and their implications for the attractiveness of these programmes. The theoretical literature shows that workfare can be welfare improving in some settings but it depends on the environment, the nature of costs and the margin on which behaviour is studied.

This paper exploits changes in behaviour along the intensive margin to identify the cost associated with programme participation (e.g. changes in unemployment duration). Since the experiment is an unexpected event and does not change the inflow into unemployment and because employment separations are exogenous in the model, there will be no characterization of how selection into unemployment depends on the existence and intensity of ALMPs. It is however perfectly plausible

actual participation and lowers his reservation wage. Under a utilitarian criterion the authors show that workfare can in fact improve welfare.

[14]The type of programmes under investigation are more traditional training or job-experience programmes with longer durations. By participating in the programmes participants renew their eligibility to UI. The authors show that by abolishing the latter rule, welfare can be increased as the efficiency of the market increases (as moral hazard is reduced). In line with previous literature they find limited effects from job training programmes and modest impacts from job experience programmes.

that behaviour on both margins is driven by the the same cost (this requires that we disregard any fixed costs associated with entry into UI which depends on the intensity of future ALMPs), but naturally predictions of behaviour along the extensive margin requires a quantification of all the decision parameters related to this decision.[15]

Finally, there is no monitoring of search activity in the model presented below. One difference between this paper and earlier work related to sanctions and monitoring (see also Fredriksson and Holmlund (2006)) is therefore that the model below associate a cost to utility to each meeting at the job centre whereas the earlier literature attributes all changes in behaviour to the disutility in the case where individuals are sanctioned. The two formulations generate similar behaviour but the former is directly linked to current periods costs. While in reality both explanations are probably relevant to explain the increases in the job finding rate I report below, it is beyond the scope of this paper to separate the two. Furthermore, as the sanctioning rate is very low in the Danish labour market (see e.g. Svarer (2011)) and the stated intention of the treatments stated below was no intensification of monitoring, this could suggest that the risk of getting caught is maybe the less relevant channel.

1.3 Data, Institutions and the Experiment

This section presents the Danish institutional setting, the social experiment and the data used in the analysis. The Danish labour market is rather flexible and is referred to as an example of the Flexicurity model.[16] It has less employment protection legislation than most continental European countries and much higher labour turnover (see e.g. OECD (2009)). At the same time a tight social security net with near-universal eligibility for income transfers keeps income security high. Finally active labour market policies are seen as an important part of this model.[17] Today ALMPs are among the most intensive in OECD, with around 1.3% of GDP spent per year on active policies and more than 12 billion Dkk on ALMPs alone (see Board (2014)). There are two types of benefits for unemployed workers, UI benefits and social assistance. Approximately 80% of the labour force are members of a UI fund and therefore eligible for UI benefits. The remaining 20% may receive means tested social assistance. The policies that apply to UI recipients are presented below, they constitute the target group of the experiment presented in the next section. UI benefits are essentially a flat rate due to an upper bound on payments (see e.g. Lentz (2009)) and the duration of benefits in the period under study (2008-2009) is 4 years. A 'right and duty' principle governs labour market policies. Unemployed individuals have the *right* to compensation for the loss of income, but also the *duty* to take action to get back into employment and follow instructions from the job centre (public employment service). Interactions between public authorities and unemployed individuals take place in job centres and activities are mainly

[15] Due to data limitations such a quantification is outside the scope of the current paper, in particular data on the reason for employment separations would be required to model this margin.

[16] Before the financial crisis the EU commission recommended this model to its member states, referring to Denmark as a model case (Commission (2007))

[17] The 1980s were characterized by persistently high unemployment rates and a low intensity of ALMPs. As the intensity of ALMPs grew, structural unemployment fell, and therefore observers have seen intensive ALMPs as an important part of the Flexicurity model (see e.g. Andersen and Svarer (2007)).

contact (meetings) and activation (see Maibom et al. (2014)). At inflow into unemployment a UI eligible individual has to register at the local job centre. She then has to attend a meeting with a caseworker every 3rd month and to participate in an activation programme after 9 months (6 if below 30 years old) of unemployment and subsequently every 26 weeks. For the experiment outlined below these are the labour market policies that will be faced by individuals in the control groups. Treated individuals are obliged to participate in further activities beyond the activities presented here.

In order to increase the knowledge about the effectiveness of current labour market policies the National Labour Market Authorities have conducted a series of experiments. Evaluations have established that there are potentially favourable gains from earlier and intensive active labour market programmes (ALMPs) in the form of either meetings or activation programmes (see e.g. Graversen and van Ours (2008a); Maibom et al. (2015)). But, importantly the evaluations says nothing about the effect of these interventions on welfare.

Experimental Design

The experiment was conducted in two different regions in Denmark in 2008. Each region had a separate treatment (either an intensification of individual meetings or early activation) and each region also had their own treatment and control group. The experiment is presented and analysed in Maibom et al. (2015) and I refer to their paper for details on the setting beyond what is presented below.[18]

The target population of the experiments were UI eligible individuals who became unemployed during weeks 8-29 in 2008. The assignment to treatment or control groups was based on the date of birth. Individuals born on the $16^{th} - 31^{st}$ were assigned to the treatment groups, while those born on the $1^{st} - 15$ were assigned to the control groups. No information was given to the unemployed workers on the selection rule. Once immigrants are excluded from the sample Maibom et al. (2015) find no deviations from random assignment, and therefore I treat it as such. See also Appendix B Table 1.13, for balance of means tests and descriptives.

At inflow into the experiment treated individuals received a letter explaining the new treatment to which they will be exposed. The information letter marks the start of the treatment, since the worker may react to the information on the new regime. Table 1.1 presents an overview of the activities in the treatment group beyond the regular activities presented above. Individuals in the treatment group from the region around the capital city, Copenhagen (R1), had to participate in individual meetings with a caseworker every other week for the first 13 weeks of unemployment, a total of 6-7 meetings during the first 13 weeks of the experiment. The stated intention of the individual meetings was counselling of the unemployed - no extra monitoring was required to take place, but naturally this says nothing about the perception of the meetings from the point of view of the unemployed nor the actual content. Individuals in the treatment group from the region around the second largest city, Aarhus (R2), were required to participate in an activation

[18] The experiments investigated here were a part of a larger experiment 'Quickly Back to Work 2' which consisted of four separate experiments, each with its own treatment and control group. See Maibom et al. (2015) for details.

programme for at least 25 hours per week from week 14 in unemployment until week 26. This experiment - the activation wall - was designed specifically to investigate the presence of ex ante effects due to the knowledge of having to participate in an activation programme, as well as ex post effects of actually having participated.[19,20]

From the presentation it is clear that the experiment have some important features which should

Table 1.1: Content of the experiments

Weeks	Meetings (R1)		Activation (R2)	
0-1	Recieve Information Letter	W	Recieve Information Letter	W
1-13	Individual Forthnightly Meetings	T		W
14-26		PT	Participation in activation programme	T
26-		PT		PT

Note: The table presents the content of activities individuals in the treatment group has to participate in *beyond* any regular activities (see above). R1 denotes the meetings region and R2 the region with activation.

be incorporated into the structural representation to model the incentives faced by unemployed workers accurately - and thus estimate key decision parameters credibly. In particular, the unemployed treated individuals progress through three different phases with different duration (phases are outlined in Table 1.1): i) a waiting phase (W) which starts with the information letter and stops when actual treatment begins (in R1 this constitutes 1-2 weeks and in R2 this will be 13 weeks), ii) the actual treatment phase (T) and iii) ex-post treatment (PT) which marks the end of the experiment. The model presented below is set up to account for the fact that incentives change as individuals progress through the experiment. For instance individuals might be more likely to increase their search effort as T approaches and similarly the incentive to leave unemployment declines as PT approaches and the future intensity of activities declines.

Data and Definitions

The data are extracted from administrative registers merged by the National Labour Market Authority into an event history data set, which records and governs the payments of public income transfers, records participation in ALMPs, and has information on periods of employment. The data includes detailed weekly information on: labour market status and history (employment, unemployment, in education, on leave, etc.). Labour market status is calculated based on information from the register on payments of public income transfers. This data is subsequently merged with two other datasets BFL and IDA[21] in order to obtain further information, in par-

[19]Note that in order to test specifically for the ex ante effect in an experimental setting, there should have been no actual treatment taking place from week 13 onwards. For our analysis the assumptions implied by the model allows us to test for the existence of such effects namely through the presence of a substantial utility cost.

[20]An important advantage of the available data in Maibom et al. (2015) is that it allows evaluators to assess the extent to which the planned treatment was actually implemented. Their analysis documents that the intended treatment was implemented to a large extend. There are also some deviations from perfect compliance as the meetings and activation intensity is not as high as planned (80% versus 100 % by design). While there can be several explanations this issue is ignored below as agents might still react solely to the threat of participation. This corresponds to assuming that non-participation in treatment in a given week is truly exogenous and unexpected (for instance due to administrative changes or other events).

[21]IDA: Integrated Database for Labour Market Research. IDA is a matched employer-employee panel containing socio-economic information on the entire Danish population. Both persons and firms can be monitored from

ticular monthly wages before taxes, hours and the education level of workers.[22]
The raw sample excluding immigrants consists of 3385 individuals who are either assigned to the treatment or control group. To have a more homogeneous sample I disregard workers below the age of 22 and above the age of 58.[23] This leaves 3099 individuals in the sample. The data is divided into sub-groups depending on the educational level of the individual. There are 3 educational levels: low (individuals with only primary education and less than 12 years of education), medium (individuals with vocational education and 12-14 years of education), high (individuals with further education and above 14 years of education). Table 1.2 shows the division into sub-groups defined by region, treatment status and education levels.

The final data identifies individuals in any public support schemes at a given point in time -

Table 1.2: Number of observations

Education Groups	Low	Medium	High
Control (R1)	211	376	137
Treatment (R1)	212	399	141
Control (R2)	102	298	396
Treatment (R2)	92	307	428

R1: meetings region, R2: activation

these will be unemployed in the model presented below. The data used does not allow for a meaningful distinction between individuals in regular employment (where the registers contain wage information etc.) and individuals who are in a residual 'self-sufficient' group where there is no information on either wages or public support (this group contains self-employed, black-sector workers and workers out of the labour force). Individuals transitioning to the self-sufficiency state are therefore treated as individuals transitioning into employment as these are individuals who have opted out of any public support scheme (UI eligibility is 4 years at the time of the experiment).[24] Figure 1.16 in Appendix B shows how the fraction in the residual group evolves over time. Unsurprisingly changing the outcome makes the impact of treatment a little larger, but I show below that the important data features are similar regardless of the used employment definition.
In the next section I provide more details on the impact of the experiment in relation to the model developed below.[25]

1980 onwards. BFL: Employment Statistics for Employees. BFL contains monthly data on jobs, paid hours of work and total wage to employees throughout the year. BFL is available from 2008 and onwards. Both data sets are available through servers at Statistics Denmark (see dst.dk).

[22]The analysis below uses wages after imputed taxes, assuming a tax rate of 37.5 % for all workers (this corresponds to the average tax rate for individuals on UI in 2008, see Maibom et al. (2015))

[23]The age-restrictions allow me to ignore decisions about retirement and entry into education. I treat entry into educatoin after the age of 22 as any other public support scheme (in the data less than 4 % of workers transit into some kind of education which is supported by the state)

[24]The definition of employment is thereby slightly different from the definition used in Maibom et al. (2015) as the model thus captures the decision of whether to stay in public transfers or not. In Maibom et al. (2015) time spent in the employment state is compared across treatment and control group. Individuals not in employment are either self-sufficient or in public support.

[25]In general, the findings in Maibom et al. (2015) is that meetings lead to a significant increase in employment rates. Furthermore a positive and statistically significant effect on accumulated weeks spent employed remains

1.4 Data and model

This section presents features of the data which serve as the motivation for the specification of the model presented in the next section. Table 1.13 in Appendix B shows average characteristics for treated and controls in each region and the p-value associated with a test of equality of means. In general the sample is balanced both in terms of past earnings, demographics and employment history. The descriptives show that while the experiment was directed at newly unemployed workers, 20-30 % of the participants came from other states than employment (e.g. education, unemployed or previously sick-listed). To interpret the generated impacts of the experiment it is therefore important to keep in mind that some of the treated individuals were in fact previously unemployed which could affect the size of impacts. There are also some important regional differences (the distribution across cells in Table 1.2 also differs depending on the region) which implies that comparing impacts across regions requires further assumptions. The estimated structure of the model can be used to analyse the sensitivity of the raw impacts to these differences.

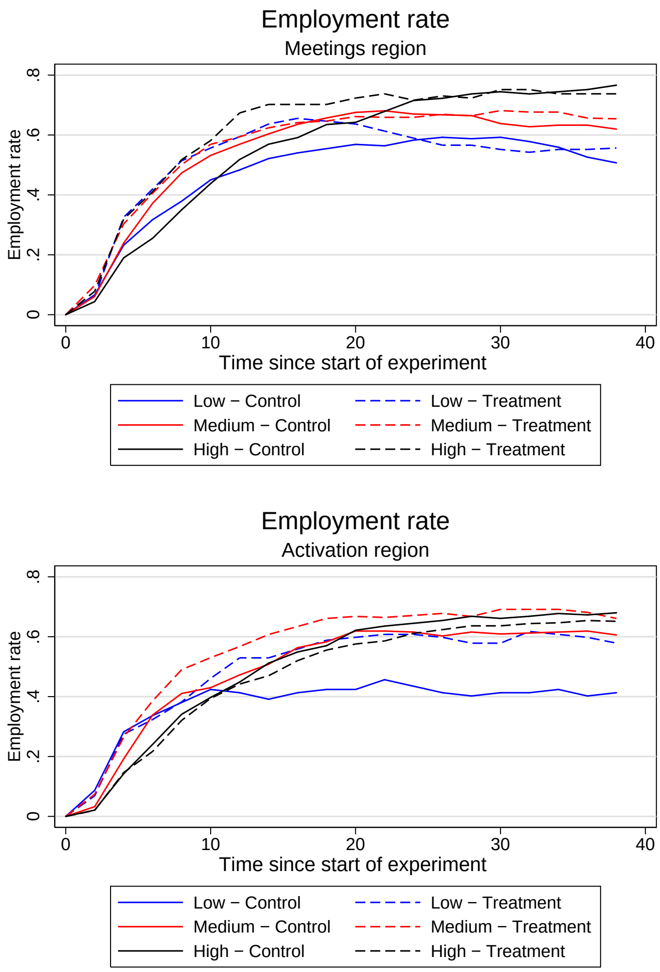
Data patterns

Figure 1.1 shows the employment rates from inflow into the experiment and onwards for the two regions. The figure shows that employment rates increase rapidly within the first 20 weeks hereafter the employment level stabilizes. There are educational differences in the inflow rates and in the "stable" employment levels. In particular there is a clear educational ordering in the employment level after 30 weeks: the employment rate is around 70% for individuals with high education, and slowly increasing, whereas the employment level is around 55% (40 %) for individuals in the medium (low) group and stable or slightly deceasing. The hazard rate out of unemployment for the control groups (see Figure 1.2) is declining with duration in unemployment. This implies that even in relative terms the initial outflow is high compared to the pool at risk.

Figure 1.1 also shows some interesting differences between treatment and control groups. In particular across both regions (with one exception) it appears that treated individuals are in employment to a larger extent. The difference is large initially and then it substantially decreases over time, except for low educated in the activation region. Table 1.3 show the result of a regression of employment status on treatment status for different regions and time periods. The table shows that already after 2(4) weeks in the experiment treated individuals in the meetings region (R1) are significantly more employed. At this point unemployed individuals *may* have participated in 1(2) meeting(s) and therefore the results indicate either a very productive first meeting or the presence of ex ante or threat effects. In the activation region (R2), where treated individuals only start participation in activation after 13 weeks (see Table 1.1), the results are more mixed after 4 weeks. When I run the same regression 10 (14) weeks after inflow into the

significant over the whole 5 year horizon studied. The activation wall produces results which are positive but insignificant, but for certain subgroups and especially young workers the impacts are large and statistically significant. Estimates from a duration model suggest both the presence of effects ex ante and subsequent employment duration effects. There are also interesting gender differences where females generally respond faster than males.

Figure 1.1: Employment rates and inflow (treated and controls)



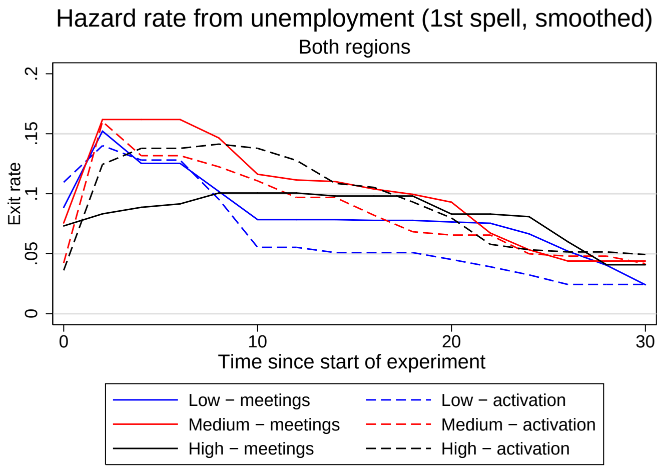
Note: time since start of experiment is measured in weeks

experiment the results are much larger for treated individuals in R2. The regressions therefore suggest that the timing of the treatment is important and that differences are large in the very early stages of treatment which could be a combination of both threat effects and programme effects. The fact that the effects accumulate this early (and also before treatment starts) indicates that the existence of a utility cost could be an important channel.[26]

Figure 1.3 shows the average hourly wage as a function of duration in employment. Wages generally increase with employment duration. The level and the growth rate of wages differ by

[26] As earlier mentioned the "employment criterion" used here defines anyone who do not receive public support as employed. Table 1.15 in Appendix B performs the same analysis using a stricter employment criterion which was also used in Maibom et al. (2015). The effects are very similar and the main findings and significance remains although some of the effects are smaller in magnitude which suggests that a part of the response to treatment goes through self-sufficiency or self-employment and then later employment (a part of individuals in self-sufficiency could also be employed due to data limitations).

Figure 1.2: Hazard rate for individuals in the control group



Note: time since start of experiment is measured in weeks

Table 1.3: Employment results

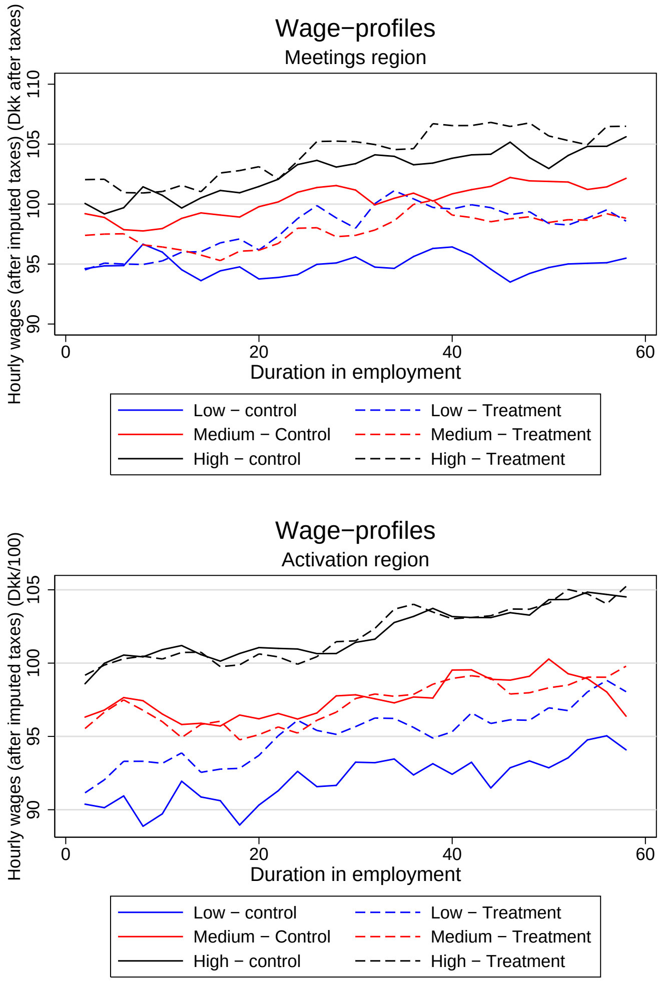
	Meetings			Activation		
After 2 weeks	Low	Medium	High	Low	Medium	High
Treatment indicator	0.0887*	0.0620*	0.123*	0.0256	0.0634+	0.0257
	(0.0410)	*(0.0314)*	*(0.0506)*	*(0.0627)*	*(0.0349)*	*(0.0265)*
After 4 weeks						
Treatment indicator	0.102*	0.0446	0.120*	-0.0156	0.0496	-0.0274
	(0.0468)	*(0.0349)*	*(0.0567)*	*(0.0690)*	*(0.0384)*	*(0.0298)*
After 10 weeks						
Treatment indicator	0.120*	0.00197	0.149*	0.115	0.105*	-0.0408
	(0.0481)	*(0.0356)*	*(0.0589)*	*(0.0717)*	*(0.0405)*	*(0.0343)*
After 14 weeks						
Treatment indicator	0.130*	0.0351	0.133*	0.166*	0.0936*	-0.0547
	(0.0473)	*(0.0348)*	*(0.0577)*	*(0.0712)*	*(0.0399)*	*(0.0348)*
Observations	423	775	278	194	605	824

Note: The results are from separate OLS regressions after 2, 4, 10 and 14 weeks. The dependent variable is employment status. Huber/White standard errors, $+\ p < 0.10$, $*\ p < 0.05$

education, and there is also variation within educational groups (the standard deviation is around 20-25% of the mean). Wage-profiles in treatment and control groups are generally similar, but wages seem slightly higher (lower) for low (medium) educated individuals in the treatment group. Table 1.18 in Appendix B shows the results from a regression of wages on treatment status after 10 weeks in employment.[27] The results show that the differences across groups are insignificant except for medium educated individuals in the meetings region.

[27]Differences in wage profiles (or lack of) can also be flawed by selection as treatment status is no longer exogenous in post-unemployment spells. In the presence of any impact or behavioural change associated with the experiment the composition of individuals in employment will differ between control and treatment.

Figure 1.3: Wage-profiles for employed workers



Note: time since start of experiment is measured in weeks

Summary and relation to model:

The model contains different explanations for decreasing outflow rates and differences across education levels documented in Figures 1.1 and 1.2. These are duration dependence in job offer probabilities, differences in wage offers and differences in preferences (both in terms of observables and unobservables). The estimated parameters will be informative about what drives the declining pattern.

Changes in the average wage of employed workers in the model can be driven by two explanations: dynamic selection out of employment as low wage individuals leave for unemployment or true skill gains which imply higher wages. The model allows for these features through a search sensitive component of wages (different wage offers) and stochastic skill accumulation while employed. The skill level will be unobserved to the econometrician and changing over time. Differences in wages and wage growth will be important for how individuals value employment (and therefore also lead to differences in the compensating variation associated with programme participation).

1.5 Model

This section presents the model in more detail. Each subsection presents different elements: The dimensions of actions and heterogeneity (state variables). The different primitives of the model: the utility function, the wage function and the evolution of skills. The decision rules which determine individual behaviour, and finally the timing of the model. The next section explains how the model is solved and the estimation proceeds.
Individuals in the model, are forward looking and infinitely lived. They maximize the discounted sum of all future pay-offs by making discrete choices in a dynamic environment. The environment is stationary and ergodic (conditional on state variables) and a time period in the model corresponds to 2 weeks in the data.[28] State variables are discrete and the transition probabilities for state variables depend on the characteristics of the agents in ways that will be specified below.[29] The environment is characterized by duration dependent job offer rates, search sensitive wages, stochastic skill accumulation in employment and depreciation at inflow into unemployment. Employed individuals face a probability of a lay-off which is independent of individual choices but depends on their skill level. Unemployed receive UI and participate in ALMPs which consist of two elements: meetings and activation. Participation in a given programme is associated with a potential loss of utility while it can also increase the probability of receiving a job offer. To estimate these components a non-stationary and finitely lived experiment is introduced into this environment (see more below). Differences in technology and preferences generate heterogeneous impacts of the experiment and therefore heterogeneous treatment effects are endogenous to the model.

Choices and State Space

Table 1.5 contains an overview of the parameters to be estimated, this entails preference parameters and parameters which affect the transition of stochastic state variables. Table 1.17 in

[28]To maintain focus on the individual behaviour and utility costs the model is cast in partial equilibrium. The inclusion of GE effects is still seldom in the literature and the interventions considered here are relatively short which could complicate the analysis further as firms simply do not have time to respond to the change in the environment (alternatively they may know that the intervention is temporary). Gautier et al. (2012) consider the GE effects in an earlier Danish experiment where the treatment period and intensity is longer. See also Lamadon et al. (2004) who focus on the Self-sufficiency Project conducted in Canada. They calibrate a search and matching model using data on the control group and use the data on the treatment group to validate their predictions about the equilibrium effects of the SSP.

[29]The model presented below use the same overall framework as in Ferrall (2012, 2002) which also contains the necessary assumptions and requirements to the primitives (e.g. environment, transition functions and utility) enabling the researcher to solve the model and deal with a problem of initial conditions in an environment with unobserved state variables which evolve in a non-iid fashion. While developing model primitives it is therefore ensured that these assumptions are met. Primarily this implies ensuring the existence of an ergodic distribution - the main requirement is that transitional dynamics for each state variable is either ergodic, invariant or dependent.

the appendix provides an overview of other model parameters which are not estimated (e.g. the meetings intensity). The next subsections contains more detail on how primitives of the model depend on parameters and state variables.[30]

Let α contain current actions and let θ contain the value of the state, i.e. the collection of variables which summarize all information about the past needed in the forward-looking optimization problem. The action space consists of two variables: a search activity choice ($ac \in \left\{0, \frac{1}{7}, .., 1\right\}$) and a working status choice ($wc \in \{0,1\}$) if a job offer arrives. Individuals choose search activity along both the extensive margin and the intensive margin while they only make a choice at the extensive margin of employment if a job offer is available - the intensive margin (e.g. hours worked) is assumed fixed and constant across jobs.[31] When employed there are no choices to be made and any potential wage increase is explained through stochastic skill accumulation (this also includes changes related to job to job transitions in the data). While there could be important effects from job quieting behaviour the data will not allow us to determine the reason for job separations.

The state space (θ) summarizes all relevant information in the environment influencing individuals in their decision making (wages, employment status, wage process). Table 1.4 contains an overview of the elements of the state space which consists of a collection of state variables describing the "normal" social environment and another collection of state variables which describe the experiment, thus $\theta = (\theta_{\text{enviroment}}, \theta_{\text{experiment}})$.

$\theta_{\text{enviroment}}$ consists of a time-invariant part and a time-varying part. While the time-invariant part of the state space is unaffected by choices made by individuals, time-varying states may change as a result of choices. Since a part of the time-varying state space evolves stochastically, individuals do not know the future position of the state space θ' with certainty but form rational expectations. The time-invariant state variables divide individuals into experimental, educational, regional and 'preference for leisure' groups. A state variable **g** marks the treatment status of individuals (control, treatment), **e** marks the education level of the individual (low, medium or high skilled) and **r** the region (R1 or R2) - all variables are observed by the econometrician. The environment is also composed on a (finite) number of types (**k**) who differ in how costly searching for a job is (or their value of leisure). Type status is unobserved to the econometrician and the distribution of types differ across educational and regional groups.

Time-varying state variables are variables for unemployment duration (**cu**), meetings or activation participation status (**mp**/**ap**), a potential job offer (**j**), the level of skills/human capital (**hc**) and the employment status of the individual (**em**). Unemployment duration (**cu**) counts the duration of the current unemployment spell (since last job loss). The meetings (activation) variable (mp/ap) indicates whether individuals currently participate in one of the programmes. A job offer (j) is a draw of firm productivity which is mapped into an actual wage offer through

[30] $\underline{\kappa}^e$ implies that κ is a vector with an education specific entry (see Table 1.5). In this case this implies that the cost associated with working is estimated separately for each education group.
#points(hc) gives the number of different values the state variable hc may take. State variables always take values $\{0, 1, ..., points(hc) - 1\}$ unless otherwise stated.

[31] To model the intensive margin of employment further characteristics of the employment situation would be necessary, for instance detailed data on working hours, other benefits and tax schemes.

a wage function. Wages are also a function of the current level of skills (hc).
The experiment is included into the model by adding two state variables (collected in $\theta_{\text{experiment}}$) to the state space. These state variables serve as "accounting variables". They cosists of a treatment phase indicator (**p**) and a counting variable for the time spent in the current phase (**c**). The inclusion of these variables allows the incentives to change as individuals progress through the experiment: for instance the incentive to leave unemployment may increase as the individual progress through the "waiting phase" knowing that in 6 periods an early activation scheme begins.

Table 1.4: Elements of the state space

Sub-space	State variable	Symbol	Type	Transition	#points()	Data
$\theta_{\text{enviroment}}$	Education group	e	Time-invariant	none	3	Observed
$\theta_{\text{enviroment}}$	Regional group	r	Time-invariant	none	2	Observed
$\theta_{\text{enviroment}}$	Treatment group	g	Time-invariant	none	2	Observed
$\theta_{\text{enviroment}}$	Effort Type	k	Time-invariant	none	2	Unobserved
$\theta_{\text{enviroment}}$	Unemployment duration	cu	Time-varying	deterministic	10	Observed
$\theta_{\text{enviroment}}$	Job offer	j	Time-varying	stochastic	8	Unobserved
$\theta_{\text{enviroment}}$	Meetings status	mp	Time-varying	stochastic*	2	Observed
$\theta_{\text{enviroment}}$	Activation status	ap	Time-varying	stochastic*	2	Observed
$\theta_{\text{enviroment}}$	Skill level	hc	Time-varying	stochastic	6	Unobserved
$\theta_{\text{enviroment}}$	Employment status	em	Time-varying	stochastic	2	Observed
$\theta_{\text{enviroment}}$	Lost job	l	Time-varying	stochastic	2	Observed
$\theta_{\text{experiment}}$	Treatment phase	p	Time-varying	deterministic	3	Observed
$\theta_{\text{experiment}}$	Clock (time in current p)	c	Time-varying	deterministic	6	Observed

* in the treatment phase π_{mp} and π_{ap} are set to 1. #points(hc) gives the number of different values the state variable hc may take. State variables always take values $\{0, 1, ..., points(hc) - 1\}$ unless otherwise stated.

Utility, Costs and Wages

The current payoff is described as a function of generated income and costs (pecuniary and non-pecuniary):

$$U(\alpha, \theta) = \log(Income(\alpha, \theta) - Cost(\alpha, \theta))$$

The formulation keeps costs in monetary units while ensuring that agents are risk averse - thereby an insurance motive can exist in the economy.[32]Costs vary with education levels (e) and unobserved type (k). They depend on effort and mandatory programme participation:[33]

$$\hat{Cost}(\alpha, \theta) = \xi^{\underline{\psi}^k} \cdot ac + \underline{\kappa}^e \cdot wc + \underline{\phi}^e_{ap} \cdot ap + \underline{\phi}^e_{mp} \cdot mp$$

[32]The formulation is similar to Shimer and Werning (2008), $u(c_t - v(e))$, if we assume that capital markets do not exist or workers are liquidity constrained such that they consume all income each period. Shimer and Werning (2008) use CARA utility, here I use the log().

[33]Note that in the current version ξ^{ψ_k} would be easier expressed as $\underline{\xi}^k$ (thus just estimate type specific search costs). In a future version costs will be formulated as $cost(\alpha, \theta)^{\psi_k}$, therefore I stick to the separation between ψ_k and ξ below.

Table 1.5: Estimated parameters:

Preference or wage parameters

Symbol	Model	Note	Dimensions
ξ	Utility	Search cost	1
$\underline{\kappa}^e$	Utility	Work cost	dim(e)
$\underline{\phi}^e_{mp}$	Utility	Meetings cost	dim(e)
$\underline{\phi}^e_{ap}$	Utility	Activation cost	dim(e)
$\underline{\psi}^k$	Utility	Leisure preference	dim(k)
$\underline{\pi}^{r,e}_k$	Type	Fraction of type 2	dim(e*r)
μ	Wages	Wage constant	1
$\underline{\sigma}^e$	Wages	Return to J	dim(e)
η	Wages	Return to hc	1
ρ	Smoothing	Smoothing kernel	1

*dim(k): variable varies with the number of unobserved types (2)

** To ensure the existence of an ergodic distribution this parameter must be strictly larger than 0

Transition functions

Symbol	Model	Note	Dimensions
$\underline{\pi}^r_{w,1}$	Job offers	Duration dependence	dim(r)
$\underline{\pi}^e_{w,2}$	Job offers	Long term job offer**	dim(e)
$\pi_{w,mp}$	Job offers	Productive effect (meeting)	1
$\pi_{w,ap}$	Job offers	Productive effect (activation)	1
$\underline{\pi}^e_{lj,1}$	Job loss	Risk of job loss, hc impact**	dim(e)
$\underline{\pi}^r_{lj,1}$	Job loss	Regional effect**	dim(r)
$\underline{\pi}^e_{hc,1}$	Skill level	Appreciation of hc	dim(e)
$\pi_{hc,2}$	Skill level	Loss of hc**	1

*dim(r): variable varies with the number of regions (2)

** To ensure the existence of an ergodic distribution this parameter must be strictly larger than 0

Costs are incurred from exerting effort either through search activity ($ac \neq 0$), working (wc = 1) or participation in programmes. Costs are linearly increasing in the intensity of effort. The education specific cost connected to the participation in ALMPs ($\underline{\phi}^e_i$) depends on the type of programme (i.e. either meetings or activation) as programmes are different in content and scope. The costs associated with working are education specific while the ξ is the same for all individuals. The total cost associated with searching (ξ^{ψ_k}) vary across types and change according to the estimate of $\underline{\psi}^k$ which is estimated separately for each type (k). One type has linear costs as $\underline{\psi}^{k=1}$ is normalized to 1.

Utility is only meaningfully defined when income succeed costs. To avoid taking the log of a negative number and to keep parameters in the relevant area for optimization,[34] costs are expressed as a fraction of maximum attainable earnings for an individual with the highest education level, wage offer and skills. W_{max}[35] therefore does not vary between different types of agents. Total

[34] I.e. in the range where changes in parameters leads to changes in actual behaviour and thus changes in the fit between data and model. Even in the absence of log utility any parametrization of costs exceeding income generates predictions by the model which are indistinguishable. In the optimization process $\xi, \underline{\kappa}, \underline{\phi}_{ap}, \underline{\phi}_{mp}$ are therefore restricted to $[-1, 1]$.

[35] $W_{max}(\alpha, \theta) = \exp\left(\mu + \underline{\sigma}^{e=3}1 + \eta \cdot 1\right)$

costs are therefore expressed as:

$$Cost(\alpha, \theta) = W_{max} \cdot \hat{Cost}(\alpha, \theta)$$

Income consists of the wage when working and UI when unemployed:

$$Income\,(\alpha, \theta) = \begin{cases} W(\alpha, \theta) & \text{if wc=1} \\ UI & \text{if wc=0} \end{cases}$$

When unemployed individuals receive UI which is determined as a fixed amount assuming that all individuals qualify for the maximum amount of benefits. Lentz (2009) estimates that around 90% of the unemployed workers in the labour market qualifies for this amount.[36] UI eligibility is not modelled here since enrolled unemployed are newly unemployed (with some deviations as documented above), the study period is relatively short and eligibility is 4 years in this period. The wage function is similar in some dimensions to Ferrall (2012), and is modelled as:

$$W\,(\alpha, \theta) = \begin{cases} 0 & \text{if j=0} \\ \exp\left(\mu + \underline{\sigma}^e \Phi^{-1}\left(\frac{j}{\#\text{points}(j)}\right) + \eta \cdot \left(\frac{\text{hc}}{\#\text{points(hc)}}\right)\right) & \text{if j>0} \end{cases}$$

μ is a wage constant and represents the deterministic part of wages, η measures the return to skills and $\underline{\sigma}^e$ measures the importance of the frictional or search sensitive component of wages (a draw of firm productivity). The transformation of values of j into percentiles of the normal cdf ensures that the distribution of wages is not uniform and that the wage dispersion does not depend on the dimension of job offers. The presence of a search sensitive component in wages (through different job offers) implies that individuals form reservation wages as optimal stopping rules. The reservation wage will be revised as unemployment duration increase and therefore an analytical expression is not obtainable as in the more standard case (see e.g. Wolpin (1987)). Differences in wages across educational levels are generated by differences in skill accumulation (presented below) and in the return to search. As $\underline{\sigma}^e$ varies across educational groups it allows the within group variance to be different and this is also an important channel through which experimental impacts can differ as the cost of accepting lower wage offers differs depending on the estimate of $\underline{\sigma}^e$.

Jobs and Skills

At inflow into unemployment individuals have no job offers (j = 0), thereafter a job offers arrive each period with probability π_w. Arrival rates are determined as a function of search activity, unemployment duration and programme participation. Following the literature on endogenous search (see e.g. Mortensen and Pissarides (1999)) job offer rates are proportional to search activity:

[36] The replacement level of a worker earning 150% above average earnings is around 0.6, see Bjoern and Hoej (2014). Therefore UI is set to $0.6 \cdot W_{max}$ in the model.

$$\pi_w = \text{ac}\cdot\left[\Phi\left(\underline{\pi}^r_{w,1}\cdot\text{uedur}\right)+\underline{\pi}^e_{w,2}+\pi_{w,mp}\cdot\text{mp}+\pi_{w,ap}\cdot\text{ap}\right]$$

The probability of receiving an offer consists of a regional specific duration dependent term $\left(\underline{\pi}^r_{w,1}\right)$ and constant terms $\underline{\pi}^r_{w,2}, \pi_{w,mp}$ and $\pi_{w,ap}$. $\pi_{w,mp}$ ($\pi_{w,ap}$) represents a potential increase in job offer arrival rates the period after participation in a meeting (activation). The duration dependent term is similar to Wolpin (1987). If employers use duration in unemployment as a screening device, which has been suggested in the literature (see e.g. Kroft et al. (2013) and Belzil (1995)), $\underline{\pi}^r_{w,1}$ will be negative. In this case the first term goes to 0 as uedur increase and $\underline{\pi}^e_{w,2}$ is then the probability that a long term unemployed receives a job offer. Note that the model also allows for "spurious" negative duration dependence in the form of dynamic selection generated by changes in the composition of unobservable types and the stock of skills across remaining unemployed individuals. The observation that outflow rates are declining with unemployment duration (as documented in Figure 1.2) can thereby also result from the more "able" (high paid or low cost) types leaving unemployment early, while the remaining stock consists of a consecutively weaker group of unemployed.

The concept of skills included in this model can be thought of as a mixture of general and specific skills - sometimes skills are transferable to new jobs, other times skills are specific to past jobs.[37] Skills are included to generate differences in the value of a job (both through payment and stability) across agents which are unobserved and change over time. It is an important channel through which the incentive to leave unemployment differs across both time and individuals. While employed the stock of skills appreciates every period with an education specific probability $\underline{\pi}^e_{hc,1}$ reflecting skill improvements through learning on the job. When separated from a job, skills are lost with probability $\pi_{hc,2}$. This captures that acquired skills have become obsolete in the market and therefore expected future wages will be lower for instance because individuals will have to start in a new job without any prior experience in the specific tasks.

Finally, the level of skills also affects the expected duration of a job. Jobs end with probability π_{lj}:

$$\pi_{lj} = \underline{\pi}^r_{\text{lj},2}\cdot\left[\underline{\pi}^e_{lj,1}\cdot\left(1-\frac{\text{hc}}{\#\text{points}\,(\text{hc})}\right)\right]$$

The job separation process is allowed to differ between education levels and regions where the region with meetings is set as the reference category ($\pi^{meeting}_{lj,2}=1$). Job separation probabilities decline (or increase) in how skilled workers are. This generates a source of duration dependence in employment as workers who have been employed for longer periods are also likely to have accumulated more skills and thus less (more) likely to exit to unemployment. The link between skills and job destruction implies that a random sample of workers at inflow at a given point in time will be a selected from the underlying ergodic distribution of workers in terms of skills and willingness to work.

[37] Since I do not focus on human capital accumulation in general, lasting experience or life cycle effects are not included in the model (a time period in the model is 2 weeks). The ergodic distribution of skills is therefore constant over time (while some individuals loose skills and others accumulate skills) although individuals in the sample become older (here 80 weeks). In larger samples workers could potentially be distinguished by age groups to allow for differences in the level of skill.

Active Labour Market Programs

ALMPs enter the model in two ways. Firstly programme participation is associated with extra effort and lost leisure measured by ϕ_i in the utility function.[38] Secondly there can be productive effects ($\pi_{w,mp}, \pi_{w,ap}$) from programme participation through an increase in job offer arrival rates. Individuals have to participate in ALMPs and the only way to escape programme participation is by becoming employed. In both control groups meeting participation is random and happens with probability $\pi_{mp} = 0.15$. The probability of participation in an activation programme is 0 during the first 10 weeks of unemployment, hereafter it increases with unemployment duration until an intensity of 0.35. The parameters are chosen in order to match the meetings and activation intensity in the control group documented in Maibom et al. (2015). The treatment group face the same participation probabilities as the control group in the waiting and post-treatment phase (see Table 1.1). In the treatment phase they participate in programmes with certainty.

Dynamic Program and Choices

The value of a (α, θ) combination at a given point in time is the sum of the current reward and an expected future reward which is affected by current choices and the position in the state space. Individuals have perfect knowledge with regard to the probability distribution from which future realizations will be drawn (each element of α,θ, $U()$ and $P()$ has been presented above):

$$\begin{aligned} \forall \alpha \in A(\theta), \quad v(\alpha, \theta) &= U(\alpha, \theta) + \delta E[V(\theta')] \\ &= U(\alpha, \theta) + \delta \sum_{\theta'} P\{\theta' | \theta, \alpha\} V(\theta') \end{aligned}$$

At each point in time the individual solves this decision problem choosing the actions that give him the highest value. The value function can be determined as:

$$\forall \theta, \; V(\theta) = \max_{\alpha} v(\alpha, \theta) \tag{1.1}$$

Conditional on a position in the state space θ (and ignoring even cases) the model generates a strong prediction about individual behaviour as one action maximizes the equation above. There are two approaches in the literature to allow "observationally" similar agents to make different choices and thus increase the correspondence with real data. One approach adds further dimensions of unobserved heterogeneity to θ while the other introduces uncertainty in the predictions of behaviour ex post. In particular Rust (1987) add a 'taste shifter' - an additive and unobserved continuous state variable to the utility function - while Eckstein and Wolpin (1999) smooth choice probabilities ex post. The procedure followed here is a mixture. Firstly, the existence of discrete unobserved state variables (human capital and wage offers) provides an explanation for why two observationally similar individuals make different choices. Secondly,

[38]Note that in principle ϕ_i could also be negative (and this is allowed for in the estimation) such that programme participation generates utility gains. If this is the case a reverse threat effect may exist for individuals who unexpectedly experience an increase in the intensity of interactions (a so called attraction effect). The empirical literature suggests effects in the opposite direction.

to allow for zero-probability or unanticipated events choice probabilities are smoothed ex post. Choice probabilities are smoothed using a logistic kernel ($\rho > 0$):

$$\begin{aligned} \tilde{v}(\alpha, \theta) &= \exp\{\rho[v(\alpha, \theta) - V(\theta)]\} \\ P\{\alpha|\theta\} &= \frac{\tilde{v}(\alpha, \theta)}{\sum_{\alpha} \tilde{v}(\alpha, \theta)} \end{aligned} \tag{1.2}$$

where ρ determines the importance of smoothing. The smoothing of choice probabilities implies that if the value associated with an in-optimal choice is close to the value of an optimal choice ($\tilde{v}(\alpha, \theta) \approx 1$) the probability of either choice will be similar. Choice probabilities connected to actions which are far from optimal ($\tilde{v}(\alpha, \theta) \approx 0$) will be close to zero. As ρ increase the probability that agents make unexpected/in-optimal choices decrease, as the distance from optimal values receives higher weight and $\tilde{v}(\alpha, \theta)$ is pushed towards zero.[39] Smoothing ex post introduces a wedge between the decision rule agents *anticipate* to follow and what happens in reality. Basically the current formulation allows agents to make zero probability or unanticipated events.

Timing

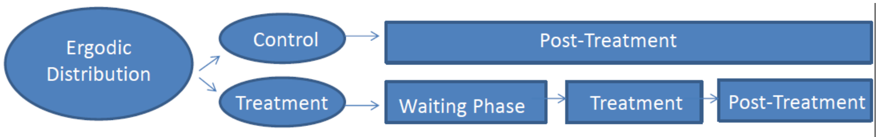
Figure 1.4 illustrates the timing of the model: from the ergodic distribution the outflow from employment into unemployment in a given period is selected into either control or treatment groups. Due to the design of the experiment the distribution over both observable and unobservable states is identical at inflow into the experiment. Individuals in the control group enter an environment without treatment (post-treatment world) and progresses through unemployment making choices according to the structure laid out above. Conditional on $\theta_{enviroment}$ their environment is stationary. For the treatment group this does not hold as the accounting variables in $\theta_{experiment}$ variables c and p change over time which affects the likelihood of present or future programme participation. At inflow into the experiment, individuals in the treatment group enter the waiting phase (see Table 1.1). Their future differs from what was expected at outflow from employment: while unemployed they will go through a waiting phase and a treatment phase before they enter the phase without treatment. In the later phase the environment is identical to the control group, but the distribution over states is potentially different due to the impact of the experiment.

1.6 Solution, Estimation and Identification

This section contains a brief presentation of how the model is solved. It is discussed in more detail how previous work is extended with additional calculations that increases the set of predictions

[39] While the expression of choice probabilities above looks almost identical to the one in Rust (1987) there is one fundamental difference. Here smoothing is ex post while the standard Rust model adds a taste shifter to the model such that individuals take the existence of shocks to utility into account when they solve for optimal values. When this taste shifter follows the extreme value distribution an expression of the choice probabilities can be analytically solved for. This leads to a slight modification of the contraction mapping (it now becomes a log sum instead of the sum above) in the calculation of choice probabilities of the model. The main argument for adding the taste shifter is to smooth choice probabilities.

Figure 1.4: Timing in the model



from the model which can be compared with data. Next the estimation procedure is presented, and a discussion of identification of central parameters of the model is provided. The model is estimated using the method of moments. Table 1.6 contains a summary of the chosen moments including the mean and standard deviation of the time series of moments. The moments capture employment, unemployment and wage dynamics which are informative about the structural parameters of the dynamic program.

Solution of the model and initial conditions

To generate predictions to compare with data the model is solved in a series of steps which will be briefly commented on below. More details are outlined in Appendix A. The solution procedure consists of 5 steps, similar to the steps presented in Ferrall (2002), and one additional step which will be presented in the last paragraph of this subsection:
i) Solve for $V(\theta)$ in (1.1).
ii) Calculate the policy function $P(\alpha|\theta)$ as given in (1.2).
iii) Use the transition function for state variables ($P(\theta'|\theta,\alpha)$) and the policy function (from ii) to solve for how the distribution over states evolve from one period to the next unconditional on choices ($P(\theta'|\theta)$).
iv) Use the state-to-state transition matrix (determined in iv) to solve for the ergodic distribution across states. [40]
v) From the ergodic distribution and the state-to-state transition matrix create a sample of unemployed workers which matches the data on observables (e.g. unemployment duration) and also takes account of the dynamic selection on unobservables.
Step v) takes into account that the data is not a random sample of workers from the ergodic distribution, but endogenously sampled as to enter the experiment individuals had to become unemployed - and some even had to remain unemployed for a longer period of time (see the Data section above). This makes the sample negatively selected in terms of both observables and unobservables compared to the average worker in the ergodic distribution. Naturally neither of the above invalidates the experimental design but it is important to take into account in an analysis focused on quantifying parameters in the decision process and extrapolating the results into other settings with different individuals or policies. If this initial conditions problem (see e.g. Aguirregabiria and Mira (2010)) is not accounted for, estimates of the decision parameters will be biased as unobservables correlate with observables and decisions in ways which are unaccounted

[40] To solve for the ergodic distribution solve for the fixed point (vector) in $\pi(\theta) = P(\theta'|\theta)\pi(\theta)$. Ferrall (2002) shows the conditions which are required for the existence of an ergodic distribution - the main requirement is that transitional dynamics for each state variable is either ergodic, invariant or dependent.

for.[41]
The five steps outlined above can be used to solve for $\Omega(\theta|t)$, the distribution over states at time t since inflow into the experiment for the selected group of individuals enrolled into the experiment. $\Omega(\theta|t)$ can be solved for by iterating on the initial distribution across states defined in v) using the state-to-state transition matrix defined in iv).[42]
Finally a further step vi) is added to the solution procedure. In order to increase the number of predictions from the model which can be compared to data, the fraction of individuals which have followed specific paths or spells and thier distribution across spells can be determined by extending the solution procedure outlined above.
vi) Determine the fraction of workers satisfying certain spell requirements and determine their distribution across state variables for each t
This final step is illustrated by example. To economize on state variables there is no state variable which counts employment duration in the model presented above. In order to calculate moments related to employment duration the inflow into employment for each t, and the distribution across states, is determined. This implies that the distribution of employment durations for each period t can be obtained. In practice an inner "reduced" Markov chain is added to the solution of the model which determines $\Omega^{RED}(\theta|t)$, the distribution across states satisfying certain spell requirements (here employment for a period of time).[43] These calculations thereby allow me to include moments which are only indirectly linked to a state variable.[44]In relation to the model specified above, these calculations allow the inclusion of moments describing both employment durations and wages conditional on employment durations to the set of moments which is matched on. Adding these moments has the advantage that the unobserved state variable human capital is now more directly linked to data moments which strengthens the identification of the job separation and skills accumulation processes.

Estimation

The parameters of the model are estimated using the method of moments. The estimation proceeds as follows: For a set of parameters the model generates a series of behavioural predictions which are translated into moments and compared to data. The difference between these moments is now minimized by changing the parameters, and resolving the model to generate new predic-

[41]For example individuals with long unemployment spells at inflow into the experiment, may accept low wages either because they have a low cost of working or low skills. If we do not take into account that this group is a negatively selected group in terms of skills we prescribe all behaviour to the former.

[42]$\Omega(\theta|t) = P(\theta'|\theta)^t \omega(\theta)$ where $\omega(\theta)$ denotes the initial distribution over states at the start of the experiment.

[43]$\Omega^{RED}(\theta|t+k) = P^{RED}(\theta'|\theta)^k \omega_t^{INFLOW}(\theta)$ where $\omega_t^{INFLOW}(\theta)$ denotes the inflow into employment in period t. $P^{RED}(\theta'|\theta)^k$ is a transition matrix which have non-zero entries for transitions which implies that the individual stays employed. $\Omega^{RED}(\theta|t+k)$ gives the fraction of individuals who has been employed for k periods at time t+k and the distribution across statevariables θ.

[44]The procedure is basically the method of moment equivalent to the *simulated* method of moment estimators where the simulated data enables the researcher to condition on moments not directly linked to the model. The inner chain calculates the distribution of e.g. employment duration over time although employment duration is not a state variable in the model. The modification shows how further moments can be added to the model without increasing the state space or having to simulate the model (but naturally it still affects computational speed). A recent paper by Eisenhauer et al. (2015) documents that the simulation error that exists in models exploiting simulated moments can affect the estimates in non-trivial ways.

tions, until a minimum is found. To calculate the predictions of the model start by calculating the expected value of a certain outcome (for instance the wage) conditional on a position in the state space θ.

$$E\left(Y|\theta\right)=\sum_{\alpha}P\left(\alpha|\theta\right)Y\left(\alpha,\theta\right)$$

Where $Y\left(\alpha,\theta\right)$ is a given outcome which may vary with both choices and position in the state space. $E\left(Y|\theta\right)$ gives us the expected value of a moment conditional on a position in the state space. The initial distribution over states and the Markovian structure of the problem determine how the conditional moment evolves over time. Next the conditional moments are weighted with the corresponding distribution over states at a given point in time. Predictions are determined conditional on time (t) and also conditional on the time invariant states: unobserved type (k), educational (e), regional (r) and treatment groups (g).[45] This results in a time series of moments:

$$E\left[Y_M|t,e,r,g,k\right]=\sum_{\theta_{|k,e,r,g}}\Omega(\theta_{|k,e,r,g}|t,k,e,r,g)E\left(Y|\theta\right)$$

Where to calculate moments related to employment duration (#12-#14 in Table 1.6) $\Omega\left(\theta|t,k,e,r,g\right)$ is substituted with $\Omega_{RED}(\theta|t,k,e,r,g)$ which was defined above. Next, moments are weighted with the distribution over unobserved types (k):

$$E\left[Y_M|t,e,r,g\right]=\sum_{k}\lambda\left(k,e,r\right)E\left[Y_M|t,e,r,g,k\right]$$

where $\lambda\left(k,r,e\right)$ is the proportion of type k individuals within educational group e and region r. Finally model predictions are compared to data predictions:

$$\left(E\left[Y_D|t,e,r,g\right]-E\left[Y_M|t,e,r,g\right]\right)'W\left(E\left[Y_D|t,e,r,g\right]-E\left[Y_M|t,e,r,g\right]\right)$$

The weight matrix is an inverted diagonal matrix with the variance of the data moments in the sample. The method of moments now proceeds by minimizing this objective. Standard errors are calculated for the standard one stage GMM case (see e.g. Cameron and Trivedi (2005)):

$$\mathrm{Var}\left(\Theta\right)=\frac{1}{N\cdot T}\left(\hat{G}'W\hat{G}\right)^{-1}\hat{G}'W\hat{S}W\hat{G}\left(\hat{G}'W\hat{G}\right)^{-1}$$

where Θ denotes the vector of parameters to be estimated, $\hat{G}$ is the Jacobian matrix and $\hat{S}$ the sample variance covariance matrix of the matrix of moments over time.

Identification

The parameters of the model are identified by restrictions (of both the behavioural and functional form) generated by the model on how the moments can vary over time, within and across educational and regional groups. The existence of experimental variation generates exogenous

[45] The notation $\theta_{|e}$ therefore refers to the set of state variables excluding the state variable e.

Table 1.6: Included time series of moments

			Meetings			Activation		
			Low	Medium	High	Low	Medium	High
#	Data Moment	Model Moment	Mean (S.D)	Mean (S.D)	Mean (S.D)	Mean (S.D)	Mean (S.D)	Mean (S.D)
1	Job separations	lj	0.026 *(0.08)*	0.027 *(0.09)*	0.022 *(0.09)*	0.026 *(0.09)*	0.027 *(0.10)*	0.025 *(0.11)*
2	Wages Squared	$W(\alpha,\theta)^2$	4322 *(986)*	6150 *(1355)*	7747 *(2156)*	3570 *(689)*	5255 *(1192)*	6833 *(1932)*
3	Wages	$W(\alpha,\theta)$	45.89 *(8.36)*	60.39 *(10.50)*	73.46 *(16.99)*	37.14 *(5.31)*	53.52 *(9.79)*	65.81 *(15.38)*
4	UE dur squared	cu^2	64.74 *(19.90)*	46.30 *(13.81)*	31.27 *(5.64)*	76.00 *(26.00)*	53.32 *(18.15)*	40.36 *(9.06)*
5	E-inflow* UE dur	$wc \cdot (1-\mathrm{e}) \cdot \mathrm{cue}^2$	1.68 *(0.86)*	1.56 *(0.76)*	2.09 *(1.75)*	1.35 *(1.11)*	1.50 *(0.62)*	1.73 *(0.76)*
6	U	$(1-\mathrm{em}) + wc \cdot \mathrm{lj}$	0.51 *(0.09)*	0.41 *(0.11)*	0.30 *(0.16)*	0.57 *(0.06)*	0.45 *(0.10)*	0.36 *(0.14)*
7	E	em	0.46 *(0.11)*	0.57 *(0.13)*	0.67 *(0.18)*	0.40 *(0.08)*	0.52 *(0.12)*	0.61 *(0.16)*
8	Inflow into E	$wc \cdot (1-\mathrm{em})$	0.02 *(0.03)*	0.02 *(0.03)*	0.02 *(0.03)*	0.02 *(0.03)*	0.02 *(0.03)*	0.02 *(0.02)*
9	Inflow wages (iw)	$wc \cdot (1-\mathrm{em}) \cdot W(\alpha,\theta)$	2.26 *(2.68)*	2.49 *(3.32)*	2.44 *(2.80)*	1.99 *(2.83)*	2.24 *(2.72)*	2.21 *(2.41)*
10	iw * UE dur	$wc \cdot (1-\mathrm{em}) \cdot W(\alpha,\theta) \cdot \mathrm{cu}$	15.58 *(8.51)*	15.70 *(9.99)*	18.92 *(16.88)*	12.09 *(9.43)*	14.85 *(8.13)*	16.34 *(8.74)*
11	Wages for E only	#3/#7	91.03 *(26.05)*	96.89 *(27.53)*	100.43 *(28.69)*	88.15 *(25.17)*	93.60 *(26.61)*	98.86 *(28.11)*
12	E-Dur	emdur	7.65 *(4.16)*	10.40 *(6.09)*	13.26 *(8.53)*	6.33 *(3.65)*	9.37 *(5.53)*	11.70 *(7.70)*
13	Wages * E-dur	$W(\alpha,\theta) \cdot \mathrm{emdur}$	722 *(390.32)*	1063 *(628)*	1435.93 *(955)*	584.27 *(331)*	918.82 *(551)*	1241 *(839)*
14	Wages * E-dur^2	$W(\alpha,\theta) \cdot \mathrm{emdur}^2$	20267 *(17377)*	31028 *(27807)*	43430 *(41712)*	15607 *(13559)*	26557 *(24281)*	36723 *(35725)*

Note: The table contains the subgroup specific means and standard deviations on the time series of moments exploited in estimation (time series is 80 weeks). Wages are after imputed taxes. UE: unemployment, E: employment, dur: duration, emdur: is not a state variable in the model but calculated for each period variation in programme participation which is useful in the identification of costs, ϕ_i, and the experiment also generates variation in incentives across time which can be useful to separate $\pi_{w,i}$ and ϕ_i. To assess whether the imposed structure and the selected moments are sufficient to recover the structural parameters a "baby-version" of the model, with the main central mechanisms, have been simulated and subsequently the generated data were used to confirm that the chosen parameter values could be recovered from estimation. Although this is by no means a formal proof of identification nor an actual Monte Carlo exercise it still provides a good indication of whether the model is identified.

Some central issues related to identification can be illustrated by discussing some "exclusion restrictions" in the model (see e.g. Wolpin (2013, 1987)). For instance, the model predicts that unemployed individuals are unemployed *either* because a job offer has been rejected or because no job offer was available. As the data contain no information on whether the unemployed received a job offer nor the size of the wage offered, the two explanations should be distinguished using data on job finding and accepted wages only. This distinction requires either functional form restrictions or the existence of a variable which affects the availability of job offers without directly affecting the decision to accept a job offer or not, and similarly a state variable in the wage function should affect wages without affecting preferences. Here the impact of unemployment duration on job offer rates, the impact of skills on wage offers serve as such restrictions. Furthermore the restrictions across regions, time[46] and educational groups, and the functional form of duration dependence implies that the probability of receiving a job offer and the parameters of the wage function can be separated from preferences. This also serve as motivation for why some of the moments matched on are unemployment duration squared and duration dependent outflow rates from unemployment. If two individuals are identical besides different unemployment duration and they have different inflow rates, this is thus informative about how the probability of a job offer changes with unemployment duration. Similarly information on the level and evolution of skills provides information regarding the parameters of the wage function and thus provides further justification for the importance of step vi) outlined in the solution procedure above. This step ensures that unobserved state variables can be linked to certain patterns in the data which improves identification of the parameters of the process. In particular, the (squared) interaction between employment duration and wages is informative of how the distribution of human capital evolves over time while matching on employment duration and the rate of job loss is informative of the parameters determining job loss.

However, the most important source of variation is generated by the existence of the experiment. The experiment generates exogenous variation in the cost of being unemployed which allows us to distinguish between competing cost structures. The experimental variation allows a separation of different environments such as i) an environment with large costs of programme participation and search intensity from ii) an environment with low costs. These two explanations could gen-

[46] Repeated spells of unemployment are also useful in this context since tastes are constant over time in the model. Magnac and Thesmar (2002) show how stationarity of the utility function serve as another exclusion restriction in dynamic models. In particular, as the utility of choices does not change with time itself observing individuals make choices at different points in time (and in different enviroments (e.g. after longer spells of employment or unemployment)) therefore also serves as identifying variation.

erate the same size of impact but will have very different welfare implications: in i) agents incur substantial utility costs as they have to participate in "harmful programmes" without us directly observing it in the data, whereas in ii) the agents utility changes very little. Non-experimental data on individuals participating in ALMPs will not allow us to distinguish these explanations without assumptions that allow us to evaluate what individuals would have done in the absence of treatment. Identifying costs thus requires a comparison of participants with non-participants at a given point in time, and conditional on a limited set of state variables. This requires the assumption that agents are similar in all other dimensions than programme participation.

Furthermore the experimental variation in incentives generates an additional channel of variation which is useful. One way to distinguish between $\underline{\phi_i}$ and $\pi_{w,i}$ (two different explanations for why programmes work) is by looking at the time profile of individual behaviour. An increase in the inflow to employment in the weeks prior to programme participation or within the first weeks is informative of the size of the utility cost component, whereas outflow rates after programme participation informs us about the qualification effect.

Overall the experiment generates the opportunity to observe identical agents in different settings and from their differential behaviour and, by the imposed structure of the model, to analyse the way that the treatment affects individuals. Contrasting moments from the treatment group with the control group allow us to keep other time varying confounders such as duration dependence and differences in skills fixed and attribute differences in behaviour to the exogenous difference in programme participation. The model and the experimental data therefore allow us to assess the importance of utility costs in an environment which is very rich in terms of heterogeneity between participants: some will face different returns to searching (due to different cu), some will face different wage offers (hc or j differences) or different job stability (differences in hc), some will differ in preferences for leisure etc. This also affects the size and distribution of the compensating variation associated with the experiment or other policies as it introduces differences in the prospects (and thus in the value associated with alternative choices) for future employment across individuals.

1.7 Results

Below some evidence on the fit of the model are presented with a series of figures that compare model predictions with data. Initially focus is on the fit of the model in the control group. Subsequently some key implications of the estimated model are shown and the long term predictions of the model is compared to data. The section proceeds by comparing the impact of the experiment generated by the model with data, and then some key channels through which participation in ALMPs affect job finding are presented. Next the implications of the estimates and the importance of the individual level costs of participating in ALMPs is discussed. The section concludes with a welfare analysis of the experiments under investigation incorporating the costs associated with increased production and participation in ALMPs. To incorporate the later the compensating variation is calculated.

Table 1.7 presents the estimated parameters and associated standard errors. The table shows that wages are generally increasing in the education level of individuals. The return to job offers ($\underline{\sigma}^e$) is higher for high educated individuals which increases both the average wage offer and the dispersion in the wage offer distribution for high educated workers. There are also differences in the cost of working ($\underline{\kappa}^e$) across educational levels - while low educated workers get lower wages their cost associated with working is also lower. There are substantial costs associated with participation in ALMPs, and the gains in terms of an increase in job offers are very small and insignificant (below I discuss the implications in more detail). The estimates also show that the environment is composed of two different types with different search costs determined by ξ^{ψ_k}: T1 with linear costs have the highest costs of searching which amounts to around 25 % of his UI, while T2 incur lower costs ($\xi^{\underline{\psi}^{k=2}}$) of around 10 % of his UI when searching at the highest intensity.

Model fit

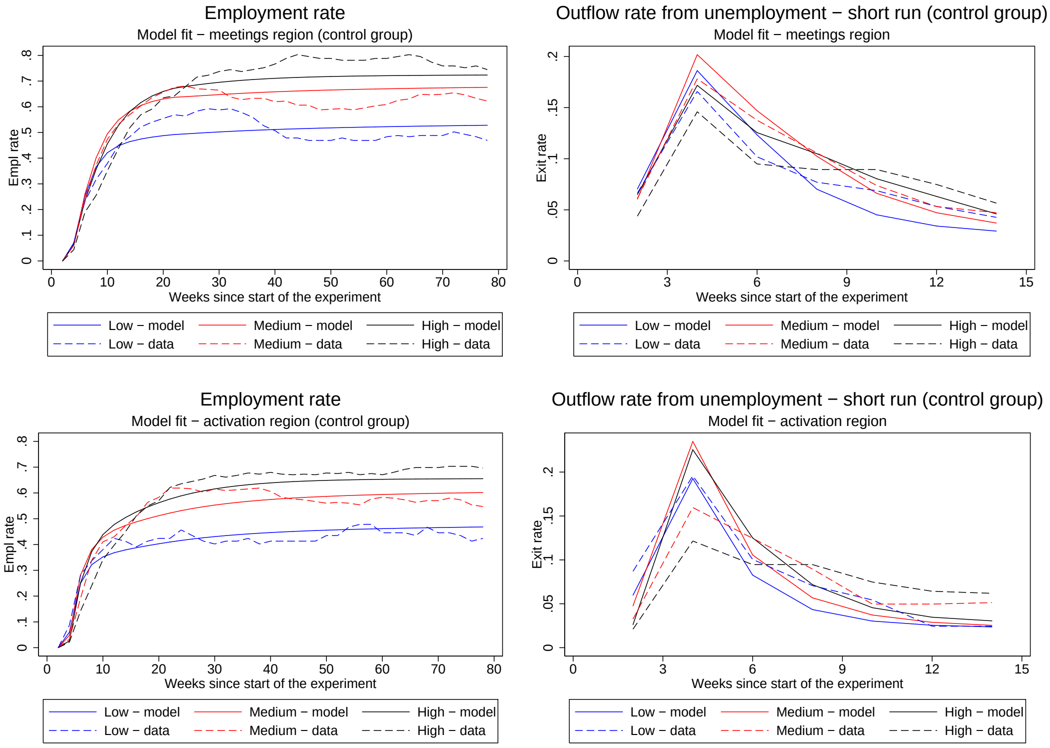
Below some evidence on the fit of the model for the control group is presented. For comparison Appendix C contains the same graphs for the treatment group, the difference between control and treatment predictions will be discussed in the subsection on the impact of the experiment. Figure 1.5 shows the correspondence between predictions of the model and the data for employment rates (see Appendix B Figure 1.19 for more figures on the inflow rate). Generally the fit is reasonable and the figure shows that the model is able to generate both high initial outflow and subsequently decreasing inflow rates. The model fits inflow rates in the long run suggesting that the main predictions in terms of in and outflow to employment are well explained by the model (I show evidence on job separations below). Figure 1.6 shows that the model matches the distribution of unemployment durations reasonably well although beyond 20 weeks the model predicts too high unemployment duration in the region with meetings (this is related to the fact that the model does not fit inflow rates sufficiently well in this region). Fitting unemployment duration is important as the model presents duration dependence as an important characteristic of the labour market. Figure 1.24 shows the fit of the model with the distribution of accepted squared wages and Figure 1.20 in Appendix B shows the fit of the interaction between duration in employment and wages. Generally the model fits both moments well, but wages for highly educated are too high initially. Finally Figure 1.8 shows the fit in terms of job loss. The figure shows that job loss in the data is generally very volatile due to a limited number of observations. The model predictions are smoother but seems to predict the average level of job loss in the data reasonably well. There is a clear educational ordering in terms of the fraction of individuals in the sample who are separated from a job. Since employment levels are also increasing in the education level, and more individuals are therefore at risk of losing their job, the difference in individual probabilities of a job loss across education groups must be substantial.

Table 1.7: Estimated parameters Θ:

Parameter	**Model**	**Note**	**Region**	**Low**	**Medium**	**High**
ξ	Utility	Search cost	Both	0.147 *(0.02)***	-	-
$\underline{\kappa}^e$	Utility	Work cost	Both	0.231*(0.011)***	0.358 *(0.014)***	0.375 *(0.017)***
$\underline{\phi}^e_{mp}$	Utility	Meetings cost	Both	0.241 *(0.099)***	0.192*(0.013)***	0.288 *(0.021)***
$\underline{\phi}^e_{ap}$	Utility	Activation cost	Both	0.123 *(0.019)***	0.246 *(0.010)***	0.287 *(0.015)***
$\underline{\psi}^{k=2}$	Utility	Leisure preference	Both	1.665 *(0.155)***	-	-
$\underline{\pi}^{k=2}_k$	Type Proportion	Fraction of type 2	Meetings	0.607 *(0.023)***	0.777*(0.034)***	0.791 *(0.057)***
-	-	-	Activation	0.584 *(0.035)***	0.730 *(0.051)***	0.718 *(0.048)***
μ	Wages	Wage constant	Both	3.645 *(0.040)***	-	-
$\underline{\sigma}^e$	Wages	Return to J	Both	0.237 *(0.018)***	0.366 *(0.110)***	0.475 *(0.100)***
η	Wages	Return to hc	Both	0.715 *(0.15)***	-	-
ρ	Smoothing	Smoothing kernel	Both	15.554 *(1.228)***	-	-
$\underline{\pi}^r_{w,1}$	Job-offers	Duration dependence	Meetings	-0.160 *(0.16)***	-	-
-	-	-	Activation	-0.376 *(0.09)***	-	-
$\underline{\pi}^e_{w,2}$	Job-offers	Long term job offer	Both	0.089 *(0.009)***	0.102 *(0.009)***	0.155 *(0.012)***
$\pi_{mp,3}$	Job-offers	Productive effect (meeting)	Both	0.002 *(0.012)*	-	-
$\pi_{ap,4}$	Job-offers	Productive effect (activation)	Both	0.014 *(0.029)*	-	-
$\underline{\pi}^e_{lj,1}$	Job-loss	Impact from hc	Both	0.102 *(0.008)***	-0.053 *(0.001)***	-0.071 *(0.001)***
$\underline{\pi}^r_{lj,1}$	Job-loss	Regional effect	Meetings	1	-	-
-	-	-	Activation	0.91358 *(0.069)***	-	-
$\underline{\pi}^e_{hc,1}$	Skills evolution	Appreciation of hc	Both	0.160 *(0.012)***	0.027 *(0.03)*	0.012 *(0.016)*
$\pi_{hc,2}$	Skills evolution	Loss of hc	Both	0.076 *(0.026)***	-	-

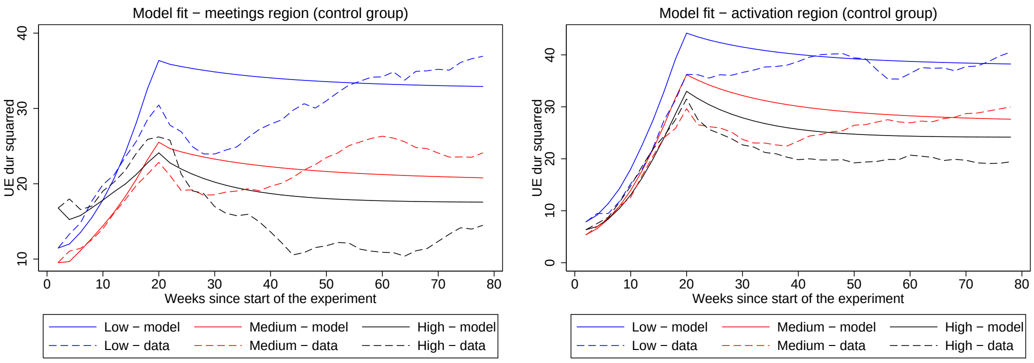
Note: Estimation window is 80 weeks. Standard Error in parenthesis, ** denotes significance at the 5 % level, * denotes significance at the 10 % level

Figure 1.5: Employment (data and model comparison)



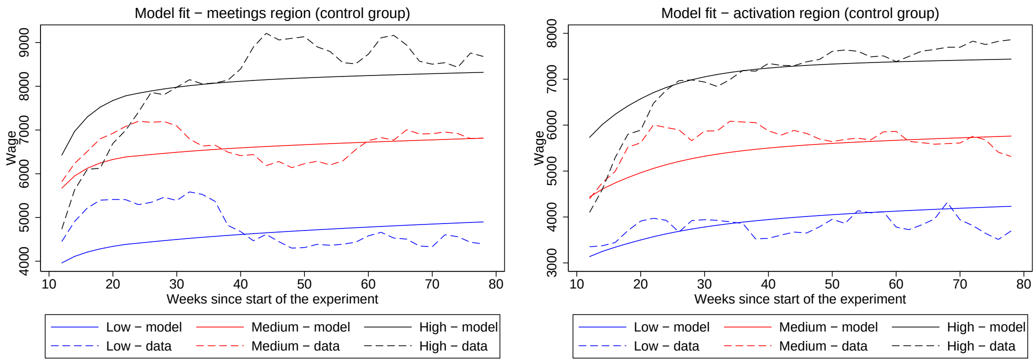
Note: see the appendix for further graphs on inflow rates for later weeks. See Appendix C Figure 1.22 for the same set of figures for the treatment group.

Figure 1.6: Average (squared) unemployment duration (data and model comparison)



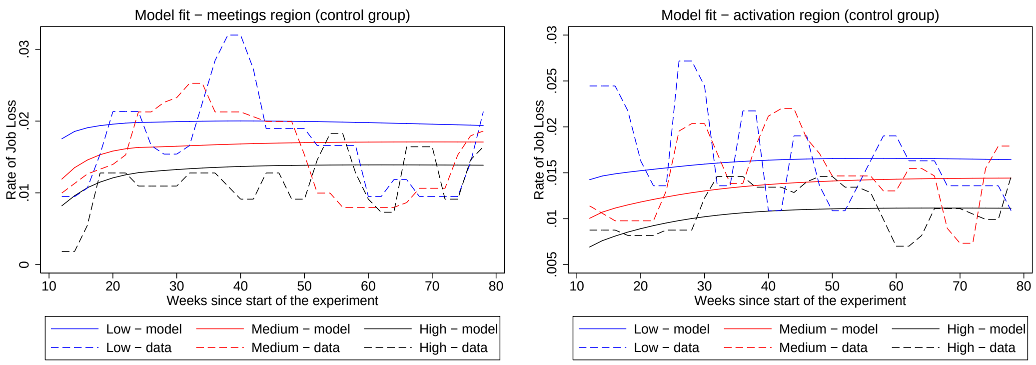
Note: See Appendix C Figure 1.23 for the same set of figures for the treatment group.

Figure 1.7: Squared wages (data and model comparison)



Note: See Appendix C Figure 1.24 for the same set of figures for the treatment group.

Figure 1.8: Job loss (data and model comparison)



Note: See Appendix C Figure 1.25 for the same set of figures for the treatment group.

Primitives of the model

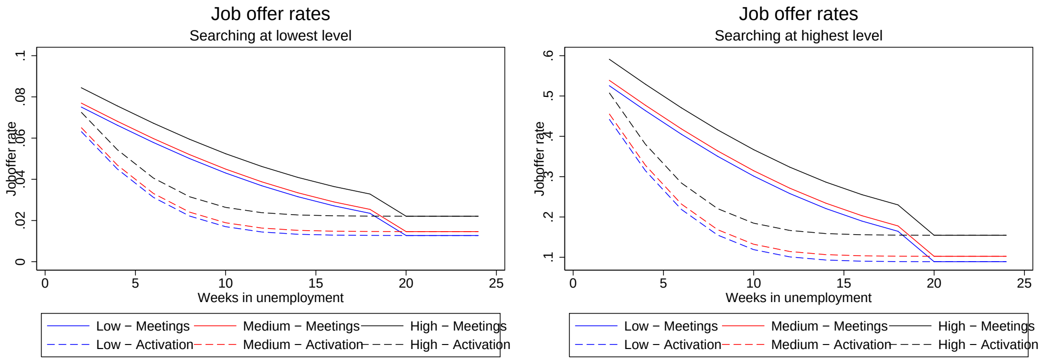
Figure 1.9 presents how job offer arrival rates varies with the duration in unemployment. The figure displays clear duration dependence in job offer rates - after 20 weeks of unemployment the likelihood that a job offer arrives is around 20 % of the rate at inflow into unemployment. Duration dependence implies that the return to job search substantially decreases over time and therefore individuals are likely to search more in the initial phases of unemployment. Negative duration dependence (see Figure 1.9) is an important part of the explanation for the decrease in outflow rates documented above. Figure 1.21 in Appendix B shows the fit of the model in an environment without duration dependence. This specification is not able to generate the spike initially in outflow rates and subsequently lower rates in the longer run. Either the initial outflow from unemployment is too low or alternatively the model predicts that long run employment levels will be too high. By including duration dependence in job offers the model fits outflow rates in both the short and longer run.

Figure 1.10 shows how wage offers vary as a function of the level of skills. There is a clear difference in wages across education levels and it is growing in the level of skills due to the education specific returns to skills. The level of skills also affects the probability of a job separation, Figure 1.11 presents the rate of job loss as a function of skills. Generally high skilled individuals face around 50 % of the risk of losing their job than low skilled. The risk of a job loss is declining in the level of education. Thus, although Table 1.7 shows that skill accumulation is faster for low educated workers, the higher risk of job separations also makes the risk of losing skills larger. As can be seen from the figures on model fit and model primitives, the environment therefore produces substantial returns to education - highly educated individuals receive higher wages, more stable employment and higher wage growth.

Table 1.8 presents some key moments in the ergodic distribution across education level, regions and types. As mentioned earlier, the environment consists of two very different types of individuals. In particular one type (with linear costs) have high costs of searching and therefore chooses not to search and instead just claim UI. This implies that they have high levels of unemployment duration and low skills. The table documents large differences in employment rates and skills both across and within educational levels. The heterogeneity across individuals is likely to generate heterogeneity in the effectiveness of ALMPs.

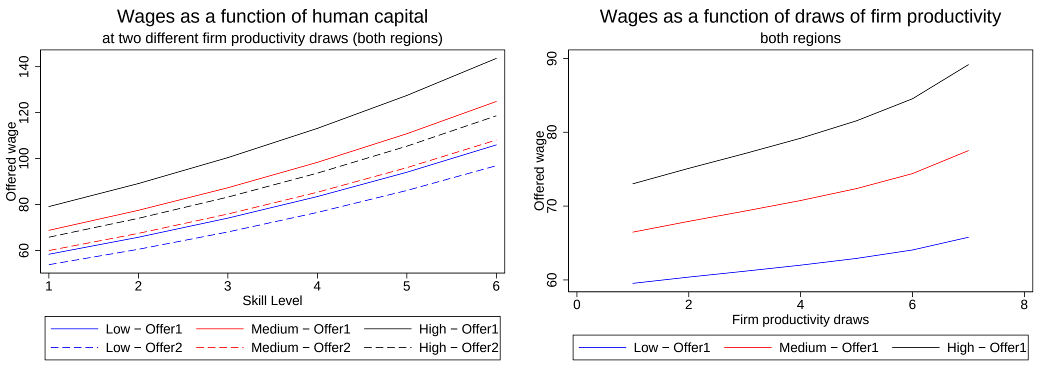
Finally the figures and table presented above show that there are some differences across regions (R1 and R2): job offer rates are generally lower in the region with early activation programmes (R2) and duration dependence is more pronounced. The probability of losing a job is also slightly smaller in this region. The environment is therefore less dynamic which is likely to affect the impact of the treatment and suggests that a "raw" comparison of the size of impacts across regions should be done with caution. This concern is further attenuated by the fact that the distribution of types differs slightly between the regions.

Figure 1.9: Job offer rates



Note: The figure gives the probability of receiving a job offer next period if searching at the lowest ($ac = \frac{1}{7}$, left), and the highest ($ac = 1$, right) level in the current period

Figure 1.10: Wage offer function



Left: The figure shows the offered wage across education levels as a function of skills (hc) for two different draws of j.
Right: The figure shows the offered wage for different draws of j across education levels.

Out of sample fit

Figure 1.12 illustrates how the model matches the data out of the current sample window. To generate predictions the time series of moments is solved beyond the estimation window of the first 80 weeks. The predictions are then compared to corresponding moments in the data. The figure shows how employment rates continue to decline beyond the estimation window in the data, and the model is not able to capture this decline fully. To the extent that this decline is not driven by changes in the environment after the experiment, this implies that the predictions of the ergodic distribution may be upward biased implying the future employment prospects of individuals are overstated.

Figure 1.11: Job Loss

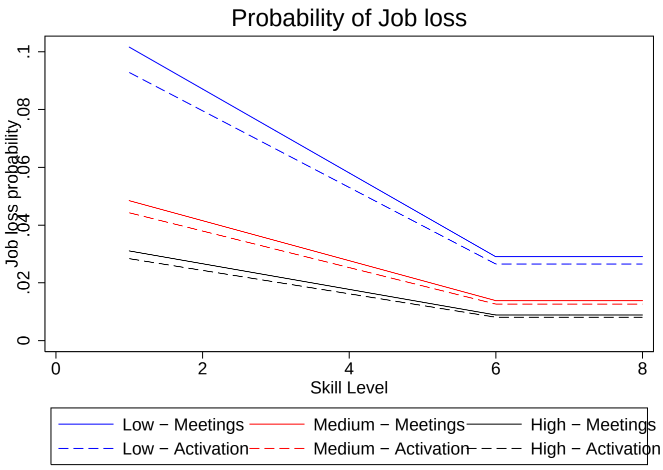


Table 1.8: Key moments describing the ergodic distribution

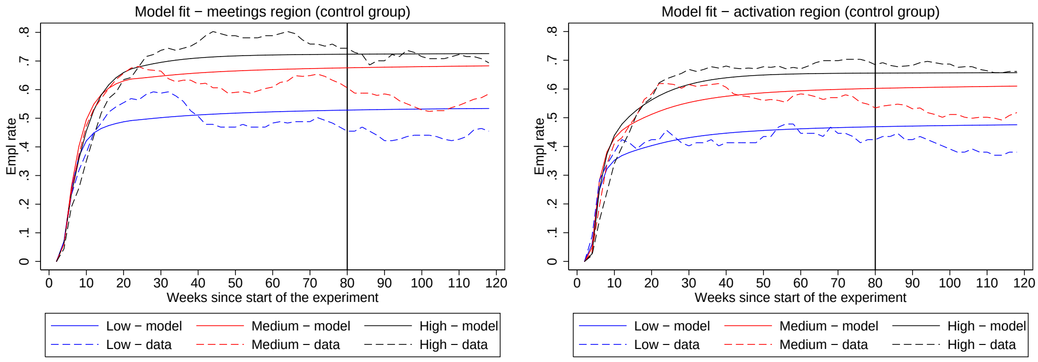
Region	Meetings					
Educational group	Low (R1)		Medium (R1)		High (R1)	
Preference for leisure group	T1	T2	T1	T2	T1	T2
Job loss rate	0.00	0.02	0.01	0.02	0.01	0.01
Employment rate	0.01	0.67	0.39	0.88	0.29	0.89
Unemployment duration***	8.99	2.19	5.22	0.56	6.19	0.56
Average skill level**	0.01	3.48	1.85	4.38	1.07	3.75
Region	Activation					
Educational group	Low (R2)		Medium (R2)		High (R2)	
Preference for leisure group	T1	T2	T1	T2	T1	T2
Job loss rate	0.00	0.02	0.01	0.01	0.01	0.01
Employment rate	0.01	0.45	0.20	0.85	0.25	0.87
Unemployment duration***	8.99	4.39	7.02	0.85	6.59	0.73
Average skill level**	0.01	2.29	0.93	4.18	0.85	3.60

Note: The table describes the ergodic distribution from which the inflow into unemployment is sampled (see estimation-section). R1: meetings region, R2: activation region,
* T1: high preference for leisure types ($\psi_k = 1$), T2: low preference for leisure types ($\psi_k \neq 1$), ** ce grid is $\{0, ...5\}$, *** cu grid is $\{0, ..., 9\}$

Impact of the experiment

Figure 1.13 compares the impact of the experimental intervention in the model to the data. The figure shows that there are clear regional differences in the response to being enrolled into the experiment in both the model and the data. In the model, individuals in the activation region display smaller responses as participation lies further in the future. In both regions the current fit of the model generates too small impacts, especially for low educated individuals. The impacts are largest in the region with meetings where employment increases with up to 15 % within the

Figure 1.12: Out sample predictions

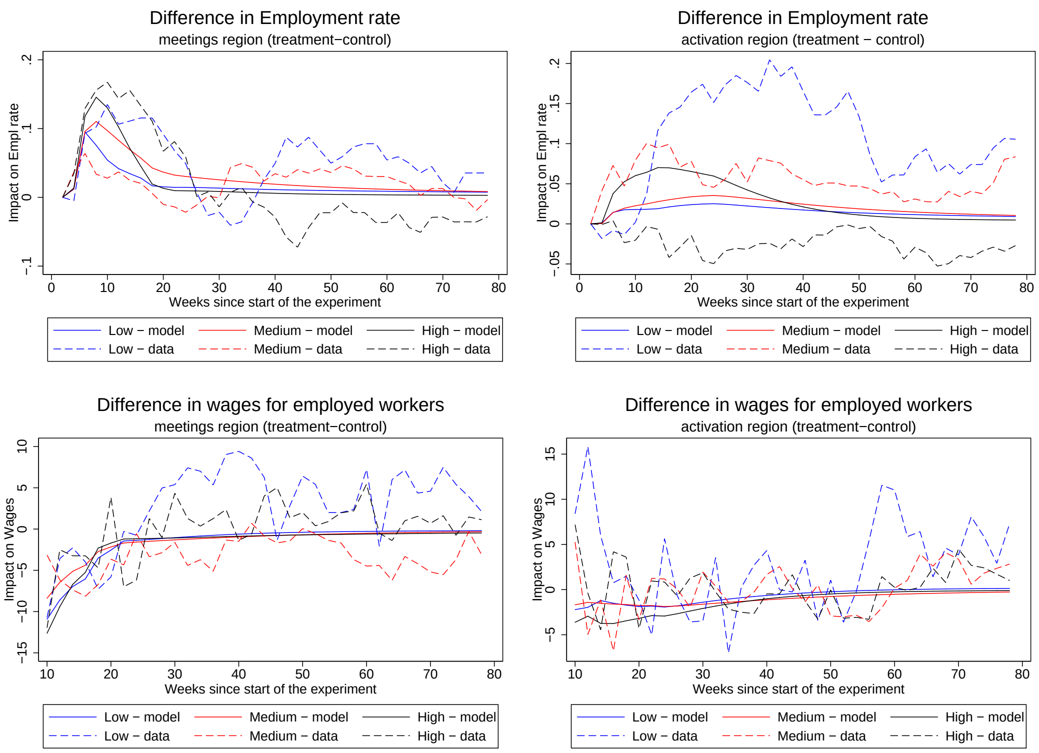


* The vertical bar (at 80 weeks) marks the end of the estimation window

first 5 weeks. The lower panel in the figure displays the impact on average wages for employed workers. Both in the data and in the model impacts are very small. In the meetings regions average wages are slightly lower initially both in the model and the data - over time the difference disappears.

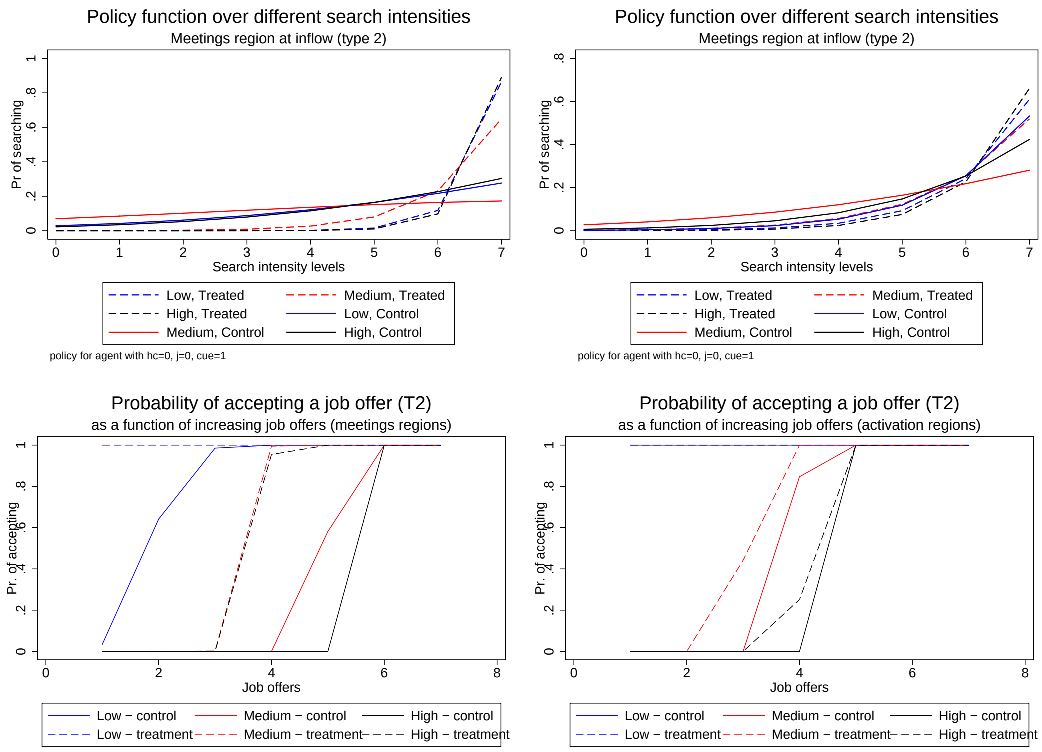
The impact of the experiment - or an increase in ALMP participation - is generated through several behavioural changes in the model. Figure 1.14 shows the policy function (defined in 1.2) over different levels of search activity and wage offers for members in the treatment and control groups at inflow into the experiment in the meetings region. Two things emerge from the figures: individuals in the treatment group change both their reservation wage and search behaviour in response to future participation in programmes which entails utility costs. The policy function over search intensities shows that treated individuals search more, and at higher levels. The likelihood of searching at the highest intensity is around 0.8 for high and low educated treated individuals compared to 0.2 in control group. The lower panel in Figure 1.14 shows that individuals in the treatment group also have lower reservation wages.

Figure 1.13: Impact of the experiment



Note: The upper panel gives the difference in employment rates for treatment and control groups in the model and data. The lower panel gives the difference in wages scaled with the fraction in employment

Figure 1.14: Channels of behaviour



* The figure shows the probability that agents choose either a given search level (upper panel) or to accept a given wage offer (lower panel). The figures show the policy function for individuals of type 2 (non-linear costs) in treatment and control groups. The response from Type 1 is generally smaller as effort is more costly and skills are lower.

Welfare calculations

Table 1.9 reports the monetary estimates of ϕ_{ap}, ϕ_{mp} and the equivalent reduction in UI. The estimates imply that individuals would be willing to reduce the level of UI with up to 50 % in the week of participation in order to escape participation.[47] The costs are increasing in the education level. The cost of participating in a meeting is higher than the corresponding one for activation when it is taken into account that a meeting is a much shorter intervention. Therefore unemployed would prefer an hour in activation compared to an hour in a meeting.

Table 1.9: Estimates of the cost of ALMP participation

Meetings	Low	Medium	High
$W_{max} \cdot \phi_{mp}$	32	30	45
Equivalent reduction in UI	35%	32%	49%
Activation	Low	Medium	High
$W_{max} \cdot \phi_{ap}$	25	45	43
Equivalent reduction in UI	27%	49%	48%

Note: The table reports the monetary value ($W_{max} \cdot \phi_i$) and the corresponding equivalent reduction in UI in the week of participation in ALMPs.

ϕ_{mp} and ϕ_{ap} measure the utility cost associated with *participating* in a programme in a given period. In order to calculate the total impact on treated individuals in the experiment under investigation further calculations are required. Generally two relevant calculations can be performed: i) the impact of the experiment on realized cost and ii) the effects on expected utility. i) only includes costs associated with actual programme participation in the treatment group, whereas ii) use the change in the value functions between treated and controls at inflow into the experiment as an estimate of the total effect on treated individuals. ii) therefore also includes the impact of the threat of future participation although such costs may never be realized because individuals change behaviour and increase their efforts to leave unemployment. The importance of this later channel depends crucially on the value of alternative actions; if for instance the cost associated with leaving unemployment earlier than "planned" is small the compensation needed to equalize the value functions is also smaller.

To calculate the monetary equivalent of the difference in expected utility the compensating variation (CV) is calculated for each value of θ and weighted with the distribution over states at inflow into the experiment. The CV is the monetary compensation which leaves individuals in the treatment groups indifferent between belonging to the treatment or control group at inflow into the experiment. It can be defined as:

$$V\left(\theta_{\text{enviroment}|g} + CV\left(\theta_{\text{enviroment}|g}\right) | \text{g=treatment, p=0, c=0}\right) = V\left(\theta_{\text{enviroment}|g} | \text{g=control}\right) \tag{1.13}$$

[47] For comparison Svarer (2011) reports that the size of sanctions related to failure to meet eligibility criteria (e.g. participating in a meeting) ranges from a loss of benefits for 2-3 days to 3 weeks (in severe cases benefits may be removed until new eligibility through employment has been established). The cost associated with not attending a meeting should be an upper bound on the size of utility costs (else participants should simply not attend).

where the LHS is equal to:

$$\max_{\alpha} \left\{ \log \left(Income\left(\alpha, \theta\right) - Cost(\alpha, \theta) + CV\left(\theta\right) \right) + \delta E\left[V\left(\theta'\right) \right] \right\}$$

In practice the state specific parameter (CV) is solved for in an optimization problem where the objective is the difference in value functions between treated and controls. Under a given set of monetary compensation levels the value function for individuals in the treatment group is calculated and compared it to the control group, This continues until the difference is 0.[48] Table 1.10 presents the average CV (within education, region and preference for leisure groups) associated with the experiment under investigation. The table shows that the average CV is substantial and up to 28 times larger than the monetary costs associated with participation in programmes in a given week. The compensation differs across educational groups and types where types with lower costs of searching require lower compensation. Furthermore the compensating variation is larger in the region with meetings as the probability of future participation is higher and closer in time (participation starts immediately after inflow).

Several features of the environment explain the high CV. Most importantly, as individuals are risk averse the utility function displays declining marginal utility of income which implies that the efficiency of compensation at inflow (before programme participation) is lower than giving it at the time of actual programme participation. This implies that the compensation exceeds expected future costs because the outcomes are risky and individuals prefer an environment without risks (the control group). Furthermore the average group specific CV masks substantial heterogeneity within groups. Table 1.18 in the appendix displays how the CV varies for different durations of unemployment (cu) and skills (hc) for low educated individuals in the meetings region. The table documents substantial heterogeneity in the CV, and two findings emerge. First the CV is increasing in unemployment duration and secondly it is declining in the level of skills or the value of future employment. A long term unemployed (cu=9) needs twice as high compensation as a newly unemployed. Similarly a high skilled individual only needs 50 % of the compensation given to a low skilled counterpart. The table thereby shows that the high average compensating variation is partly driven by individuals with low employment prospects who needs to be compensated much more.

Welfare analysis

The estimates of the CV can be used to analyse the overall impact on welfare associated with the experiment under investigation. The impact on welfare is analysed incorporating the value of lost non-market time both in terms of costs associated with participation in ALMPs and in terms of costs from an increase in production. In the CBA the gains to society of running the

[48] Since the utility function is non-linear solving for the CV implies that the contraction mapping should be resolved for each guess of compensation. Any compensation may change current actions and thus the expectations about the future. Furthermore, this calculation cannot be performed solving for the value functions in the control group due to the ergodic structure of the problem which implies that individuals eventually end up in the same state as their inflow state and thus receive the compensation again (individuals take this into account calculating the value associated with the given state). To ensure a one time compensation the compensation should therefore be calculated for the treatment group at inflow into the experiment.

Table 1.10: Compensating variation associated with the experiment

Meetings Region						
Education groups	Low		Medium		High	
Type	T1	T2	T1	T2	T1	T2
Compensating variation (CV)	912	634	794	434	941	314
CV relative to utility cost	28.5	19.8	26.5	14.5	20.9	6.9
Activation Region						
Education groups	Low		Medium		High	
Type	T1	T2	T1	T2	T1	T2
Compensating variation	227	67	444	181	796	204
CV relative to utility cost	9.1	2.7	9.9	4.0	18.5	4.7

Note: The CV (defined in 1.3) is weighted with the initial distribution across states at inflow into the experiment. T1: high preference for leisure types ($\psi_k = 1$), T2: low preference for leisure types ($\psi_k \neq 1$)

experiments are calculated. The gains include the value of increased production and in addition it is assumed that the marginal cost of public funds is either 20% or 0%[49]. The former means that to finance a given transfer to the unemployed the loss to society is 20%. When reducing transfers (by bringing individuals into employment) the gain to society amounts to 20% (0%) of the saved transfers. The saved transfers as such are not included in the CBA as this is simply a transfer internally in society. The costs are the direct costs of running the programme and in addition to the marginal costs of public funds needed to finance the extra costs.[50] The calculations are reported in Table 1.11. The table shows that a traditional CBA substantially overestimates the value of social programmes by assuming that the value of lost leisure is 0. This is especially true in cases where the programme requires some effort from the individual which he regards as unpleasant, as in such cases the non-market wage is substantially different from 0. In the case of meetings the gain of the programme falls by 50% and in the case of activation the gain is reduced with 80 % and is close to 0.

[49]There is a discussion in the literature (see e.g. Kreiner and Verdelin (2012)) on whether marginal costs of public funds should be included. Below calculations with 20 % and 0 % are therefore presented

[50]See Maibom et al. (2015) for further details. The same employment definition as reported in their paper is used for determining the increase in production. Maibom et al. (2015) analyse the impact of the experiment on government budgets and a simple welfare analysis (assuming that individuals do not value leisure) is conducted.

Table 1.11: Cost Benefit Analysis

in EURO pr. participant	Input	Costs MCPF=20%	Costs MCPF=0
Meetings: ***			
Saved income transfers	3631	726	0
Saved programme costs	-47	-57	-47
Saved total costs		669	-47
Acc. gain in employment (weeks)	7.44		
Gain of increased production		6508	6508
Costs from increase in production*		-1130	-1130
Value of increased production		5378	5378
CBA before welfare effects (in €)		6047	5331
Loss in welfare**		3007	3007
Net result of CBA (in €)		**3040**	**2291**

*Costs associated with the increase in production are the average value of $\underline{\kappa}$ averaged over types and education **I use the average cost (compensating variation) obtained by averaging over types and educational groups. *** The time frame is 237 weeks as in Maibom et al. (2015)

Table 1.12: Cost Benefit Analysis

in EURO pr. participant	Input	Costs MCPF=20%	Costs MCPF=0
Activation:			
Saved income transfers	1392	278	0
Saved programme costs	-440	-528	-440
Saved total costs		-250	-440
Acc. gain in employment (weeks)	2.98		
Gain from increased production		2607	2607
Costs from increase in production*		484	484
Value of increased production		2123	2123
CBA before welfare effects (in €)		1873	1683
Loss in welfare**		1482	1482
Net result of CBA (in €)		**391**	**201**

*Costs associated with the increase in production is the average value of $\underline{\kappa}$ averaged over types and education (work week 37 hours) **I use the average cost (compensating variation) obtained by averaging over types and educational groups. *** The time frame is 237 weeks as in Maibom et al. (2015)

1.8 Conclusion

Active Labour Market Programs (ALMPs) such as meetings at the job centre or shorter workfare (activation) programmes have been presented as a way to improve efficiency and reduce moral hazard in the labour market. The empirical literature has documented the existence of so-called threat effects which are consistent with the existence of a costs associated with programme participation. These costs arise because individuals spend a part of their non-market time at the job centre where they have to exert effort and potentially do unpleasant work (and maybe even feel stigmatized). Although costs are an important driver behind generated impacts, the previous literature has mainly focused on the gains of these programmes in terms of increasing job finding rates. However, in order to assess whether such programmes are indeed worthwhile social investments and whether better alternatives exist, gains associated with the programmes should be contrasted to costs including individual level costs. Therefore knowledge about the magnitude and distribution of such costs is needed.

Determining individual level costs is challenged by the fact they are generally unobservable. Furthermore since these programmes often serve as conditionalities for receiving UI the individual valuation is not directly observable from the individual decision of whether to enter the programme or not. Costs therefore have to be determined indirectly from individual behaviour such as job finding rates and accepted wages. In order to generate a link between behaviour and individual level costs an economic model of behaviour and an accurate description of the incentives faced by potential participants must be specified.

In order to quantify how individuals value ALMPs this paper developed a dynamic model with discrete choices capturing key behavioural channels which can be affected through interactions between unemployed and public authorities (the job centre). The model was estimated using data from a Danish randomized experiment which provides exogenous variation in the intensity of interactions. Thereby the costs agents incurs when they have to go into either activation or a meeting at the job centre can be estimated. To analyse effects on welfare, the structure of the model is used to calculate the compensating variation associated with the experimental intervention. The model incorporates several sources of heterogeneity and the analysis shows that the corresponding estimates of the compensating variation varies greatly among states. In particular some individuals require very large compensations at inflow into the experiment.

Overall the results suggest that traditional CBA calculations which do not take the individual loss of non-market time into account overstate the gain from having these programmes. The individual level costs are substantial and amounts to up to 50 % of UI in a given week of participation. The analysis shows that individual costs and associated compensating variation are important to quantify in order to assess whether the current mix between ALMPs and UI is optimal. Ignoring the existence of these costs implies that we put excessive weight on the efficiency of UI systems while overall welfare may be deteriorated.

Adda, J., Costa Dias, M., Meghir, C., Sianesi, B., 2007. Labour market programmes and labour market outcomes: a study of the swedish active labour market interventions. Tech. rep., Working Paper, IFAU-Institute for Labour Market Policy Evaluation.

Aguirregabiria, V., Mira, P., 2010. Dynamic discrete choice structural models: A survey. Journal of Econometrics 156 (1), 38 – 67, structural Models of Optimization Behavior in Labor, Aging, and Health.

Albrecht, J., van den Berg, G. J., Vroman, S., 2009a. The aggregate labor market effects of the swedish knowledge lift program. Review of Economic Dynamics 12 (1), 129 – 146.

Andersen, T., Svarer, M., 2007. Flexicurity - labour market performance in denmark. CESifo Working Paper Series, CESifo Group Munich 2108, CESifo Group Munich.

Andersen, T. M., Svarer, M., 2014. The role of workfare in striking a balance between incentives and insurance in the labour market. Economica 81 (321), 86–116.

Attanasio, O. P., Meghir, C., Santiago, A., 2012. Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA. Review of Economic Studies 79 (1), 37–66.

Beaudry, P., Blackorby, C., Szalay, D., 2009. Taxes and employment subsidies in optimal redistribution programs. American Economic Review 99 (1), 216–42.

Belzil, C., 1995. Unemployment duration stigma and re-employment earnings. The Canadian Journal of Economics / Revue canadienne d'Economique 28 (3), pp. 568–585.

Besley, T., Coate, S., March 1992. Workfare versus Welfare Incentive Arguments for Work Requirements in Poverty-Alleviation Programs. American Economic Review 82 (1), 249–61.

Bjoern, N. H., Hoej, A. K., 2014. UnderstÃžttelse ved ledighed i syv lande. Tech. rep., Danish Economic Council.

Black, D. A., Smith, J. A., Berger, M. C., Noel, B. J., 2003a. Is the threat of reemployment services more effective than the services themselves? evidence from random assignment in the ui system. American Economic Review 93 (4), 1313–1327.

Board, A., 2014. Veje til job - en arbejdsmarkedsindsats med mening. Tech. rep., Ministry of Employment.

Cameron, A. C., Trivedi, P. K., 2005. Microeconometrics: methods and applications. Cambridge university press.

Card, D., Kluve, J., Weber, A., 2010. Active labour market policy evaluations: A meta-analysis. The Economic Journal 120 (November), F452–F477.

Chetty, R., Finkelstein, A., 2013. Chapter 3 - social insurance: Connecting theory to data. In: Alan J. Auerbach, Raj Chetty, M. F., Saez, E. (Eds.), handbook of public economics, vol. 5. Vol. 5 of Handbook of Public Economics. Elsevier, pp. 111 – 193.

Cohen-Goldner, S., Eckstein, Z., May 2010. Estimating the return to training and occupational experience: The case of female immigrants. Journal of Econometrics 156 (1), 86–105.

Commission, E., 2007. Towards common principles of flexicurity: More and better jobs through flexibility and security. Tech. rep., Directorate-General for Employment, Social Affairs and Equal Opportunities, Brussels.

Cuff, K., 2000. Optimality of workfare with heterogeneous preferences. Canadian Journal of Economics 33 (1), 149–174.

Eckstein, Z., Wolpin, K. I., August 1999. Estimating the effect of racial discrimination on first job wage offers. The Review of Economics and Statistics 81 (3), 384–392.

Eisenhauer, P., Heckman, J. J., Mosso, S., 2015. Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. International Economic Review 56 (2), 331–357.

Ferrall, C., 2002. Estimation and inference in social experiments. Queen's University, Institute for Economic Research.

Ferrall, C., 2012. Explaining and Forecasting Results of the Self-sufficiency Project. Review of Economic Studies 79 (4), 1495–1526.

Fredriksson, P., Holmlund, B., 2006. Improving incentives in unemployment insurance: A review of recent research. Journal of Economic Surveys 20 (3), 357–386.

Gautier, P. A., Muller, P., van der Klaauw, B., Rosholm, M., Svarer, M., 2012. Estimating equilibrium effects of job search assistance, IZA Discussion Paper No. 6748.

Graversen, B., van Ours, J. C., 2008a. Activating unemployed workers work: Experimental evidence from denmark. Economics Letters 100, 308–310.

Greenberg, D. H., Robins, P. K., April 2008. Incorporating nonmarket time into benefit-cost analyses of social programs: An application to the self-sufficiency project. Journal of Public Economics 92 (3-4), 766–794.

Hagglund, P., 2011. Are there pre-programme effects of active placement efforts? evidence from a social experiment. Economics Letters 112 (1), 91 – 93.

Heckman, J. J., Lalonde, R., Smith, J., 1999. The Economics and Econometrics of ALMP. Vol. 3 of Handbook of Labor Economics. North-Holland, Amsterdam.

Judd, K. L., June 1998. Numerical Methods in Economics. Vol. 1 of MIT Press Books. The MIT Press.

Kluve, J., 2010. The effectiveness of european active labor market programs. Labour Economics 17, 904–918.

Kreiner, C. T., Tranaes, T., 2005. Optimal workfare with voluntary and involuntary unemployment*. Scandinavian Journal of Economics 107 (3), 459–474.

Kreiner, C. T., Verdelin, N., 2012. Optimal provision of public goods: A synthesis*. The Scandinavian Journal of Economics 114 (2), 384–408.

Kroft, K., Lange, F., Notowidigdo, M. J., 2013. Duration dependence and labor market conditions: Evidence from a field experiment*. The Quarterly Journal of Economics.

Lamadon, T., Lise, J., Seitz, S., Smith, J., February 2004. Equilibrium policy experiments and the evaluation of social programs. Working Paper 10283, National Bureau of Economic Research.

Lentz, R., 2009. Optimal unemployment insurance in an estimated job search model with savings. Review of Economic Dynamics 12 (1), 37 – 57.

Magnac, T., Thesmar, D., 2002. Identifying dynamic discrete decision processes. Econometrica 70 (2), 801–816.

Maibom, J., Svarer, M., Rosholm, M., 2014. Can active labour market policies combat youth unemployment? Nordic Economic Policy Review.

Maibom, J., Svarer, M., Rosholm, M., 2015. Experimental evidence of the effects of early meetings and activation. IZA discussion paper DP.

Moffitt, R., 1983. An economic model of welfare stigma. American Economic Review 73 (5), 1023–1035.

Mortensen, D. T., Pissarides, C. A., 1999. New developments in models of search in the labor market. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics. Vol. 3 of Handbook of Labor Economics. Elsevier, Ch. 39, pp. 2567–2627.

OECD, 2009. Oecd employment outlook - tackling the jobs crisis. Tech. rep., OECD.

Roed, K., 2012. Active social insurance. IZA Journal of Labor Policy 1 (1), 8.

Rust, J., January 1986. Structural estimation of markov decision processes. In: Engle, R. F., McFadden, D. (Eds.), Handbook of Econometrics. Vol. 4 of Handbook of Econometrics. Elsevier, Ch. 51, pp. 3081–3143.

Rust, J., September 1987. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. Econometrica 55 (5), 999–1033.

Shimer, R., Werning, I., December 2008. Liquidity and Insurance for the Unemployed. American Economic Review 98 (5), 1922–42.

Svarer, M., 2011. The effect of sanctions on exit from unemployment: Evidence from denmark. Economica 78 (312), 751–778.

Van Den Berg, G. J., Van Der Klaauw, B., 2006. Counseling and monitoring of unemployed workers: Theory and evidence from a controlled social experiment. International Economic Review 47 (3), 895–936.

Wolpin, K. I., 1987. Estimating a structural search model: The transition from school to work. Econometrica 55 (4), pp. 801–817.

Wolpin, K. I., 2013. The Limits of Inference without Theory. Tjalling C. Koopmans Memorial Lectures. The MIT Press.

Wolpin, K. I., Todd, P. E., December 2006. Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility. American Economic Review 96 (5), 1384–1417.

1.9 Appendix A: Solution of the model

I solve the model in a series of steps outlined below. The solution procedure is similar to the one presented in Ferrall (2002).

i) Solve for $V(\theta)$ *in* (1.1) *using the contraction mapping properties.*

The method of successive approximations and error bounds suggested by McQuad and Porteus is used (see Rust (1986)). Ferrall (2002) gives the conditions under which $V(\theta)$ is a contraction mapping. This have also been tested numerically by starting the fixed point equation from a series of different initial conditions, the resulting value function is indistinguishable across iterations.

ii) Calculate the policy function, $P(\alpha|\theta)$ *as given in* (1.2).

The policy function specifies how agents behave given a position in the state space. Given a distribution across state variables, aggregate behaviour in a given period can be determined (in a later stage this is then compared to data through moments). To determine behaviour across time the policy function and the transition functions presented in the main text can be combined to specify how the distribution over the state space evolves over time.

iii) Solve for the state-to-state transition matrix:

$$P_{sts}(\theta'|\theta) = \sum_{\alpha} P(\alpha|\theta)\, P(\theta'|\alpha,\theta)$$

The state-to-state transition function allows us to track the evolution of the state space from some t to some $t+k$ exploiting that the model is Markovian (i.e. iterating on a Markov chain). Given an initial distribution over states the distribution of states at a given point in time can be solved for. The remaining challenge is therefore to specify an initial distribution across states. This is further complicated by the fact that some state variables are unobserved and therefore an initial distribution over states is also unobservable. As explained earlier this problem is solved exploiting the existence of an ergodic distribution.

iv) Solve for the ergodic distribution:

$$P_{ergodic}(\theta) = \sum_{\theta'} P(\theta'|\theta)\, P_{ergodic}(\theta) \tag{1.4}$$

The ergodic distribution specifies how individuals are distributed across states in the economy in steady state. From this distribution the inflow into unemployment can be determined. The ergodic distribution is found by solving for the fixed point in(1.4), see also Judd (1998). The existence and uniqueness of the ergodic distribution have also been tested numerically.

v) Apply sample selection rules to the the ergodic distribution.

This final step creates a sample that matches the data on observable terms (e.g. unemployment duration) but also takes account of the dynamic selection on unobservables since inflow into unemployment. Using the state-to-state transition function and the corrected initial distribution over observable and unobservable states we can now solve for the distribution over states for each time period since t=0.

These 5 steps now enable me to relate the predictions of the model to the actual data and thus learn about the structural parameters. In the main text the content of a final step vi) is presented:

vi) Determine the fraction of workers satisfying certain spell requirements and determine their distribution across state variables for each t.

1.10 Appendix B: Further Figures and Tables

Further data descriptives

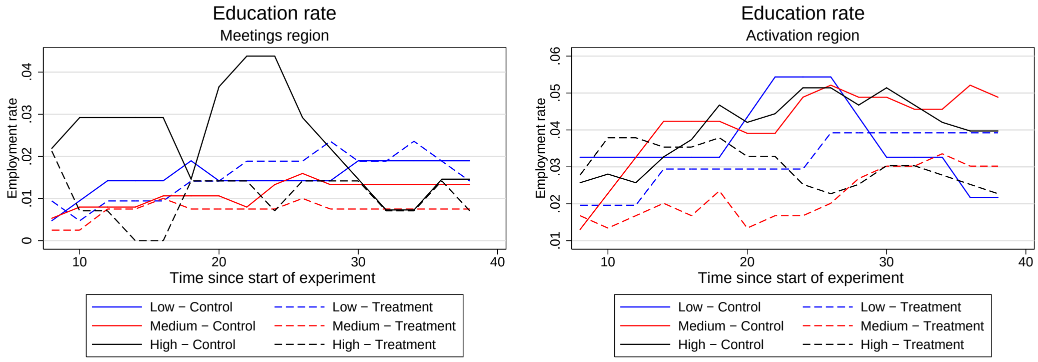
Table 1.13: Descriptives in the Meetings Region

Variable Type	Variable	Treatment	Control	P-value
Demographics	Age	39.18	39.12	0.91
	Below 30	0.21	0.23	0.18
	Above 45	0.27	0.29	0.39
	Fraction Males	0.52	0.55	0.25
	Education length	11.73	11.68	0.79
State before inflow	Newly Non-employed	0.80	0.77	0.18
	Sick-listed	0.12	0.12	0.89
	Education	0.04	0.03	0.30
	Earnings 2007	283890	273857	0.22
	Hourly Wage 2007	192.28	187.50	0.35
Previous Employment	Public sector	0.33	0.33	0.97
	Trade	0.55	0.56	0.90
	Construction	0.12	0.11	0.88
Employment history	Weeks in E (year -1)	40.35	39.70	0.48
	Weeks in E (year -2)	37.82	39.05	0.20
	Weeks in E (year -3)	36.41	36.88	0.64
	Weeks in E (year -4)	35.57	36.46	0.39
	Weeks in E (year -5)	36.30	36.65	0.74
	Weeks in NE (year -1)	11.65	112.30	0.48
	Weeks in NE (year -2)	14.18	12.95	0.20
	Weeks in NE (year -3)	15.59	15.12	0.64
	Weeks in NE (year -4)	16.43	15.54	0.39
	Weeks in NE (year -5)	15.69	15.35	0.74
Observations		752	724	

Table 1.14: Descriptives in the Activation region

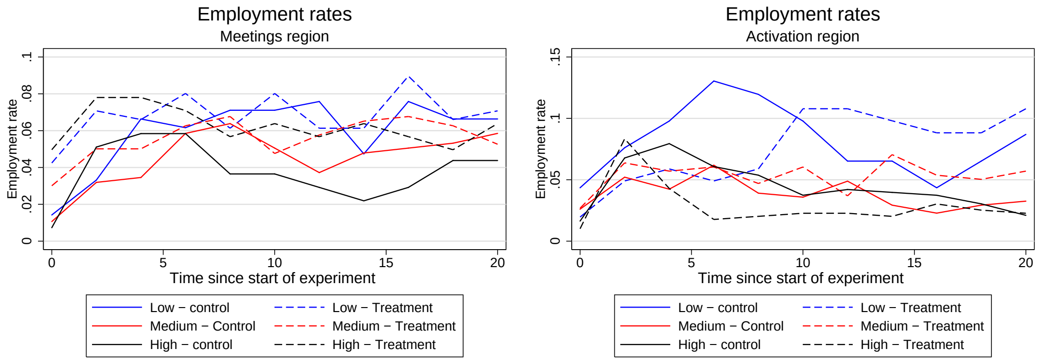
Variable Type	Variable	Treatment	Control	P-value
Demographics	Age	35.64	35.34	0.53
	Below 30	0.38	0.35	0.23
	Above 45	0.20	0.17	0.17
	Fraction Males	0.47	0.48	0.57
	Education length	13.63	13.78	0.26
State before inflow	Newly Non-employed	0.68	0.68	0.83
	Sick-listed	0.06	0.07	0.77
	Education	0.21	0.21	0.74
	Earnings 2007	234322.17	232753.77	0.87
	Hourly Wage 2007	158.30	151.50	0.28
Previous Employment	Public sector	0.42	0.47	0.03
	Trade	0.51	0.45	0.03
	Construction	0.07	0.8	0.87
Employment history	Weeks in E (year -1)	28.70	29.58	0.44
	Weeks in E (year -2)	26.69	26.13	0.63
	Weeks in E (year -3)	25.02	23.49	0.18
	Weeks in E (year -4)	23.10	21.53	0.16
	Weeks in E (year -5)	24.63	23.21	0.24
	Weeks in NE (year -1)	23.30	22.41	0.44
	Weeks in NE (year -2)	25.31	25.87	0.63
	Weeks in NE (year -3)	26.98	28.51	0.18
	Weeks in NE (year -4)	28.91	30.47	0.16
	Weeks in NE (year -5)	27.37	28.68	0.24
Observations		796	827	

Figure 1.15: Fraction of individuals entering education



Note: time since start of experiment is measured in weeks

Figure 1.16: Fraction in self-support



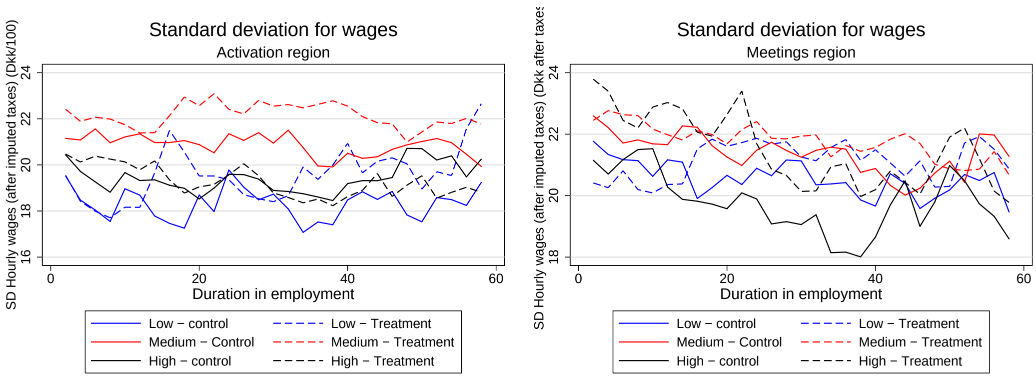
Note: time since start of experiment is measured in weeks

Table 1.15: Replicating Table 1.3 using an alternative employment criterion

Employment status	(1)	(2)	(3)	(4)	(5)	(6)
After 4 weeks	Low, R1	Medium, R1	High, R1	Low, R2	Medium, R2	High, R2
Treatment indicator	0.103*	0.0291	0.1000+	0.0234	0.0349	0.00911
	(0.0445)	(0.0340)	(0.0535)	(0.0643)	(0.0367)	(0.0268)
Constant	0.251*	0.322*	0.226*	0.261*	0.267*	0.175*
	(0.0299)	(0.0241)	(0.0359)	(0.0460)	(0.0253)	(0.0184)
After 14 weeks						
Treatment indicator	0.116*	0.0178	0.0911	0.134+	0.0525	-0.0352
	(0.0482)	(0.0356)	(0.0591)	(0.0702)	(0.0405)	(0.0349)
Constant	0.488*	0.561*	0.540*	0.337*	0.515*	0.502*
	(0.0345)	(0.0256)	(0.0427)	(0.0495)	(0.0286)	(0.0242)
Observations	423	775	278	194	605	824

Note: The results are from separate OLS regressions after 2, 4, 10 and 14 weeks. The dependent variable is employment status (not counting individuals in self sufficiency). Huber/White standard errors, + $p < 0.10$, * $p < 0.05$

Figure 1.17: Standard deviation for wages



Note: time since start of experiment is measured in weeks

Table 1.16: Wage growth

	(1)	(2)	(3)	(4)
Employment duration	20 weeks	40 weeks	60 weeks	80 weeks
Low Education	-0.00284	0.0237*	0.0180	0.0214+
	(0.00963)	(0.0109)	(0.0119)	(0.0130)
Medium Education	-0.00646	0.0215*	0.0188*	0.0371*
	(0.00600)	(0.00670)	(0.00722)	(0.00772)
High Education	0.0226*	0.0470*	0.0648*	0.0830*
	(0.00645)	(0.00714)	(0.00754)	(0.00801)
N	2694	2490	2303	2085

Robust standard errors in parentheses, + $p < 0.10$, * $p < 0.05$. Data is pooled across regions. Wages for employed workers are compared to their own inflow wage after 20,40,60,80 weeks

Figure 1.18: Treatment impact for employed workers

20 weeks in employment	(1)	(2)	(3)	(4)	(5)	(6)
	Low, R1	Medium, R1	High, R1	Low, R2	Medium, R2	High, R2
Treatment indicator	2.397	-3.778*	1.830	2.426	-0.255	0.103
	(2.419)	(1.700)	(2.450)	(3.285)	(1.945)	(1.373)
Constant	95.60*	101.2*	103.4*	93.24*	97.84*	101.4*
	(1.703)	(1.209)	(1.691)	(2.478)	(1.320)	(0.957)
N	307	646	256	130	501	754

Robust standard errors in parentheses, + $p < 0.10$, * $p < 0.05$. The sample is employed after 10 weeks of employment

Further details about the model and fit:

Table 1.17: Other parameters in the model (not estimated)

Symbol	Model	Value (Control group)
π_{mp}	Meetings probability	$\pi_{mp} = 0.10$
π_{ap}	Meetings probability	$\pi_{ap} = \min\{0.1 \cdot \text{uedur}, 0.35\}$
δ	Discount rate	0.995
UI	UI level	$0.6 \cdot W_{max}$

* in the treatment phase π_{mp} and π_{ap} are set to 1. ** parameters are set to match features of the data, in particular meetings and activation intensities as documented in Maibom et al. (2015). The replacement level of a worker earning 150% above average earnings is around 0.6, see Bjoern and Hoej (2014). Therefore, in the model UI is set to $0.6 \cdot W_{max}$.

Figure 1.19: Model fit: Inflow rates

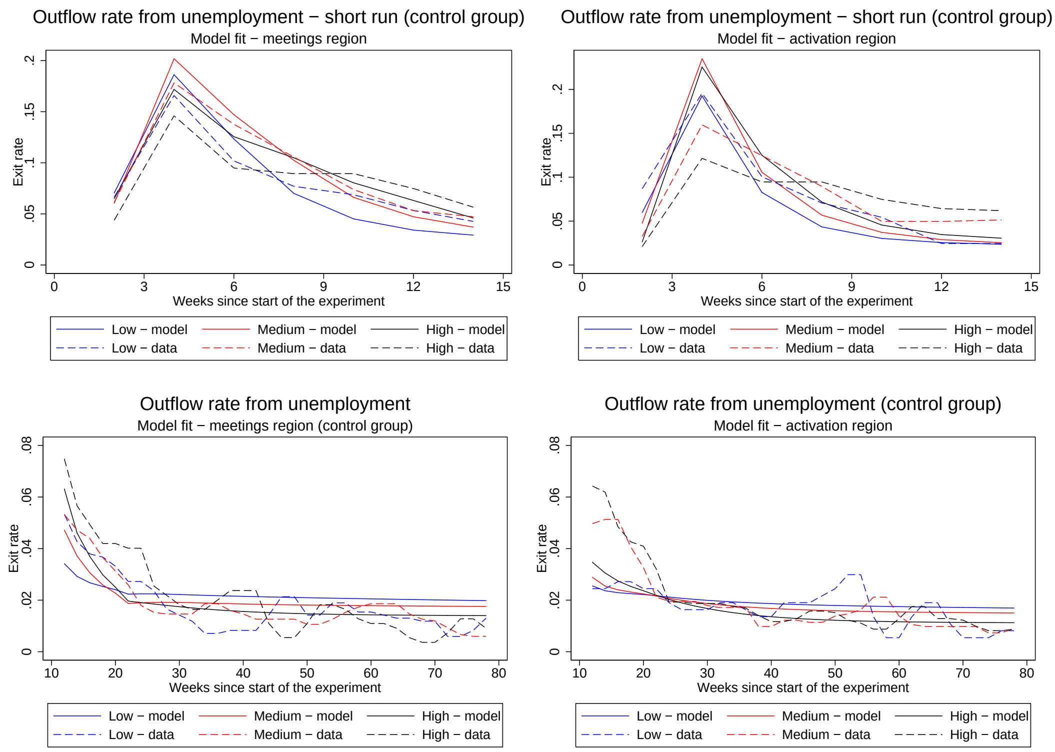
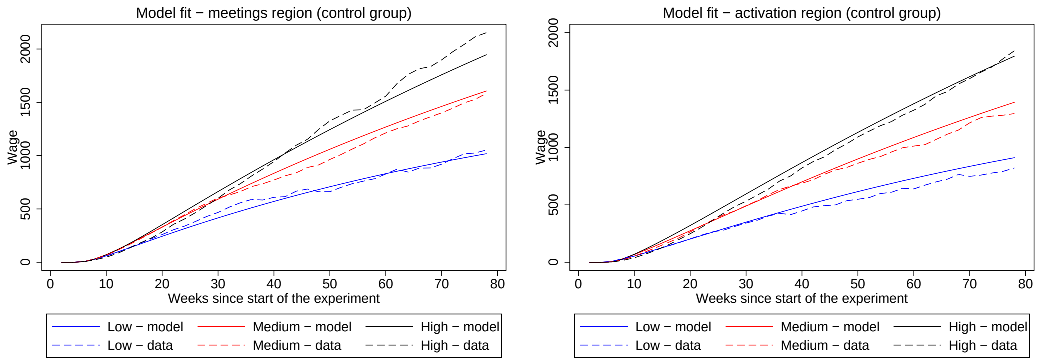
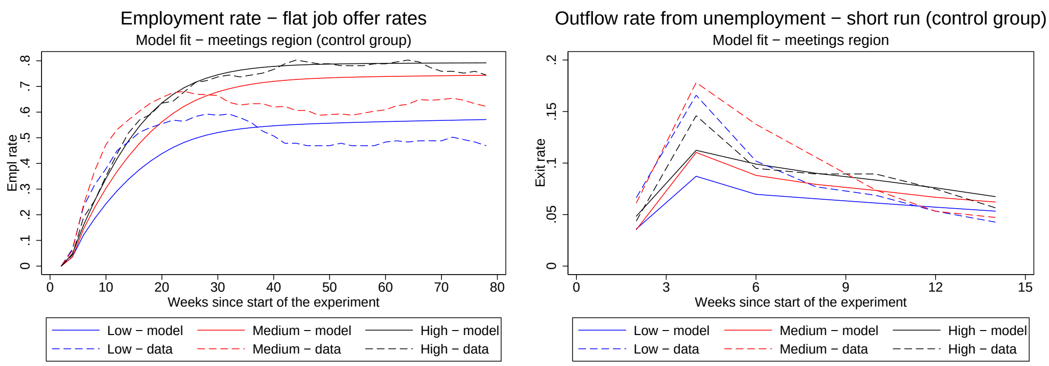


Figure 1.20: Model fit: Employment duration and wages



Note: the figure gives the data and model prediction for moment 13 defined in Table 1.6

Figure 1.21: Eliminating duration dependence



Note: The figure compares data to model predictions in a model where job offer arrival rates do not vary with unemployment duration

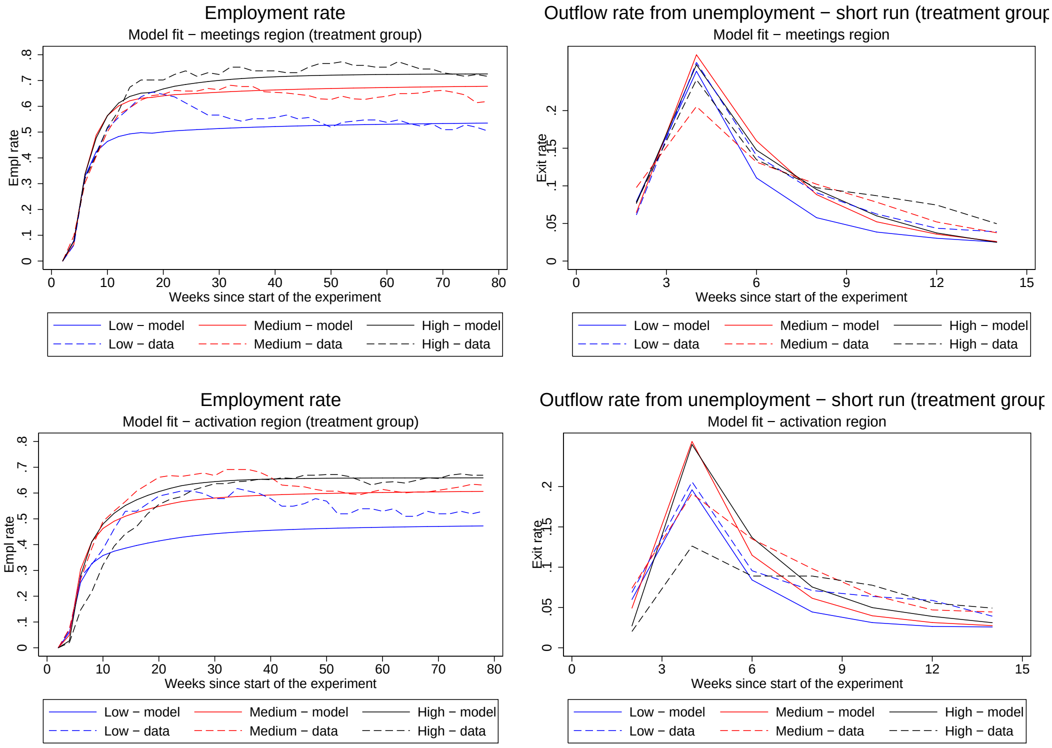
Table 1.18: Heterogeneity in the compensating variation (meetings region)

State variables		Compensating variation	
Unemployment duration (cu)	Skill level (hc)	Low (type 1)	Low (type 2)
0	0	1014	652
1	0	1068	947
2	0	1160	1047
3	0	1297	1121
5	0	1747	1531
7	0	2086	1809
9	0	2138	2025
0	1	914	592
0	2	863	104
0	3	674	111
0	4	469	104
0	5	414	50

Note: The table reports the CV (defined in 1.3) for different unemployment durations and skills for low educated individuals in the meetings region. All other state variables are set to 0.

1.11 Appendix C: Model fit for the treatment group

Figure 1.22: Employment (data and model comparison for the treatment group)



Note: see the appendix for further graphs on inflow rates for later weeks.

Figure 1.23: Average (squared) unemployment duration (data and model comparison for the treatment group)

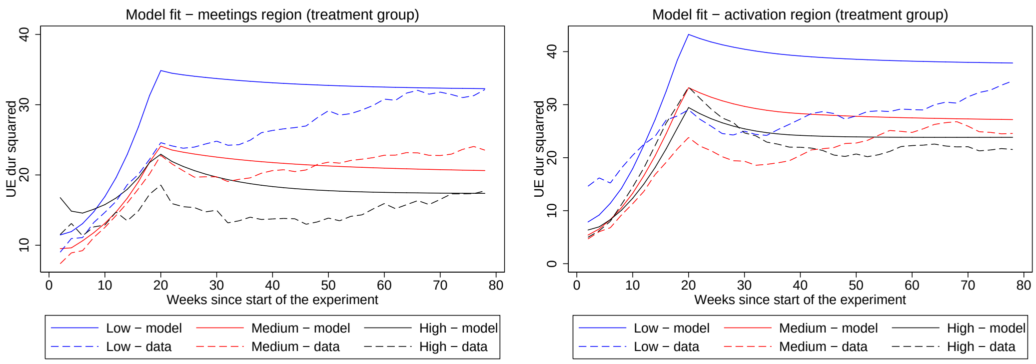


Figure 1.24: Squared wages (data and model comparison for the treatment group)

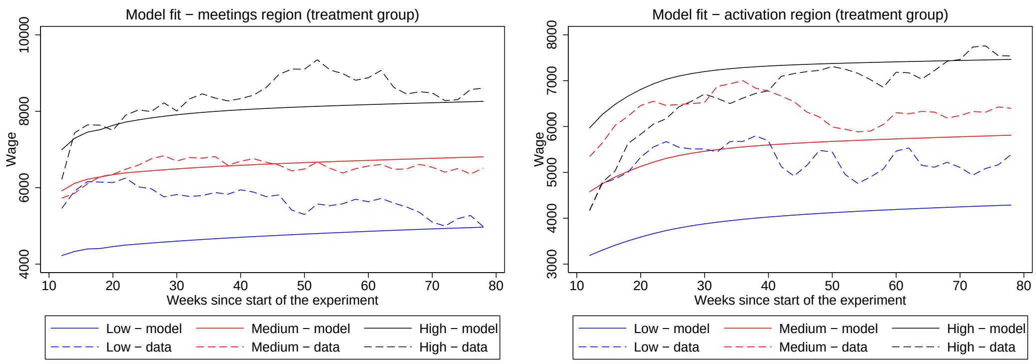


Figure 1.25: Job loss (data and model comparison)