

Essays on the Evaluation of Social Programmes

Mónica Sofia Rodrigues da Costa Dias

A dissertation submitted to the Department of Economics
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy
University College London

London
September 2002

ProQuest Number: U642551

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest U642551

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

This thesis focus on the evaluation of social programmes, a subject attracting renewed interest given the deteriorating labour market position of unskilled workers in many industrialised countries and the large financial requirements programmes designed to move individuals away from welfare place on public budgets. Non-experimental methods face the difficult missing counterfactual problem, where individuals are either participants or non-participants, not both simultaneously. I use different approaches to assess the impact of interventions using the New Deal for Young Persons (NDYP) as role model. This UK mandatory program involves extensive job-search assistance and options that include tax credit schemes and education subsidies.

Chapters 2 and 3 address the evaluation of direct effects of interventions. I start by reviewing the recent methodological developments discussing social experiments, natural experiments, matching methods and instrumental variables in chapter 2. Chapter 3 presents the evaluation of the employment effect of the NDYP. Identification exploits the differential timing of the programme introduction across regions and age-related eligibility rules and estimation uses a variety of techniques combining “difference in differences” and matching procedures.

Being a global programme of wide implementation, the NDYP is expected to impact on prices, indirectly affecting the whole economy and challenging the validity of counterfactuals constructed from observed data. Chapters 3 and 4 focus on the structural overall evaluation of a stylised version of the NDYP initiative, avoiding “no-indirect-effects” assumptions. A general equilibrium model of savings, skills and human capital with labour supply is developed within an overlapping generations set up, allowing for idiosyncratic uncertainty under risk aversion, fixed costs and discrete working and studying choices. The model’s parameters are identifiable with currently available data and a procedure to estimate the human capital production function is proposed. The model was numerically solved for the SS and the effects of tax credit policies were simulated.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 13 |
| 1.1 | The NDYP within the British labour market | 16 |
| 1.2 | Outline of this thesis | 20 |
| 2 | Alternative empirical approaches to evaluation problems | 22 |
| 2.1 | Which Parameter of Interest? | 26 |
| 2.1.1 | Homogeneous Treatment Effects | 29 |
| 2.1.2 | Heterogeneous Treatment Effects | 29 |
| 2.2 | Experimental And Non-Experimental Data | 30 |
| 2.2.1 | Experimental Data | 30 |
| 2.2.2 | Non-Experimental Data | 31 |
| 2.3 | Methods For Non-Experimental Data | 31 |
| 2.3.1 | The Instrumental Variables (IV) Estimator | 32 |
| 2.3.2 | The Difference In Differences (DID) Estimator | 37 |
| 2.3.3 | The Matching Estimator | 40 |
| 2.3.4 | Matching and DID | 44 |
| 2.4 | Interpreting the Evidence | 45 |
| 2.4.1 | The LaLonde Study and the NSW Evaluation | 45 |
| 2.4.2 | A Critique of the LaLonde Study | 48 |
| 2.4.3 | A Simulation Study | 52 |
| 2.4.4 | Matching and Difference in Differences: An Area Based Evaluation of the British NDYP | 55 |
| 2.5 | Conclusions | 55 |

| | | |
|----------|---|-----------|
| 3 | The employment impact of job-search assistance | 57 |
| 3.1 | The NDYP and the choice of the outcome variables | 61 |
| 3.2 | Identification and Estimation Methods | 63 |
| 3.2.1 | Implementation | 69 |
| 3.2.2 | Other estimation issues | 70 |
| 3.3 | Data | 73 |
| 3.4 | Results | 76 |
| 3.4.1 | Pilot Study: men's results | 76 |
| 3.4.2 | National Roll Out: men's results | 82 |
| 3.4.3 | The impact of the programme on women | 86 |
| 3.5 | Conclusions | 89 |
| 3.6 | Appendix to chapter 3 | 89 |
| 3.6.1 | Data | 89 |
| 3.6.2 | Gateway employment effects under different propensity score matching techniques | 92 |
| 3.6.3 | Estimation methods | 93 |
| 4 | The dynamics of earnings in a life-cycle model of labour supply | 97 |
| 4.1 | A human capital model of labour supply | 100 |
| 4.1.1 | The decision rule | 102 |
| 4.1.2 | The dynamics of the state variables | 104 |
| 4.2 | Properties of the model | 105 |
| 4.3 | Estimation process | 109 |
| 4.3.1 | Identification using the structural model | 110 |
| 4.3.2 | Practical estimation procedure | 116 |
| 4.4 | The data | 119 |
| 4.5 | Estimation results | 121 |
| 4.5.1 | Earnings model at the beginning of the working life. | 121 |
| 4.5.2 | The rates of human accumulation | 126 |
| 4.5.3 | Sensitivity analysis | 129 |
| 4.6 | Conclusions | 133 |

| | | |
|----------|---|------------|
| 4.7 | Appendix to chapter 4 | 135 |
| 4.7.1 | Proof of Lemma 2 | 135 |
| 4.7.2 | Proof of Lemma 3 | 138 |
| 4.7.3 | The NCDS58 data | 138 |
| 5 | The overall evaluation of labour market interventions | 140 |
| 5.1 | The Model | 145 |
| 5.1.1 | The production sector | 146 |
| 5.1.2 | The individual's problem | 147 |
| 5.1.3 | Characterising the equilibrium | 152 |
| 5.2 | Identification issues | 154 |
| 5.2.1 | The aggregate technology | 155 |
| 5.2.2 | Education decisions and unemployment benefit | 157 |
| 5.2.3 | Other parameters | 159 |
| 5.3 | Labour market policies: assessing the overall effects of wage subsidies | 160 |
| 5.3.1 | Directions of the programme effects | 161 |
| 5.3.2 | Results from the simulations | 163 |
| 5.4 | Conclusions | 171 |
| 5.5 | Appendix to chapter 5 | 172 |
| 5.5.1 | Proof of lemma 7 | 172 |
| 5.5.2 | Proof of lemma 8 | 173 |
| 5.5.3 | Proof of lemma 9 | 173 |
| 5.5.4 | Proof of lemma 10 | 174 |
| 5.5.5 | Proof of lemma 11 | 175 |
| 6 | Further developments | 177 |

List of Figures

| | | |
|-----|--|-----|
| 1.1 | Simplified flow diagram of the NDYP. | 18 |
| 3.1 | Outflows from the JSA conditional on completing 6 months on the claimant count. Effect by the end of the 10th month on JSA. Men only. | 71 |
| 3.2 | Unemployment - claimant count and ILO measures | 76 |
| 3.3 | Difference in differences estimates over time. Outflows to all destinations. Men only. | 82 |
| 3.4 | Women outflows from the JSA conditional on completing 6 months. Effect by the end of month 10. | 87 |
| 3.5 | Distribution of the PS before and after matching: men aged 19-24 in Pathfinder Pilot and control areas. | 91 |
| 3.6 | Distribution of the PS before and after matching: single men aged 19-24 in Pathfinder Pilot and control areas. | 92 |
| 3.7 | Distribution of the PS before and after matching: men w/ less than 6 months on unemployment in the last 2 years in Pathfinder Pilot and control areas. | 93 |
| 3.8 | Distribution of the PS before and after matching: men with less than 2 unemployment spells in the last 2 years in Pathfinder Pilot and control areas. | 94 |
| 4.1 | Rates of human capital accumulation by ability and educational levels. | 128 |
| 4.2 | Sensitivity test 1 - Upper bound to the rates of human capital accumulation by ability and education. | 132 |
| 4.3 | Sensitivity test 2 - Upper bound to the rates of human capital accumulation by ability and education. | 134 |
| 5.1 | Impact of the wage subsidies on the working decisions. | 166 |

5.2 GE effects of the wage subsidy (policy 2) on the working decisions by treatment status and working status in the baseline case. 167

5.3 Impact of the wage subsidies on the investment in education decisions. 168

5.4 GE effects of the wage subsidy (policy 2) on the education decisions by treatment status and working status in the baseline case. 169

List of Tables

| | | |
|-----|--|-----|
| 2.1 | Observable characteristics for NSW males. | 46 |
| 2.2 | Annual earnings for NSW males. | 47 |
| 2.3 | Effects of treatment on the treated for the NSW males. | 47 |
| 2.4 | Bias associated with alternative estimates of the TTE among NSW males. | 49 |
| 2.5 | Bias decomposition of the simple differences estimator. | 51 |
| 2.6 | Bias in non-experimental estimates of the TTE. | 54 |
| | | |
| 3.1 | Effect of treatment on the treated men. Pilot period. | 77 |
| 3.2 | NDYP effect on the ineligible and the whole economy. Pilot period. | 78 |
| 3.3 | Flows from the claimant count into employment. | 79 |
| 3.4 | Effect of treatment on the treated men. Pilot and National Roll Out periods. | 83 |
| 3.5 | Robustness of the results. | 85 |
| 3.6 | Effect of treatment on the treated women. Pilot period. | 88 |
| 3.7 | Descriptive statistics for different treatment and control groups. | 90 |
| 3.8 | Effect of treatment on the treated men. Pilot period. | 95 |
| 3.9 | NDYP effect on the ineligible and the whole economy. Pilot period. | 96 |
| | | |
| 4.1 | Selection model of earnings. | 123 |
| 4.2 | Selection model of earnings. | 124 |
| 4.3 | Standard error for the productivity shock. | 125 |
| 4.4 | Initial rate of human capital accumulation. | 126 |
| 4.5 | Adjustment rate of the rate of human capital accumulation. | 127 |
| 4.6 | Distribution of the productivity shock. | 127 |
| 4.7 | Sensitivity test 1. | 130 |

| | | |
|------|--|-----|
| 4.8 | Sensitivity test 2. | 133 |
| 4.9 | Observations in the NCDS58. | 139 |
| 4.10 | Observations used to compute the rate of human capital accumulation. | 139 |
| 5.1 | Aggregate production function parameters. | 156 |
| 5.2 | Education related parameters. | 158 |
| 5.3 | Investment in education: Data vs. simulated moments. | 158 |
| 5.4 | Working decisions: Data vs. simulated moments. | 158 |
| 5.5 | Labour market transitions: Data vs. simulated moments. | 159 |
| 5.6 | Steady state prices. | 164 |
| 5.7 | Programme participation rates. | 164 |
| 5.8 | Lifetime net wealth and welfare: descriptives. | 170 |
| 5.9 | Proportions benefiting from the programme. | 170 |

Acknowledgements

Costas Meghir and Richard Blundell exceeded by far all my most optimistic expectations about supervisors. Costas has been greatly supportive, enthusiastic, generous and friendly at literally all times throughout my PhD. He was always available when I needed and proved of infinite patience in dealing with my most unbelievable worries and doubts. I cannot fairly measure his contribution to the present work as it would still be in its early stages, and I would still be lost in the wilderness of the economics' subject, if it was not him. Richard was the person introducing me to the evaluation problem, the Institute for Fiscal Studies and the world of research. He has equally added to this research and to my development by discussing an incredibly wide range of topics in equal detail, always providing the most insightful comments. This is even more impressive when taken together with his fantastic ability of cheer up and motivate the most ran-down student. I have got no words to express my gratitude to both Costas and Richard. I feel the most fortunate person for having had the honour of being their student and interacting with them throughout the last few years. They have been, and will continue to be, a major source of inspiration on both my work and my life.

This thesis benefited immensely from the interaction with John Van Reenen. He is not only the co-author of a paper that is the substance of chapter 3. John has participated in a large number of discussions about other aspects of the work presented here and has always contributed with precious comments and suggestions. To him I am also sincerely grateful.

Other persons have largely contributed to this research as well. Hamish Low has discussed parts of it with me in several occasions, always being extremely helpful. Orazio Attanasio, James Banks, James Heckman, Hide Ichimura, Derek Neal and Gianluca Vi-

olante have provided invaluable comments at different stages of this work, from which I profited to a great extent. I am especially indebted with Hide for his care and dedication to make PhD students at UCL a group of researchers able to handle the tools required for a formal analysis of economic problems. I took great benefit from having participated in his “maths meetings”. An enormous thanks goes to Barbara Sianesi as well, my dearest PhD colleague. Discussions with her have been most important at this stage of my life as she has the sharpest mind and is the most cheerful person.

I must also express my gratitude to Professors Stephen Nickell and Christopher Taber for agreeing to formally examine my thesis and for their most helpful, insightful and constructive comments.

The Institute for Fiscal Studies has provided the most enjoyable and friendly environment, making my student life so much easier. I am especially indebted with Erich Battistin, Mike Brewer, Lorraine Dearden, Alissa Goodman, Amanda Gosling and Howard Reed for help and advice on different aspects of the British labour market and education systems, available data and resources at IFS.

Nuno, my husband, was always the most supportive person. There has never been anyone as able to transform my most unsociable moods in big laughs by making fun of me in such a sweet and irresistible way. I can never thank him enough for having been on my side so unconditionally, always putting my interests above his own ones. Mariana, my sweet little girl, was the best thing that happened to me during my PhD years and, in fact, during my whole life. She filled my time with smiles and laughs, short nights and full days that I can never forget. I thank her for existing and being such a joy. To my two beloved ones, I will add my father, to whom I dedicate this thesis for his love, affection and enormous interest, and mother, with whom I spent hours on the phone sharing mine and hers experiences throughout these years.

This work would have never been possible without the financial support received from Fundação para a Ciência e a Tecnologia, grant number PRAXIS XXI/BD/11413/97 and I am extremely grateful for having been given such opportunity. I am also thankful to the Leverhulme Foundation for funding this project. This research is part of the programme of research at the ESRC Centre for the Microeconomic Analysis of Fiscal Policy at IFS. The DfEE kindly provided access to NDED and the ESRC Data Archive to the JUVOS

and GHS data sets.

Declaration

1. No part of this thesis has been presented to any University for any degree.
2. Chapter 2 was undertaken as joint work with Richard Blundell.
3. Chapter 3 was undertaken as joint work with Richard Blundell, Costas Meghir and John Van Reenen.

Mónica Costa Dias

Chapter 1

Introduction

Throughout the industrialized countries there has been a growing concern over the labour market position of lower educated and less skilled young people. Their labour market position in terms of jobs and wages has generally deteriorated in the last two decades. In response to this, there has been a renewed interest in labour market programmes designed to move individuals into work and away from welfare. Such programmes usually work on two fronts. First, by creating direct incentives for individuals to move to employment (e.g. through the introduction of wage subsidies or tax incentives, the imposition of tighter conditions on unemployment benefit eligibility, or the creation of better support conditions for workers with particular needs like single parents). Secondly, by providing incentives for unemployed individuals to improve their labour market skills through training and/or education schemes, making the working option more attractive.

A recent feature of such welfare policies is the call for independent evaluation. It is now recognised that different groups in the society extract different payoffs either from treatment or from the sole existence of the programme, making the size and distribution of such impacts an important input in the designing process of new interventions. In addition, much of the attention has been captured by the costs involved in implementing and running social interventions, which can be quite significant. Hence, evaluation studies have been developed on two main grounds to determine how effective welfare policies are: the individual and the aggregate levels. At the individual level, the performance of the treatment in changing participants' outcomes has received most of the attention. At the

aggregate level, the interest concentrates on the assessment of the relative size of the social costs and benefits from interventions.

The evaluation process, however, has not proved an easy task. The main problem can be regarded as one of missing data. Individuals are either treated or not-treated at each moment in time, never both simultaneously. At the economy at large, either the programme has been implemented or not, but again the two cases are not simultaneously observed. Self-selection along with heterogeneous treatment effects, different labour markets with specific economic conditions, economy-wide effects through prices' adjustments and scarcity of good detailed data all contribute to the problem of finding the adequate counterfactual in evaluation methods. Social experiments could, in principle, solve part of the problem with randomisation, but they are rarely available especially in Europe where some of the largest interventions take place. Moreover, serious criticisms have been risen on the quality of experimental data concerning the bureaucratic procedures on the actual implementation of randomisation, the response to randomisation by participants and the potential disruptions introduced by evaluation efforts during the experiment. Even when the experimental conditions have been fulfilled, it is often difficult to extract meaningful economic information from it. More likely than not, the identified effects are not generalisable to other groups, other local labour markets, other levels of implementation or other similar interventions. Notwithstanding, good quality experimental data can provide more reliable results than non-experimental one and be very useful to assess the adequacy of different empirical methods in evaluation.

Recognising that the effects from social interventions can be quite heterogeneous presents new difficulties to the evaluation problem. Numerous parameters of interest can be defined, which are inexorably linked to the data being used in empirical experimental or non-experimental settings. Such diversity complicates the interpretation process considerably when results from different studies are being compared, all based on different data, policies and underlying assumptions. An important message from the extensive review of the literature by Heckman, LaLonde and Smith (99) is that the estimation methodology is intrinsically linked to the parameter one wishes to identify given the data at hand and the economics of the problem. By clearly establishing the underlying identification assumptions one is able to precisely locate the information being disclosed when studying a

social intervention. This is a major advantage of structural models, where the economics of the problem is openly discussed.

Large programmes with wide implementation within certain groups may also bring further problems to the evaluation process. By succeeding in changing the labour market behaviour of participants, the programme may also affect prices. This happens not only because the government budgeting is likely to be affected, bringing consequent adjustments in the tax rates. Also, the composition of labour market groups is expected to change, which itself can affect prices. Moreover, a large programme that induces participants to work may also affect the market through substitution, whereby non-treated agents supplying labour substitutable for that of the treated may find their opportunities worsened by the increase in the supply of labour and a relative preference for treated agents. The additional complication with the effects just described is that they change the economic environment, impacting on both participants and non-participants. Non-participants are indirectly affected by the existence of the programme, ceasing to be a suitable source to construct the counterfactual. The plague of the evaluation problem under such circumstances is that non-participants that closely resemble participants are more likely to be affected by the indirect effects of the programme, while non-participants of different characteristics do not usually participate in the same labour market. This happens even if full control for the selection process is achieved, both on observables and unobservables.

This thesis focus on the evaluation of global and generous labour market programmes. The role model being used throughout the study is that of the New Deal for the Young People (NDYP), a British labour market programme released in 1998, one year after the new Labour government came to power. Our approach to tackle the evaluation problem is to discuss it under different perspectives, trying to address its several difficulties emanating from the missing data and the size of the intervention. Two different evaluation strategies are adopted in what follows. The first is the more common *ex-post* evaluation, where data on participants and non-participants before and after treatment is used to try to assess the direct effects of the treatment. The second strategy comprises an *ex-ante* form of evaluation, where the ‘no-indirect effects’ assumption is relaxed at the cost of additional structure.

There are gains from taking the two approaches simultaneously as they complement

each other in terms of the analysis and inform policy-making of different aspects of the impact of interventions. When applied to a well-posed problem, the *ex-post* analysis provides a simple setting to answer a precisely defined question. The question is usually formulated as ‘What is the short-run, direct impact of treatment on the outcome x among treated?’, which is convincingly answered by a number or set of numbers extracted from the actual observation of participants behaviour. On the contrary, the *ex-ante* approach uses a considerable amount of structure to provide an overall framework that guides the economic reasoning about the potential effects of social interventions. It answers the broader question of explaining the economic behaviour of individuals causing the measured effects. Simultaneously, it justifies the choice of the parameters of interest in the *ex-post* analysis by establishing its economic meaning and sheds light on the potential drawbacks of that analysis. Moreover, its focus on long-run global effects completes the analysis of the impact of interventions and provides tools for policy-making that would not otherwise be available. This procedure does not, however, provide definitive answers to the evaluation problem as the *ex-post* analysis does. It is intrinsically determined by the assumptions imposed and also dependent on the estimation methodologies adopted and data available to determine the structural parameters of the model. Such dependencies are not absent from the *ex-post* analysis but are weakened by the fundamental assumption of such approach, which rules out indirect effects from interventions.

In the rest of this introductory chapter we provide a brief overview of the NDYP within the British unemployment protection system. The design of the programme is central to much of the analysis that follows and that explains the decision of including its discussion here. We finish by overviewing the object of the following chapters in the second section.

1.1 The NDYP within the British labour market

Most of the studies focusing on the impact of social policies targeted at the unemployed respect to US programmes. Small, frequently not statistically significant treatment effects on the treated are the rule. However, US programmes are often characterised by being targeted at very disadvantaged and unskilled groups of the population. Additionally, treatment is not sufficiently long and intense in many cases to actually affect their human

capital or to truly improve participants incentives to work. On the contrary, European programmes have been targeted at wider groups of the population and are usually more generous, including a variety of possible treatments available. Under heterogeneous treatments and heterogeneous effects from treatment, it is licit to wonder about the effectiveness of labour market programmes of the European style.

The focus of this thesis is the NDYP, a recent initiative of the UK government to help the young unemployed gain work. It is targeted at the 18 to 24 years old longer-term unemployed and has been launched in two stages. It was first tried out in some areas during the Pilot period running from January till March 1998. Twelve areas were selected for this experiment, called the Pathfinder Pilots.¹ From April 1998 onwards, the programme was launched in the whole UK, and this is called the National Roll Out. On all grounds, the NDYP is a rather complex labour market policy, including features of social interventions ranging from job-search assistance, tax credit schemes, tuition subsidies, training and sanction policies. Moreover, participation is compulsory, so that every eligible individual who refuses to participate risks losing entitlement to benefits. The criteria for eligibility are simple: agents aged between 18 and 24 by the time of completion of the sixth month on Job Seekers' Allowance (JSA) are assigned to the programme and start receiving treatment. Given the stated rules, the programme can be classified as one of 'global implementation', being administered to everyone in the UK meeting the eligibility criteria. Indirect effects that spill over to other groups than the treatment group may occur, and this is a central issue to the analysis on this thesis.

The path of a participant through the NDYP is composed of three main steps (see Figure 1.1). Assignment to the programme happens at the end of the 6th month on the JSA claimant count. This is when the individual starts the first stage of the treatment called the Gateway. It lasts for up to 4 months and is composed of intensive job-search assistance and small basic skills' courses taking about two weeks each. Each individual is assigned a 'Personal advisor', a mentor who they meet at least once every two weeks to encourage/enforce job search.

The second stage is composed of four possible options. First, there is the employer option - a six-month spell on a subsidised employment. For the subsidised employment

¹See Anderson, Riley and Young, 1999.

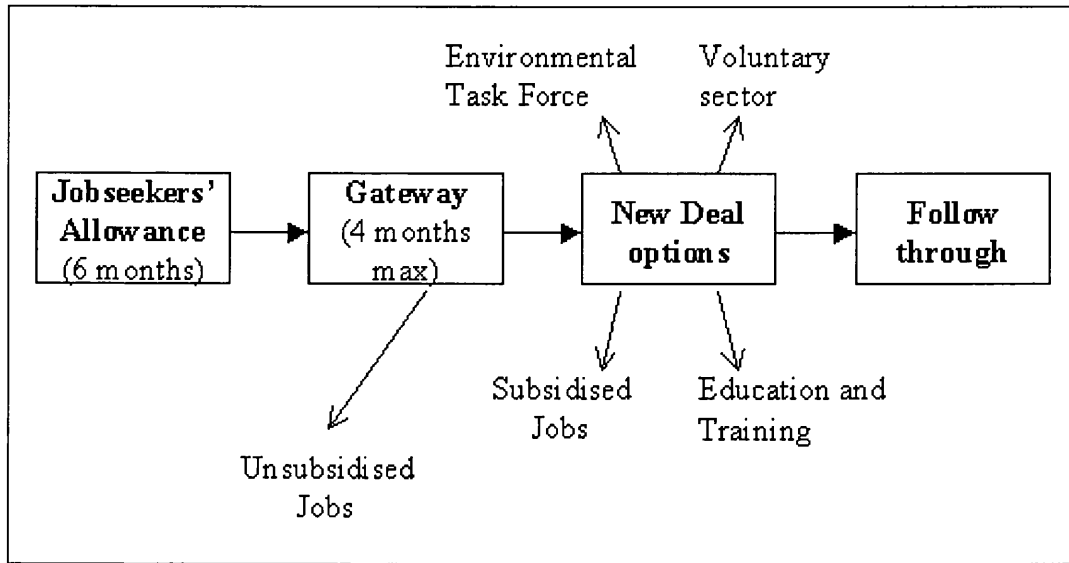


Figure 1.1: Simplified flow diagram of the NDYP.

option, the employer receives a £60 a week wage subsidy during the first six months of employment plus an additional £ 750 payment for a required minimum amount of job training equivalent to one day a week. Second, an individual can enrol in a stipulated full-time education or training course and receive an equivalent amount to the JSA payment for up to twelve months. Under this option they may be eligible for special grants in order to cover exceptional expenses. Third, individuals can work in the voluntary sector for up to six months, being paid a wage or allowance at least equal to JSA plus £400 spread over the six months. Finally, they may take a job on the Environmental Task Force, essentially composed of government jobs. The payment on this option is of a wage or allowance at least equal to JSA plus £400 spread over the six months.

Once the option period is over, if the individual has not managed to keep/find a job or leave the claimant count for any other reason, the third stage of the programme is initiated, the Follow Through. This is a process similar to the Gateway, taking up to 13 weeks, where job-search assistance is the main treatment being provided.

Finally, at any stage of the programme, a participant refusing to collaborate is liable to suffer a benefits' sanction. At a first stage, sanctions assume the form of withdrawal from benefits for a 2 weeks periods. Further refusals are liable to withdrawal from benefits

during 4 weeks' periods.

An important feature of the NDYP is that it effectively introduced a limit to the time period an unemployed is eligible for benefits in the UK. The main benefit available for unemployed young people is the Jobseeker's Allowance (JSA). It was introduced in October 1996 to replace unemployment benefit. The level of JSA is about £40 a week throughout the NDYP period, though this amount depends on the age of the applicant and the respective household income and needs. To be eligible for JSA, an unemployed person must: (i) Be 'actively seeking work', which is assessed by a fortnightly short interview taking 5-10 minutes; and (ii) Meet some conditions concerning the past two tax years working history, related to the amount of National Insurance contributions made while employed for 'contributory JSA' or, alternatively, pass a 'means test' for 'means tested JSA'. Thus, it is possible for someone who never worked before to be entitled for the benefit. In a reform in 1986 (RESTART) more intensive job focused interviews took place at six monthly intervals. If not before, receipt of JSA becomes 'means tested' after six months. Individuals with significant income from other sources, let it be assets or a partner bringing in income, have their JSA scaled down or taken away altogether. Prior to October 1996, this period of 'non-means tested' unemployment benefit was one year. Otherwise, the JSA imposes no time limit: as long as the conditions are met, an applicant is entitled to it.

The recent study by Van Reenen (01) analyses the NDYP within the UK labour market history. It is clear from the study that the UK economy is quite volatile, which translates into sharp booms and busts in the numbers out of work and education. Moreover, long-term unemployment becomes negligible during upturns but rises up to 50% of the total unemployment during the recent recessions. In particular, the NDYP was introduced during the latest boom, when unemployment is falling to an historical minimum since the seventies. This is likely to affect the impact of the programme as the pool of unemployment at this stage is different in composition from what it would be during downturns. With heterogeneous treatment effects, the estimated impact is specific to the economic conditions in which the programme is introduced. Thus, interpreting the evaluation results should take into consideration the specific conditions the programme is introduced.

1.2 Outline of this thesis

The main part of this thesis is divided into four main chapters, starting with the *ex-post* type of evaluation in chapters 2 and 3. The *ex-ante* approach is pursued in chapters 4 and 5. A final chapter sets the path to further work. The following briefly describes the object of chapters 2 through 5.

Chapter 2 overviews the recent developments in evaluation methods in empirical micro-economics, setting the ground for the evaluation study presented in the following chapter. Four alternative but related approaches to empirical evaluation of policy interventions are studied: social experiments, natural experiments, matching methods, and instrumental variables. In each case the necessary assumptions and the data requirements are considered for estimation of a number of key parameters of interest. These key parameters include the average treatment effect, the effect of treatment on the treated and the local average treatment effect. Some issues of implementation and interpretation are discussed drawing on the labour market programme evaluation literature.

Chapter 3 uses the methodological insight provided in chapter 2 to evaluate the impact of the NDYP using administrative panel data on individuals between 1982 and 1999. Given the timing of the programme, the potential duration of the treatment and the time limit in the available data, we concentrate on the effects of the first four months of treatment formed mainly of intensified job-search assistance, the Gateway. The main aim of this study is to identify the impact of treatment on the treated, which is done by exploiting both the differential timing of the introduction of the programme across regions and the age-related eligibility rules. The effects of the Gateway are measured on the employment probabilities of the treated, a major concern of the policy-makers driving the creation of the NDYP. A variety of estimation techniques exploring combined ‘difference in differences’ and matching procedures are used to identify the parameter of interest. Based on the pilot study we find that the NDYP programme raised employment by a significant 5 percentage points. However we present some evidence suggesting that this effect may not be sustained in the longer run. Potential indirect effects, including substitution and equilibrium price effects, also deserve considerable attention in the analysis. The design of the experiments ran was thought to provide as much information as possible on these type of effects. Such

analysis provides no strong indications that indirect effects are contaminating the results, an expected result given the early stages of the programme at the evaluation moment.

Chapters 4 and 5 focus on the *ex-ante* analysis. An overlapping generations general equilibrium model of savings, skills and human capital with labour supply is developed. Idiosyncratic uncertainty under risk aversion is considered, along with fixed costs and discrete working and studying choices. A major concern assisting the setup of the model was to use the lessons from the empirical literature to establish its founding assumptions. A parallel goal was, of course, to keep it at a level of simplicity that would allow the estimation and numerical solution of the problem to be carried out.

Chapter 4 addresses the individual's problem, discusses its structure and establishes some major results fundamental for the estimation and the solution procedures to be developed. The structural estimation of wage equations is the main subject of this chapter, an important issue on its own that is essential to perform overall evaluations of labour market interventions. The focus is on the earnings dynamics throughout the agent's working life, which is determined by the idiosyncratic working experience and learning ability. Heterogeneous rates of human capital accumulation are estimated using the simple procedure developed.

Chapter 5 formally presents the overall model of the economy and approximates the remaining parameters required. The model is solved for the steady state and used to simulate the impact of wage subsidies on the participants, non-participants and economy at wide. The wage subsidy policy is one of the major initiatives advanced by the NDYP and that is why we have chosen to simulate it. Notwithstanding, the model can be used to simulate other aspects of the NDYP. The results suggest that large indirect effects may result from the introduction of such programmes, of an order of magnitude comparable to that of direct partial equilibrium effects. Such outcome calls for extra caution when interpreting results from *ex-post* evaluation as prices' changes affect both participants and non-participants. Moreover, selection into the programme is a potential major issue as the composition of the treatment group may not correspond to the original target group, implying severe deadweight loss.

Chapter 2

Alternative empirical approaches to evaluation problems

In this review we discuss different approaches to the evaluation problem in empirical microeconomics. We understand these approaches as methods of constructing the missing counterfactual, the central piece in any evaluation strategy. The four distinct but closely related approaches being considered are: (i) social experiments, (ii) natural experiments, (iii) matching methods, and (iv) instrumental methods. Though the former two are generally understood as methods of collecting/interpreting data, one should bear in mind that the available data determines the appropriate characteristics of the estimation method. In the particular case of social experiments and natural experiments, it is most frequent to see simple differences and difference-in-differences associated with these data, respectively. Thus, we follow this wording in the current chapter.

The first of these approaches (i) is closest to the ‘theory’ free method of medical experimentation since it relies on the availability of a randomised control. The last approach (iv) is closest to the structural econometric method since it relies directly on exclusion restrictions. Natural experiments and matching methods lie somewhere in between in the sense that they attempt to mimic the randomised control of the experimental setting but do so with non-experimental data and consequently place reliance on independence and/or exclusion assumptions.

Our concern here is with the evaluation of a policy intervention at the microeconomic

level. This could include training programmes, welfare programmes, wage subsidy programmes and tax-credit programmes, for example. At the heart of this kind of policy evaluation is a missing data problem since, at any moment in time, an individual is either in the programme under consideration or not, but not both. If we could observe the outcome variable for those in the programme had they not participated then there would be no evaluation problem of the type we discuss here. Thus, constructing the counterfactual is the central issue that the evaluation methods we discuss address. Implicitly, each of the four approaches provides an alternative method of constructing the counterfactual.

The literature on evaluation methods in economics is vast and continues to grow. There are also many references in the literature which document the development of the analysis of the evaluation problem in economics. In the labour market area, from which we draw heavily in this review, the ground breaking papers were those by Ashenfelter (78), Ashenfelter and Card (85) and Heckman and Robb (85, 86).

In many ways the *social experiment method* is the most convincing method of evaluation since it directly constructs a control (or comparison) group which is a randomised subset of the eligible population. The advantages of experimental data are discussed in papers by Bassi (83, 84) and Hausman and Wise (85) and were based on earlier statistical experimental developments (see Cochran and Rubin (73) and Fisher (51), for example). A properly defined social experiment can overcome the missing data problem. For example, in the design of the impressive study of the Canadian Self Sufficiency Project reported in Card and Robbins (98), the labour supply responses of approximately 6,000 single mothers in British Columbia to an in-work benefit programme, in which half those eligible were randomly excluded from the programme, were recorded. This study has produced invaluable evidence on the effectiveness of financial incentives in inducing welfare recipients into work.

Of course, social experiments have their own drawbacks. They are rare in economics and typically expensive to implement. They are not amenable to extrapolation. That is, they cannot easily be used in the ex-ante analysis of policy reform proposals. They also require the control group to be completely unaffected by the reform, typically ruling out spillover, substitution, displacement and equilibrium effects on wages etc. None-the-less, they have much to offer in enhancing our knowledge of the possible impact of policy

reforms. Indeed, a comparison of results from non-experimental data to those obtained from experimental data can help assess appropriate methods where experimental data is not available. For example, the important studies by LaLonde (86), Heckman, Ichimura and Todd (97) and Heckman, Smith and Clements (97) use experimental data to assess the reliability of comparison groups used in the evaluation of training programmes. We draw on the results of these studies below.

It should be noted that randomisation can be implemented by area. If this corresponds to a local (labour) market, then general equilibrium (GE) or market level spillover effects will be accounted for. This is more likely to be true in the short run as economic agents in adjacent areas may take a while to respond to changes close by but not within their region. The use of control and treatment area designs is a feature of the New Deal for the Young People (NDYP) evaluation data base in the UK. In the discussion below the area to area comparisons are used to comment on the likely size of GE and spillover effects.

The *natural experiment approach* considers the policy reform itself as an experiment and tries to find a naturally occurring comparison group that can mimic the properties of the control group in the properly designed experimental context. This method is also often labelled “difference-in-differences” since it is usually implemented by comparing the difference in average behaviour before and after the reform for the eligible group with the before and after contrast for the comparison group. In the absence of a randomised experiment and under certain very strong conditions, this approach can be used to recover the average effect of the programme on those individuals entered into the programme - or those individuals “treated” by the programme. Thus measuring the average effect of the treatment on the treated. It does this by removing unobservable individual effects and common macro effects. However, it relies on the two critically important assumptions of (i) *common time effects across groups*, and (ii) *no systematic composition changes within each group*. These two assumptions make choosing a comparison group extremely difficult. For example, in their heavily cited evaluation study of the impact of Earned Income Tax Credit (EITC) reforms on the employment of single mothers in the US, Eissa and Liebman (96) use single women without children as one possible control group. However, this comparison can be criticized for not satisfying the common macro effects assumption (i). In particular, the control group is already working to a very high level of participation in

the US labour market (around 95%) and therefore cannot be expected to increase its level of participation in response to the economy coming out of a recession. In this case all the expansion in labour market participation in the group of single women with children will be attributed to the reform itself. In the light of this criticism the authors also use low education childless single women as a control group for which non-participation is much more common and who have other similar characteristics to those single parents eligible to EITC.

The *matching method* has a long history in non-experimental statistical evaluation (see Heckman, Ichimura and Todd (97), Rosenbaum and Rubin (85) and Rubin (79)). The aim of matching is simple. It is to select sufficient observable factors that any two individuals with the same value of these factors will display no systematic differences in their reaction to the policy reform. Consequently, if each individual undergoing the reform can be matched with an individual that has not undergone the reform and exhibits similar characteristics as measured by the matching variables, the impact on individuals of that type can be measured. It is a matter of prior assumption as to whether the appropriate matching variables have been chosen. If not, the counterfactual effect will not be correctly measured. Again experimental data can help here in evaluating the choice of matching variables and this is precisely the motivation for the Heckman, Ichimura and Todd (97) study. As we document below, matching methods have been extensively refined in the recent evaluation literature and are now a valuable part of the evaluation toolbox.

The *instrumental variable method* is the standard econometric approach to endogeneity. It relies on finding a variable excluded from the outcome equation but which is also a determinant of programme participation. In the simple linear model, the IV estimator identifies the treatment effect removed of all the biases which emanate from a non-randomised control. However, in heterogeneous models, in which the impact of the programme can differ in unobservable ways across participants, the IV estimator will only identify the average treatment effect under strong assumptions and ones that are unlikely to hold in practice. Recent work by Angrist and Imbens (94) and Heckman and Vytlacil (99) has provided an ingenious interpretation of the IV estimator in terms of local treatment effect parameters. We provide a review of these developments.

The distinction between homogeneous and heterogeneous treatments effects that is

highlighted in this recent instrumental variable literature is central to the definition of a ‘parameter of interest’ in the evaluation problem. In the homogeneous linear model there is only one impact of the programme and it is one that would be common to participants and non-participants alike. In the heterogeneous model, those that are treated may have a different mean impact of the programme from those not treated. Certainly this is likely to be the case in a non-experimental evaluation, where participation provides some gain to the participants. In this situation we can define a treatment on the treated parameter that is different from a treatment on the untreated parameter or the average treatment effect. One central issue in understanding evaluation methods is clarifying what type of treatment effect is being recovered by these different approaches.

We should note that fully structural econometric choice models are not discussed in this chapter. These have been the cornerstone of non-experimental evaluation (and simulation) of tax and welfare policies and are extensively discussed and applied in chapters 4 and 5 below. They provide a comprehensive analysis of the choice problem facing individuals deciding on programme participation. They explicitly describe the full constrained maximisation problem and are therefore perfectly suited for ex-ante policy simulation.¹ It should also be noted that the commonly used regression estimator is not directly referred to because it can be included in the matching approach.

The rest of the paper is organized as follows. In the next section we lay out the different definitions of treatment parameters and ask: what are we trying to measure? Section 2 considers the types of data and their implication for the choice of evaluation method. Section 3 is the main focus of this chapter as it presents a detailed comparison of alternative methods of evaluation for non-experimental data. In section 4 we illustrate these methods drawing on recent applications in the evaluation literature. Section 5 concludes.

2.1 Which Parameter of Interest?

We begin by presenting a general model of outcomes which can then assume particular forms depending on the amount of structure one wishes, or needs, to include. There are

¹See Blundell and MaCurdy (1999) for a comprehensive survey and a discussion of the relationship of the structural choice approach to the evaluation approaches presented here.

several important decisions to be taken when specific applications are considered, the one we are especially concerned with is whether the response to the treatment is homogeneous across individuals or heterogeneous. Typically, we do not expect all individuals to be affected by a policy intervention in exactly the same way - there will be heterogeneity in the impact across individuals. Consequently, there are different potential questions that evaluation methods attempt to answer, the most commonly considered being the average effect on individuals of a certain type. This includes a wide range of parameters such as the population average treatment effect (ATE), which would be the outcome if individuals were assigned at random to treatment, the average effect on individuals that were assigned to treatment (TTE), the effect of treatment on agents that are indifferent to participation, which is the marginal version of the local average treatment effect (LATE) discussed below, or the effect of treatment on the untreated (TU) which is typically an interesting measure for decisions about extending some treatment to a group that was formerly excluded from it. Under the homogeneous treatment effect assumption, all these measures are identical, but this is clearly not true when treatment effects depend on individual's characteristics. From now onwards, except if explicitly mentioned, anywhere we discuss heterogeneous treatment effects the analysis pertains the TTE parameter.

To make things more precise, suppose there is a policy reform or intervention at time k for which we want to measure the impact on some outcome variable, Y . This outcome is assumed to depend on a set of exogenous variables, X , the particular relationship being dependent on the participation status in each period t . Let D be a dummy variable representing the treatment status, assuming the value 1 if the agent has been treated and 0 otherwise. The outcome's equations can be generically represented as follows,

$$\begin{aligned} Y_{it}^1 &= g_t^1(X_i) + U_{it}^1 \\ Y_{it}^0 &= g_t^0(X_i) + U_{it}^0 \end{aligned} \tag{2.1}$$

where the superscript stands for the treatment status and the subscripts i and t identify the agent and the time period, respectively. The functions g^0 and g^1 represent the relationship between the potential outcomes (Y^0, Y^1) and the set of observables X and (U^0, U^1) stand for the error terms of mean zero and assumed to be uncorrelated with the regressors X . The X variables are not affected by treatment (or pre-determined) and are assumed known at the moment of deciding about participation. For this reason we have excluded the time

subscript from X . For comparison purposes, this means that agents are grouped by X before the treatment period and remain in the same group throughout the evaluation period. This is a general form of the switching regimes or endogenous selection model.

We assume that the participation decision can be parameterised in the following way: For each individual there is an index, IN , depending on a set of variables W , for which enrolment occurs when this index raises above zero. That is:

$$IN_i = f(W_i) + V_i \quad (2.2)$$

where V_i is the error term, and,

$$\begin{aligned} D_{it} &= 1 && \text{if } IN_i > 0 \text{ and } t > k \\ D_{it} &= 0 && \text{otherwise} \end{aligned} \quad (2.3)$$

Except in the case of experimental data, assignment to treatment is most probably not random. As a consequence, the assignment process is likely to lead to a non-zero correlation between enrolment in the programme - represented by D - and the outcome's error term - (U^0, U^1) . This happens because individuals participation decision is most likely based on personal unobservable characteristics that may well affect the outcome Y as well. If this is so, and if we are unable to control for all the characteristics affecting Y and D simultaneously, then some correlation between the error term and the participation variable is expected. Any method that fails to take such problem into account is not able to identify the true parameter of interest.

Under the above specification, one can define the individual-specific treatment effect to be

$$\alpha_{it}(X_i) = Y_{it}^1 - Y_{it}^0 = [g_t^1(X_i) - g_t^0(X_i)] + [U_{it}^1 - U_{it}^0] \quad \text{with } t > k. \quad (2.4)$$

and the different potential parameters of interest measured in period $t > k$,

$$\begin{aligned} \alpha_t^{ATE} &= E(\alpha_{it} | X = X_i) \\ \alpha_t^{TTE} &= E(\alpha_{it} | X = X_i, D_{it} = 1) \\ \alpha_t^{TU} &= E(\alpha_{it} | X = X_i, D_{it} = 0) \end{aligned}$$

2.1.1 Homogeneous Treatment Effects

The simplest case is when the effect is assumed to be constant across individuals, so that

$$\alpha_t = \alpha_{it}(X_i) = g_t^1(X_i) - g_t^0(X_i) \quad \text{with } t > k$$

for any i . But this means that g^1 and g^0 are two parallel curves, only differing in the level, and the participation-specific error terms are not affected by the treatment status. The outcome's equation (2.1) can therefore be re-written as

$$Y_{it} = g_t^0(X_i) + \alpha_t D_{it} + U_{it}^0 \quad (2.5)$$

2.1.2 Heterogeneous Treatment Effects

However, it seems reasonable to assume that the treatment impact varies across individuals. These differentiated effects may come systematically through the observables' component or be a part of the unobservables. Without loss of generality, the outcome's equation (2.1) can be re-written as follows

$$\begin{aligned} Y_{it} &= D_{it} Y_{it}^1 + (1 - D_{it}) Y_{it}^0 = \\ &= g_t^0(X_i) + \alpha_t(X_i) D_{it} + [U_{it}^0 + D_{it} (U_{it}^1 - U_{it}^0)] \end{aligned} \quad (2.6)$$

where $\alpha_t(X_i)$ is the expected treatment effect at time t among agents characterised by X_i , so that²

$$\alpha_t(X_i) = E[\alpha_{it}(X_i)] = g_t^1(X_i) - g_t^0(X_i) \quad (2.7)$$

Two issues are particularly important under this more general setting. The first relates to the observables and their role in the identification of the parameter of interest. It is clear that the common support problem is central to this setting:³ contrary to the homogeneous treatment effect, this structure does not allow extrapolation to areas of the support of X that are not represented at least among the treated (if a particular parameterisation of g^0 is assumed, one may be able to extrapolate among the non-treated).

²Specification (2.6) obviously includes (2.5) as a special case.

³By *common support* it is meant the subspace of individual' characteristics that is represented both among treated and non-treated.

The second problem concerns the form of the error term, which differs across observations according to the specific treatment status. If there is selection on the unobservables, the OLS estimator after controlling for the covariates X is inconsistent for $\alpha_t(X)$, identifying instead the following parameter,

$$E(\hat{\alpha}_t(X)) = \alpha_t(X) + E(U_t^1 | X, D_t = 1) - E(U_t^0 | X, D_t = 0) \quad \text{with } t > k.$$

2.2 Experimental And Non-Experimental Data

2.2.1 Experimental Data

Under ideal conditions to be discussed below, experimental data provides the correct missing counterfactual, eliminating the evaluation problem. The contribution of experimental data is to rule out bias from self-selection as individuals are randomly assigned to the programme. To see why, imagine an experiment that randomly chooses individuals from a group to participate in a programme - these are administered the treatment. It means that assignment to treatment is completely independent from a possible outcome or the treatment effect. Under the assumption of no spillover (GE) effects, the group of non-treated is statistically equivalent to the treated group in all dimensions except treatment status. The ATE within the experimental population can be simply measured by

$$\hat{\alpha}_t^{ATE} = \bar{Y}_t^1 - \bar{Y}_t^0 \quad t > k \quad (2.8)$$

where $\bar{Y}_t^{(1)}$ and $\bar{Y}_t^{(0)}$ stand for the treated and non-treated average outcomes at a time t after the programme.

However, a number of disrupting factors may interfere with this type of social experiments, invalidating the results. First, we expect some individuals to dropout, and the process is likely to affect treatments and controls unevenly and to occur non-randomly. The importance of the potential non-random selection may be assessed by comparing the observable characteristics of the remaining treatments and controls and treatment groups. Second, given the complexity of the contemporaneous welfare systems, truly committed experimental controls may actively search for alternative programmes and are likely to succeed. Moreover, observed behaviour of the individuals may also change as a consequence

of the experiment itself as, for instance, the officers may try to “compensate” excluded agents by providing them detailed information about other programmes. It is even possible that some controls end up receiving precisely the same treatment being enrolled through other programme. In such case, the parameter being estimated is likely to differ from the parameter of interest by the contamination of the control group. If treatment is defined as actually receiving some kind of “care” or “therapy”, then the effect of treatment will not be identified. If, on the other hand, treatment is understood as being enrolled in the programme, the effect should account for what alternatives exist outside the programme and is more likely to be identified from the data.

2.2.2 Non-Experimental Data

Despite the above comments, non-experimental data is even more difficult to deal with and requires special care. When the control group is drawn from the population at large, even if satisfying strict comparability rules based on observable information, we cannot rule out differences on unobservables that are related to programme participation. This is the *econometric selection problem* as commonly defined (see Heckman, 79). In this case, using the estimator (2.8) results in a fundamental non-identification problem since it approximates (abstracting from other regressors in the outcome equation),

$$E(\hat{\alpha}_{ATE}) = \alpha + [E(U_{it} | d_i = 1) - E(U_{it} | d_i = 0)].$$

Under selection on the unobservables, $E(U_{it}d_i) \neq 0$ and $E(\hat{\alpha}_{ATE})$ is expected to differ from α unless, by chance, the two final r.h.s. terms cancel out. Thus, alternative estimators are needed, which motivates the methods discussed in section 4 below.

2.3 Methods For Non-Experimental Data

The appropriate methodology for non-experimental data depends on three factors: the type of information available to the researcher, the underlying model and the parameter of interest. Data sets with longitudinal or repeated cross-section information support less restrictive estimators due to the relative richness of information. Not surprisingly, there is a clear trade-off between the available information and the restrictions needed to guarantee

a reliable estimator.

This section starts by discussing the Instrumental Variables (IV) estimator, a frequent choice when only a single cross-section is available. IV uses at least one variable that is related with the participation decision but otherwise unrelated with the outcome. Under some conditions, it provides the required randomness in the assignment rule. Thus, the relationship between the instrument and the outcome for different participation groups identifies the impact of treatment avoiding selection problems.

If longitudinal or repeated cross-section data is available, Difference in Differences (DID) can provide a more robust estimate of the impact of the treatment (Heckman and Robb, 85 and 86).⁴ We outline the conditions necessary for DID to reliably estimate the parameter of interest and discuss a possible extension to generalise the common trends assumption.

An alternative approach is the method of matching, which can be adopted with either cross section or longitudinal data although typically detailed individual information from the before the programme period is required (see Heckman, Ichimura and Todd, 97, Rosenbaum and Rubin, 85 and Rubin, 79). Matching deals with the selection process by constructing a comparison group of individuals with observable characteristics similar to the treated. A popular choice that will be discussed uses the probability of participation to perform matching - the so called Propensity Score Matching.

Finally, a joint DID and matching approach may significantly improve the quality of non-experimental evaluation results and is the last estimator discussed.⁵

2.3.1 The Instrumental Variables (IV) Estimator

The IV method requires the existence of, at least, one regressor exclusive to the decision rule, Z , satisfying the two following conditions:⁶

A1: Conditional on X , Z is not correlated with the unobservables (V, U^0) and (V, U^1) .

A2: Conditional on X , the decision rule is a non-trivial (non-constant) function of Z .

⁴This idea is further developed in Blundell, Duncan and Meghir (98) and Bell, Blundell and Van Reenen (99).

⁵This is applied in Blundell, Costa Dias, Meghir and Van Reenen, 01, as will be discussed below.

⁶The time subscript is omitted from the IV analysis since only one time period is under consideration.

Assumption (A2) means that there is independent (from X) variation in Z that affects programme participation, or, in other words, that under a linear specification of the decision rule, the Z coefficient(s) is(are) non-zero. Thus, in general

$$E(D | X, Z) = P(D = 1 | X, Z) \neq P(D = 1 | X)$$

Assumption (A1) means that Z has no impact on the outcomes equation through the unobservable component. The only way Z is allowed to affect the outcomes is through the participation status, D . Under homogeneous treatment effects, this means Z affects the level only, while under heterogeneous treatment effects how much Z affects the outcome depends on the particular values of X . The variable(s) Z is called the instrument(s), and is a source of exogenous variation used to approximate randomised trials: it provides variation that is correlated with the participation decision but does not affect the potential outcomes from treatment directly.

The IV Estimator: Homogeneous Treatment Effect

Under conditions (A1) and (A2), the standard IV procedure identifies the treatment effect α using only the part of the variation in D that is associated with Z ($\hat{\alpha}_{IV} = \text{cov}(y_i, Z_i) / \text{cov}(d_i, Z_i)$). An alternative is to use both Z and X to predict D , building a new variable \hat{D} that is used in the regression instead of D . A third possibility is directly derived by noting that, given assumption (A1) and equation (2.5),

$$E(Y | X, Z) = g^0(X) + \alpha P(D = 1 | X, Z)$$

and since, from assumption (A2), there are at least two values of Z , say z and $z + \delta$ ($\delta \neq 0$), such that $P(D = 1 | X, Z = z) \neq P(D = 1 | X, Z = z + \delta)$,

$$\begin{aligned} \alpha_{IV} &= \frac{\int_{S(X)} [E(Y | X, Z = z) - E(Y | X, Z = z + \delta)] dF(X | X \in S(X))}{\int_{S(X)} [P(D = 1 | X, Z = z) - P(D = 1 | X, Z = z + \delta)] dF(X | X \in S(X))} \\ &= \frac{E(Y | Z = z) - E(Y | Z = z + \delta)}{P(D = 1 | Z = z) - P(D = 1 | Z = z + \delta)} \end{aligned} \quad (2.9)$$

where $S(X)$ stands for the support of X where the probability of participating changes with z , $P(D = 1 | X, Z = z) \neq P(D = 1 | X, Z = z + \delta)$.

The IV Estimator: Heterogeneous Treatment Effects

Depending on the assumptions one is willing to accept, the heterogeneous framework may impose additional requirements on the data for the treatment effect to be identifiable. We start from the simpler case given by the following assumption,

A3: Individuals do not use information on the idiosyncratic component of the treatment effect when deciding about participation ($\alpha_i(X) - \alpha(X)$ where $\alpha(X) = E(\alpha_i|X)$).

Assumption (A3) is satisfied if potential participants have no *a priori* information apart from the one available to the researcher (X) and decision is based on the average treatment effect for the agent's specific group. In such case,

$$E[U_i^1 - U_i^0 | X, Z, D] = E[D_i[\alpha_i(X) - \alpha(X)] | X, Z] = 0$$

which together with (A1) and (A2) is sufficient to identify the average treatment effect $E[\alpha_i | X]$. Furthermore, there is no apparent reason for it to differ from the effect of treatment on the treated, $E[\alpha_i | X, D_i = 1]$ for as long as the estimated parameters are conditional on the observables, X .

If, however, agents are aware of their own idiosyncratic gains from treatment, they are likely to make a more informed participation decision. Selection on the unobservables is expected, making individuals that benefit more from participation to be the most likely to participate within each X -group. Such a selection process creates correlation between $\alpha_i(X)$ and Z . This is easily understood given that the instrument impacts on D , facilitating or inhibiting participation. For example, it may be that participants with values of Z that make participation more unlikely are expected to gain on average more from treatment than participants with values of Z that make participation more likely to occur. Take the case where distance from home to the treatment location is taken as an instrument. Though in general such a variable is unlikely to be related with outcomes such as earnings or employment probabilities, it is likely to be related with the idiosyncratic component of the treatment effect since agents living closer incur less travelling costs and are, therefore, more likely to participate even if expecting lower gains from treatment. Such a relationship between the instrument Z and the idiosyncratic gain from treatment

is immediately recognised formally since, from (2.1)

$$U_i^1 - U_i^0 = (Y_i^1 - Y_i^0) - \alpha(X_i) = \alpha_i(X_i) - \alpha(X_i)$$

Thus, the error term under heterogeneous treatment effect is

$$U_i = U_i^0 + D_i [\alpha_i(X_i) - \alpha(X_i)]$$

where D is, by assumption, determined by Z depending on the gain $\alpha_i(X_i) - \alpha(X_i)$.

Under such circumstances, assumptions (A1) and (A2) are no longer enough to identify the ATE or TTE. This happens because the average outcomes of any two groups differing on the particular Z -realisations alone are different not only as a consequence of different participation rates but also because of compositional differences in the participants (non-participants) groups according to the unobservables. Thus, the main concern relates to the existence and identification of regions of the support of X and Z where changes in Z cause changes in the participation rates unrelated with potential gains from treatment.

The solution advanced by Imbens and Angrist (94) is to use IV locally, for particular changes of the instrument Z . The rationale is that some local changes in the instrument Z reproduce random assignment by inducing agents to decide differently as they face different conditions unrelated to potential outcomes. To guarantee that the groups being compared are indeed comparable, Imbens and Angrist use a strengthened version of (A2),

A2': Conditional on X , the decision rule is a non-trivial monotonic function of Z .

In what follows, suppose D is an increasing function of Z , meaning that an increase in Z leads some individuals to take up treatment but no one individual to give up treatment. In an hypothetical case, where Z changes from $Z = z$ to $Z = z + \delta$ ($\delta > 0$), the individuals that change their participation decisions as a consequence of the change in Z are those that choose not to participate under $Z = z$ excluding the ones that choose not to participate under $Z = z + \delta$, or, equivalently, those that decide to participate under $Z = z + \delta$ excluding the ones that prefer participation under $Z = z$. Thus, the expected outcome under treatment and non-treatment for those affected by the change in Z can be estimated

as follows,

$$\begin{aligned} E [Y_i^1 | X_i, D_i(z) = 0, D_i(z + \delta) = 1] &= \\ &= \frac{\left\{ \begin{array}{l} E [Y_i^1 | X_i, D_i(z + \delta) = 1] P [D_i = 1 | X_i, z + \delta] - \\ E [Y_i^1 | X_i, D_i(z) = 1] P [D_i = 1 | X_i, z] \end{array} \right\}}{P [D_i = 1 | X_i, z + \delta] - P [D_i = 1 | X_i, z]} \end{aligned}$$

and

$$\begin{aligned} E [Y_i^0 | X_i, D_i(z) = 0, D_i(z + \delta) = 1] &= \\ &= \frac{\left\{ \begin{array}{l} E [Y_i^0 | X_i, D_i(z) = 0] P [D_i = 0 | X_i, z] - \\ E [Y_i^0 | X_i, D_i(z + \delta) = 0] P [D_i = 0 | X_i, z + \delta] \end{array} \right\}}{P [D_i = 1 | X_i, z + \delta] - P [D_i = 1 | X_i, z]} \end{aligned}$$

The estimated treatment effect is given by,

$$\begin{aligned} \alpha_{LATE}(X_i, z, z + \delta) &= E(Y_i^1 - Y_i^0 | X_i, D_i(z) = 0, D_i(z + \delta) = 1) \quad (2.10) \\ &= \frac{E [Y_i | X_i, z + \delta] - E [Y_i | X_i, z]}{P [D_i = 1 | X_i, z + \delta] - P [D_i = 1 | X_i, z]} \end{aligned}$$

which is the Local Average Treatment Effect (LATE) parameter. To illustrate the LATE approach, take the example discussed above on selection into treatment dependent on the distance to the treatment site. Participation is assumed to become less likely the longest the distance from home to the treatment location. To estimate the treatment effect, consider a group of individuals that differ only on the distance dimension. Among those that participate when the distance Z equals z some would stop participating if at distance $z + \delta$. LATE measures the impact of the treatment on the “movers” group by attributing any difference on the average outcomes of the two groups defined by the distance to the treatment site to the different participation frequency.⁷

The LATE parameter uses the IV estimator applied to some specific values of Z and is, therefore, different from TTE or ATE. It is intrinsically dependent on the particular values

⁷Abadie, Angrist and Imbens (1998) extend this approach to the evaluation of *quantile treatment effects*. The goal is to assess how different parts of the outcome’s distribution are affected by the policy. As with LATE, a local IV procedure is used, making the estimated impacts representative only for the sub-population of individuals changing their treatment status as a consequence of the particular change in the instrument considered.

of Z used to evaluate the treatment and on the particular instrument chosen. The group of “movers” is not in general representative of the whole treated or, even less, the whole population. For instance, agents benefiting the most from participation are more unlikely to be observed among the movers. The LATE parameter answers a different question, of how much agents at the margin of participating benefit from participation given a change in policy. That is, it measures the effect of treatment on the sub-group of treated at the margin of participating for a given $Z = z$. This is more easily seen if taking the limits when $\delta \rightarrow 0$, as in Heckman and Vytlacil (99),

$$\alpha_{MTE}(X_i, z) = \frac{\partial E[Y | X_i, Z]}{\partial P[D = 1 | X_i, Z]} \Big|_{Z=z}$$

α_{MTE} is the Marginal Treatment Effect (MTE), and is by definition the LATE parameter defined for an infinitesimal change in Z . It represents TTE for agents that are indifferent between participating and not participating at $Z = z$. All the three parameters, namely ATE, TTE and LATE, can be expressed as averages of MTE over different subsets of the Z support. The ATE is the expected value of MTE over the entire support of Z , including the values where participation is nil or universal. The TTE excludes only the subset of the Z -support where participation does not occur. Finally, LATE is defined as the average MTE over an interval of Z bounded by two values for which participation rates are different.⁸

2.3.2 The Difference In Differences (DID) Estimator

If longitudinal or repeated cross-section information is available, the additional time dimension can be used to estimate the treatment effect under less restrictive assumptions. Without loss of generality, re-write model (2.6) as follows,

$$Y_{it} = g_t^0(X_i) + \alpha_{it}(X_i) D_{it} + (\phi_i + \theta_t + \varepsilon_{it}) \quad (2.11)$$

where the error term U_{it}^0 , is being decomposed on an *individual-specific fixed effect*, ϕ_i , a *common macro-economic effect*, θ_t and a *temporary individual-specific effect*, ε_{it} . The main assumption underlying the DID estimator is the following,

⁸The importance of the monotonic assumption depends on the parameter of interest. It is not needed if one is willing to assess the effects of a change in policy on average outcomes, which includes both changes in participation and effects of participation (see Heckman, 97).

A4: Selection into treatment is independent of the temporary individual-specific effect, ε_{it} , so that,

$$E(U_{it}^0 | X_i, D_i) = E(\phi_i | X_i, D_i) + \theta_t$$

where D_i distinguishes participants from non-participants and is, therefore, time-independent (that is, $D_i = D_{it}$ for any $t > k$).

Assumption (A4) is sufficient because ϕ_i and θ_t vanish in the sequential differences.⁹ To see why, suppose information is available for a pre- and a post-programme periods - denoted respectively by t_0 and t_1 ($t_0 < k < t_1$). DID measures the excess outcome growth for the treated compared to the non-treated. Formally, it can be presented as follows,

$$\hat{\alpha}_{DID}(X) = [\bar{Y}_{t_1}^1(X) - \bar{Y}_{t_0}^1(X)] - [\bar{Y}_{t_1}^0(X) - \bar{Y}_{t_0}^0(X)] \quad (2.12)$$

where \bar{Y} stands for the mean outcome among the specific group being considered. Under heterogeneous effects, the DID estimator recovers the TTE since¹⁰

$$E(\hat{\alpha}_{DID}(X)) = E[\alpha_i(X) | D_i = 1] = \alpha_{TTE}(X)$$

In the homogeneous effect case, one may omit the covariates from equation (2.12) and average over the complete groups of treated and non-treated. The obtained estimate is consistent for α .

The DID Estimator: The Common Trends And Time Invariant Composition Assumptions

In contrast to the IV estimator, no exclusion restrictions are required under the DID methodology as there is no need for any regressor in the decision rule. Even the outcome equation may remain unspecified as long as the treatment impact enters additively.

⁹Notice that selection is allowed to occur on a temporary individual-specific effect that depends on the observables only, namely $g_t^0(X_i)$.

¹⁰It should be noticed that in a more general case where the intervention runs over time, DID will not generally identify the TTE parameter. In such case, DID is usually applied to periods where changes in policies occur and the identified parameter will be the impact of the intervention on “new” participants in the “new” regime.

However, assumption (A4) together with the postulated specification (2.11) brings two main weaknesses to the DID approach. The first problem relates to the lack of control for *unobserved temporary individual-specific components* that influence the participation decision. If ε is not unrelated to D , DID is inconsistent and in fact approximates the following parameter,

$$E(\hat{\alpha}_{DID}(X)) = \alpha_{TTE}(X) + E(\varepsilon_{it_1} - \varepsilon_{it_0} \mid D_i = 1) - E(\varepsilon_{it_1} - \varepsilon_{it_0} \mid D_i = 0)$$

To illustrate the conditions such inconsistency might arise, suppose a training programme is being evaluated in which enrolment is more likely if a temporary dip in earnings occurs just before the programme takes place (so-called Ashenfelter's dip, see Heckman and Smith, 97). A faster earnings growth is expected among the treated, even without programme participation. Thus, the DID estimator is likely to over-estimate the impact of treatment. Moreover, if instead of longitudinal data one uses cross-section data, the problem is likely to worsen as it may extend to the fixed effect (ϕ_i) component: the before-after comparability of the groups under an unknown selection rule may be severely affected as the composition of the groups may change over time, particularly due to the intervention, causing $E(\phi_i \mid D_i)$ to change artificially with t .

The second weakness occurs if the macro effect has a differential impact across the two groups. This happens when the treatment and comparison groups have some (possibly unknown) characteristics that distinguish them and make them react differently to common macro shocks or, alternatively, when the macro shocks are not common (different labour markets). Such issue motivates the differential trend adjusted DID estimator that is presented below.

The DID Estimator: Adjusting For Differential Trends

Replace (A4) by

A4' Selection into treatment is independent of the temporary individual-specific effect, ε_{it} , under differential trends

$$E(U_{it} \mid D_i) = E(\phi_i \mid D_i) + k^D \theta_t$$

where the k^D acknowledges the differential macro effect across the two groups.

The DID estimator now identifies

$$E(\hat{\alpha}_{DID}(X)) = \alpha_{TTE}(X) + (k^1 - k^0)[\theta_{t_1} - \theta_{t_0}] \quad (2.13)$$

which clearly only recovers the true TTE when $k^1 = k^0$.

Now suppose we take another time interval, say t_* to t_{**} (with $t_* < t_{**} < k$), over which a similar macro trend has occurred. Precisely, we require a period for which the macro trend matches the term $(k^1 - k^0)[\theta_{t_1} - \theta_{t_0}]$ in (2.13). It is likely that the most recent cycle is the most appropriate, as earlier cycles may have systematically different effects across the target and comparison groups. The differentially adjusted estimator proposed by Bell, Blundell and Van Reenen (99) takes the form

$$\hat{\alpha}_{TADID}(X) = \left\{ (\tilde{Y}_{t_1}^1 - \tilde{Y}_{t_0}^1) - (\tilde{Y}_{t_1}^0 - \tilde{Y}_{t_0}^0) \right\} - \left\{ (\tilde{Y}_{t_{**}}^1 - \tilde{Y}_{t_*}^1) - (\tilde{Y}_{t_{**}}^0 - \tilde{Y}_{t_*}^0) \right\} \quad (2.14)$$

and will now consistently estimate α_{TTE} .

To illustrate this approach, let's consider the case where treatments and controls belong to different cohorts. Suppose treatments are drawn from a younger cohort, making them more responsive to macroeconomic cycles. If the outcome of interest is affected by the macro conditions, we expect the time-specific effect to differ between treatments and controls. But if similar cyclical conditions were observed in the past and the response of the two groups has been kept unchanged, it is possible to find a past period characterised by the same differential, $\theta_{t_1} - \theta_{t_0}$.

2.3.3 The Matching Estimator

The third method we present is the matching approach. Like the DID, matching does not require an exclusion restriction or a particular specification of the participation decision or the outcomes equation. It also does not require the additive specification of the error term as postulated for the DID estimator. Its additional generality comes from being a non-parametric method, which also makes it quite versatile in the sense that it can easily be combined with other methods to produce more accurate estimates. The cost is paid with data: matching requires abundant good quality data to be at all meaningful.

The main purpose of matching is to re-establish the conditions of an experiment when no randomised control group is available. As we have noted, total random assignment

allows for a direct comparison of the treated and non-treated, without particular structure requirements. The matching method aims to construct *the* correct sample counterpart for the missing information on the treated outcomes had they not been treated by pairing each participant with members of non-treated group. Under the matching assumption, the only remaining difference between the two groups is programme participation.

The solution advanced by matching is based on the following assumption,

A5: *Conditional independence assumption (CIA)*: conditional on the set of observables X , the non-treated outcomes are independent of the participation status,

$$Y^0 \perp D \mid X$$

That is, the non-treated outcomes are what the treated outcomes would have been had they not been treated conditional on X , In other words, *selection occurs only on observables*. For each treated observation (Y^1) we can look for a non-treated (set of) observation(s) (Y^0) with the same X -realisation. With the matching assumption, this Y^0 constitutes the required counterfactual. Thus, matching is a process of re-building an experimental data set.

The second assumption guarantees that the required counterfactual actually exists

A6: All treated agents have a counterpart on the non-treated population and anyone constitutes a possible participant:

$$0 < P(D = 1 \mid X) < 1$$

However, assumption (A6) does not ensure that the same happens within any sample, and is, in fact, a strong assumption when programmes are directed to tightly specified groups.

Call S^* to the common support of X . Assuming (A5) and (A6), a subset of comparable observations is formed from the original sample and a consistent estimator for TTE is produced using the empirical counterpart of

$$\frac{\int_{S^*} E(Y^1 - Y^0 \mid X, D = 1) dF(X \mid D = 1)}{\int_{S^*} dF(X \mid D = 1)} \quad \text{at a time } t > k \quad (2.15)$$

where the denominator represents the expected gain from the programme among the subset of sampled participants for whom one can find a comparable non-participant (that is, over S^*). To obtain a measure of the TTE, individual gains must be integrated over the distribution of observables among participants and re-scaled by the measure of the common support, S^* . Therefore, equation (2.15) represents the expected value of the programme effect over S^* .¹¹ If (A5) is fulfilled and the two populations are large enough, the common support is the entire support of both.

The challenge of matching is to ensure that the ‘correct’ set of observables X is being used so that the observations of non-participants are what the observations of treated would be had they not participated, forming the right counterfactual and satisfying the CIA. In practical terms, however, the more detailed the information is, the harder it is to find a similar control and the more restricted the common support becomes. That is, the appropriate trade-off between the quantity of information at use and the share of the support covered may be difficult to achieve. If, however, the right amount of information is used, matching deals well with potential bias. This is made clear by decomposing the treatment effect in the following way

$$E(Y^1 - Y^0 | X, D = 1) = \{E(Y^1 | X, D = 1) - E(Y^0 | X, D = 0)\} - \\ - \{E(Y^0 | X, D = 1) - E(Y^0 | X, D = 0)\}$$

where the latter term is the bias conditional on X . Conditional on X , the only reason the true parameter, $\alpha_{TTE}(X)$, might not be identified is selection on the unobservables.

Note, however, if one integrates over the common support S^* , two additional causes of bias can occur: non-overlapping support of X and misweighting over the common support. Through the process of choosing and re-weighting observations, matching corrects for the latter two sources of bias and selection on the unobservables is assumed to be zero.

The Matching Estimator: The Use Of Propensity Score

As with all non-parametric methods, the dimensionality of the problem as measured by X may seriously limit the use of matching. A more feasible alternative is to match on a

¹¹It is simply the mean difference in outcomes over the common support, appropriately weighted by the distribution of participants.

function of X . Usually, this is carried out on the propensity to participate given the set of characteristics X : $P(X_i) = P(D_i = 1 | X_i)$ the *propensity score*. Its use is usually motivated by Rosenbaum and Rubin's result (83, 84), which shows that the CIA remains valid if controlling for $P(X)$ instead of X :

$$Y^0 \perp D \mid P(X)$$

More recently, a study by Hahn (98) shows that $P(X)$ is ancillary for the estimation of ATE. However, it is also shown that knowledge of $P(X)$ may improve the efficiency of the estimates of TTE, its value lying on the "dimension reduction" feature.

When using $P(X)$, the comparison group for each treated individual is chosen with a pre-defined criteria (established by a pre-defined measure) of proximity. Having defined the neighbourhood for each treated observation, the next issue is that of choosing the appropriate weights to associate the selected set of non-treated observations for each participant one. Several possibilities are commonly used, from a unity weight to the nearest observation and zero to the others, to equal weights to all, or kernel weights, that account for the relative proximity of the non-participants' observations to the treated ones in terms of $P(X)$.

In general the form of the matching estimator is given by

$$\hat{\alpha}_M = \sum_{i \in T} \left\{ Y_i - \sum_{j \in C} W_{ij} Y_j \right\} w_i \quad (2.16)$$

where T and C represent the treatment and comparison groups respectively, W_{ij} is the weight placed on comparison observation j for individual i and w_i accounts for the re-weighting that reconstructs the outcome distribution for the treated sample.¹²

The Matching Estimator: Parametric Approach

Specific functional forms assumed for the g -functions in (2.1) can be used to estimate the impact of treatment on the treated over the whole support of X , reflecting the trade-off

¹²For example, in the nearest neighbour matching case the estimator becomes

$$\hat{\alpha}_{MM} = \sum_{i \in T} \{Y_i - Y_j\} \frac{1}{N_T}$$

where, among the non-treated, j is the nearest neighbour to i in terms of $P(X)$. In general, kernel weights are used for W_{ij} to account for the closeness of Y_j to Y_i .

between the structure one is willing to impose in the model and the amount of information that can be extracted from the data. To estimate the impact of treatment under a parametric set-up, one needs to estimate the relationship between the observables and the outcome for the treatment and comparison groups and predict the respective outcomes for the population of interest. A comparison between the two sets of predictions supplies an estimate of the impact of the programme. In this case, one can easily guarantee that outcomes being compared come from populations sharing exactly the same characteristics.¹³

2.3.4 Matching and DID

The CIA is quite strong if admitted that individuals decide according to their outcomes' forecast. However, by combining matching with DID there is scope for an unobserved determinant of participation as long as it lies on separable individual and/or time-specific components of the error term. To clarify the exposition, let's take model (2.11).¹⁴ If performing matching on the set of observables X within this setting, the CIA can now be replaced by,

$$(\varepsilon_{t_1} - \varepsilon_{t_0}) \perp D \mid X$$

where $t_0 < k < t_1$. Since DID effectively controls for the other components of the outcomes under non-treatment, only the temporary individual-specific shock requires additional control. The main matching hypothesis is now stated in terms of the before-after evolution instead of levels. It means that controls have evolved from a pre- to a post-programme period in the same way treatments would have done had they not been treated.

The effect of the treatment on the treated can now be estimated over the common

¹³If, for instance, a linear specification is assumed with common coefficients for treatments and controls, so that

$$Y = X\beta + \alpha_{TTE}D + U$$

then no common support requirement is needed to estimate α_{TTE} - a simple OLS regression using all information on treated and non-treated will consistently identify it.

¹⁴An extension to consider differential trends can be considered similarly to what have been discussed before.

support of X, S^* , using an extension to (2.16),

$$\hat{\alpha}_{MDD}^{LD} = \sum_{i \in T} \left\{ [Y_{it_1} - Y_{it_0}] - \sum_{j \in C} W_{ij} [Y_{jt_1} - Y_{jt_0}] \right\} w_i$$

where the notation is similar to what has been used before.

Quite obviously, this estimator requires longitudinal data to be applied. However, it is possible to extend it for the repeated cross-sections data case. If only repeated cross-sections are available, one must perform matching three times for each treated individual after treatment to find the comparable treated before the programme and the controls before and after the programme. If the same assumptions apply, the TTE is identified by,

$$\hat{\alpha}_{MDD}^{RCS} = \sum_{i \in T_1} \left\{ \left[Y_{it_1} - \sum_{j \in T_0} W_{ijt_0}^T Y_{jt_0} \right] - \left[\sum_{j \in C_1} W_{ijt_1}^C Y_{jt_1} - \sum_{j \in C_1} W_{ijt_0}^C Y_{jt_0} \right] \right\} w_i$$

where T_0, T_1, C_0 and C_1 stand for the treatment and comparison groups before and after the programme, respectively, and W_{ijt}^G represent the weights attributed to individual j in group D (where $G = C$ or T) and time t when comparing with treated individual i (for a more detailed discussion with application of the combined matching and DID estimator, see Blundell, Costa Dias, Meghir and Van Reenen, 01).

2.4 Interpreting the Evidence

In this section we briefly draw on some recent studies to illustrate some of the non-experimental techniques presented in this review. The studies presented below show that the methods we have described should be carefully applied and even more cautiously interpreted (see also Blundell and Costa Dias, 00).

2.4.1 The LaLonde Study and the NSW Evaluation

LaLonde (86) aimed at assessing the reliability of the non-experimental techniques by comparing the results produced by these methods as commonly applied and the true parameters obtained using experimental data. This study used the National Supported Work Demonstration (NSWD), a programme operated in 10 sites across USA and designed to help disadvantaged workers, in particular women in receipt of AFDC (Aid for

Families with Dependent Children), ex-drug addicts, ex-criminal offenders and high-school drop-outs. Qualified applicants were randomly assigned to treatment, which comprised a guaranteed job for 9 to 18 months. Treatment and control groups summed up to 6,616 individuals. Data on all participants were collected before, during and after treatment takes place, and earnings were the chosen outcome measure.

To assess the reliability of the experimental design, LaLonde presents pre-treatment earnings and other demographic variables for male treatments and controls (see table 2.1). As far as can be inferred from the observables, treatments and controls are not different before the treatment takes place. In the absence of non-random drop-outs, no alternative treatment being offered and no changes in behaviour induced by experiment, the controls constitute the perfect counterfactual to estimate the treatment impact.

Table 2.1: Observable characteristics for NSW D males.

| Comparison between treatments and controls. | | |
|---|------------|----------|
| Variable | Treatments | Controls |
| Age | 24.49 | 23.99 |
| Years of school | 10.17 | 10.17 |
| Proportion high school dropouts | 0.79 | 0.80 |
| Proportion married | 0.14 | 0.13 |
| Proportion black | 0.76 | 0.75 |
| Proportion hispanic | 0.12 | 0.14 |
| Real earnings 1 year before treatment | 1472 | 1558 |
| Real earnings 2 year before treatment | 2860 | 3030 |
| Hours worked 1 year before treatment | 278 | 274 |
| Hours worked 2 year before treatment | 458 | 469 |
| Number of observations | 2083 | 2193 |

An analysis of the earnings evolution for treated and controls from a pre-programme year, 1975, through the treatment periods, 1976-77, until the post-programme period, 1978, is presented in table 2.2. It can be seen that the treatments' and controls' earnings were nearly the same before treatment. They then diverged substantially during the pro-

Table 2.2: Annual earnings for NSW males.

| Comparison between treatments and controls. | | |
|---|------------|----------|
| Year | Treatments | Controls |
| 1975 | 3,066 | 3,027 |
| 1976 | 4,035 | 2,121 |
| 1977 | 6,335 | 3,403 |
| 1978 | 5,976 | 5,090 |
| Number of observations | 297 | 425 |

gramme and somehow converged after it. The estimated impact one year after treatment is almost + \$900.

Table 2.3: Effects of treatment on the treated for the NSW males.

Estimates using the control group and comparison groups from the PSID and the CPS-SSA.

| Comparison group | Difference of mean | | Difference | | Two-step estimator |
|------------------|-------------------------|----------|----------------|----------|--------------------|
| | post-programme earnings | | in differences | | |
| | Unadjusted | Adjusted | Unadjusted | Adjusted | |
| (1) Controls | 886 | 798 | 847 | 856 | 889 |
| (2) PSID 1 | -15,578 | -8,067 | 425 | -749 | -667 |
| (3) PSID 2 | -4,020 | -3,482 | 484 | -650 | - |
| (4) PSID 3 | 697 | -509 | 242 | -1325 | - |
| (5) CPS-SSA 1 | -8,870 | -4,416 | 1,714 | 195 | 213 |
| (6) CPS-SSA 2 | -4,095 | -1,675 | 226 | -488 | - |
| (7) CPS-SSA 3 | -1,300 | 224 | -1,637 | -1,388 | - |

Notes: PSID 1- All male household heads continuously on the studied period (75 to 78), who were less than 55 years old and did not classify themselves as retired in 75. PSID 2 - Selects from PSID 1 group all men not working when surveyed in the spring of 76. PSID 3 - Selects from PSID 1 group all men not working when surveyed in either spring of 75 or 76. CPS-SSA 1 - All males based on Westat's criteria except those over 55 years old (the Westat's criteria selects individuals that were in the labour force in March 1976 with nominal income less than \$20,000 and household income less than \$30,000). CPS-SSA 2 - Selects from CPS-SSA 1 all males who were not working when surveyed in March 76. CPS-SSA 3 - Selects from CPS-SSA 1 all unemployed in 76 whose income in 75 was below the poverty level.

To evaluate the quality of the non-experimental techniques, LaLonde applied a set of different methods using both the control group and a number of other, non-experimentally determined, comparison groups. The choice of the comparison group is determinant, the aim being to reproduce what the participants would have been in the absence of the programme. The comparison groups were drawn from either the Panel Study of Income Dynamics (PSID) or from the Current Population Survey - Social Security Administration (CPS-SSA). Table 2.3 summarises LaLonde's main results. The first row reveals the robustness of the experimental results to the choice of estimator. Rows 2 to 7, however, show that by using comparisons from non-experimental samples significantly changes the results. Moreover, strong dependence on the adopted specification for the earnings function and participation decision is found when non-experimental data is being used.

2.4.2 A Critique of the LaLonde Study

LaLonde's results have been criticised on the basis that the chosen non-experimental comparison groups do not satisfy the necessary requests to successfully identify the correct parameter (see Heckman, Ichimura and Todd, 97a, and Heckman, Ichimura, Smith and Todd, 98). It is argued that to ask for identification of the true parameter from the data used in LaLonde (86) to construct the counterfactual is to make unfair requests on data that has not been selected to truly represent what the treated would have been without treatment. Three main reasons are pointed out: First, comparisons are not drawn from the same local labour markets; Second, data on treated and comparisons were collected from different questionnaires and do not, therefore, measure the same characteristics; and Third, data are not rich enough to clearly distinguish between individuals. Moreover, it has been argued that the observed differences in the estimates produced by different non-experimental methods does not necessarily signal their inadequacy to identify the true effect or the relative poor quality or paucity of the data (Heckman, LaLonde and Smith, 99). On the contrary, it results from the fact that each method relies on different identification assumptions and estimates different parameters. Such differences are expected in the absence of homogeneous treatment effects.

A recent study by Smith and Todd (00) is based on precisely the same data used by LaLonde. A careful evaluation of the bias present in non-experimental studies is performed

Table 2.4: Bias associated with alternative estimates of the TTE among NSW males. Comparison groups constructed from cross-sectional and longitudinal data from the PSID and the CPS.

| | Mean difference | Nearest neighbour | | Local linear matching | |
|-----------------------------|----------------------|-------------------|--------------------|-----------------------|----------------------------|
| | | w/ common support | w/o common support | simple (bw=1) | regression adjusted (bw=1) |
| PSID: cross-section results | | | | | |
| Old variables | -16,464 (-1,858%) | -3,878 (-438%) | -3,838 (-433%) | -3,689 (-416%) | -3,708 (-419%) |
| New variables | -16,676 (-1,882%) | -2,932 (-331%) | -166 (-19%) | -1,237 (-140%) | -587 (-66%) |
| PSID: DID results | | | | | |
| Old variables | -427 (-48%) | -381 (-43%) | -364 (-41%) | -204 (-23%) | -204 (-23%) |
| New variables | -383 (-43%) | -1644 (-186%) | 608 (69%) | 188 (21%) | -344 (-39%) |
| CPS: cross-section results | | | | | |
| Old variables | -10,227 (-1,154%) | -3,602 (-406%) | -3,586 (-405%) | -3,562 (-402%) | -3,435 (-388%) |
| New variables | -9,757 (-1,101%) | -555 (-63%) | -838 (-95%) | -1,380 (-156%) | -1,406 (-159%) |
| CPS: DID results | | | | | |
| Old variables | 897 (101%) | -463 (-52%) | -21 (-2%) | 192 (22%) | -145 (-16%) |
| New variables | 867 (98%) | -1,527 (-172%) | -929 (-105%) | -1212 (-137%) | -1212 (-137%) |

Notes: All estimators attempt to reproduce the experimental effect, \$886 (table 2.3). Old variables are used in LaLonde's study and include age, age squared, years of schooling, dummies for high school drop outs, black, hispanic married and employed in 1976, number of children and whether there is information on children at all. The new set of variables includes age, age squared, age cubed, years of schooling, years of schooling squared, dummies from high school drop outs, married, black, hispanic, zero earnings in 1974, zero earnings in 1975 and interception with hispanic, real earnings in 1974, real earnings in 1974 squared, real earnings in 1975, real earnings in 1975 squared and the interception between schooling and real earnings in 1974.

Nearest neighbour matching uses a single comparison to match each treatment, the one closest to the treated according to some measure of distance. Since propensity score matching is being used, the distance is taken with respect to $P(X)$. When the common support restriction is used, a maximum acceptable distance is established *a priori* and only treatment observations with close enough comparisons are used in the analysis.

Regression adjusted matching is performed on the residuals of a regression of the outcome of interest on a number of selected independent variables. This method matches both parametrically on a set of variables (the explanatory variables on the variable of interest's regression) and non-parametrically on the remaining variables (the ones determining the propensity score).

Local linear matching uses a weighted average of the outcomes for all comparisons as a counterfactual to each treatment observation. Let $W_{N_C, N_T}(i, j)$ be the weight for the comparison j when matching with the treated i and the numbers of comparisons and treatments are N_C and N_T , respectively; G_{ij} is a kernel function, $G_{ik} = G((X_i - X_k)/a_{N_C})$; a_{N_C} be the bandwidth and I_C be the sample of comparisons. The weights can now be defined as follows,

$$W_{N_C, N_T}(i, j) = \frac{G_{ij} \sum_{k \in I_C} G_{ik}(X_k - X_i)^2 - G_{ij}(X_j - X_i) \sum_{k \in I_C} G_{ik}(X_k - X_i)}{\sum_{l \in I_C} G_{il} \sum_{k \in I_C} G_{ik}(X_k - X_i)^2 - [\sum_{k \in I_C} G_{ik}(X_k - X_i)]^2}$$

More details in Fan, 1992.

by using a variety of methodologies and experimenting with the data. They use LaLonde's outcome variable, earnings, and directly compare non-experimental comparisons with ex-

perimental controls to obtain a measure of the bias. Table 2.4 presents the main results from this study.

This exercise suggests that matching may substantially improve the results when only cross-section data is available, in which case a careful choice of the matching variables is determinant for the quality of the estimates. When using longitudinal data, however, other demands on information are somewhat relaxed. The quality of the estimates improves significantly and to the same order of magnitude independently of the technique or amount of information used. It seems as if much of the information fundamental in a cross-sectional analysis pertains variables that stay relatively constant over time and that cancel out on the sequential differencing that characterises DID. That is, the third issue raised about the LaLonde's study seems to be particularly important when applied to cross-sectional studies.

Evidence on the relative importance of the other two criticism raised about the LaLonde study is provided by the study by Heckman, Ichimura and Todd (97a). The authors explore the source of bias in evaluation studies under different assumptions on the richness of available data. Information was gathered under the Job Training Partnership Act (JTPA), the main US government training programme for disadvantaged workers. It provides on-the-job training, job search assistance and classroom training to youth and adults. Eligibility is determined by a family income near or below the poverty level for six months prior to application or by participation in federal, state or local welfare and food stamp programmes. As in LaLonde's study, earnings are the outcome measure. Data resources, however, are richer than the ones available for the NSW experiment. Detailed longitudinal information was collected under an experimental setting for a group of treatments, randomised-out controls and eligible non-participants (see Devine and Heckman (96), Kemple, Doolittle and Wallace (93) and Orr et. al. (94)). All the three groups were resident in the same narrowly-defined geographic regions and were administered the same questionnaire. The relative richness of information also allowed the construction of close comparison groups from other surveys, providing the means for a formal analysis of estimated bias.

Total bias can be decomposed in three parts, denoted by B_1 , B_2 and B_3 , respectively: the bias due to non-overlapping support of X , the bias due to misweighting on the common

support of X , and the bias resulting from selection on unobservables. One can get an idea of the importance of B_1 by plotting the densities of $P(X)$, the propensity score, by treatment status. Heckman et al. (97a) do this and conclude that the common support defined by the propensity to participate is very restricted. This means that the potential non-experimental comparison group, composed of the eligible non-participants, does not reproduce the characteristics of the treated as represented by the experimental control group. Therefore, a significant source of bias when dealing with non-experimental data should come from non-overlapping support uncontrolled for. Under such circumstances and given heterogeneous treatment effects, it is obvious that non-experimental evaluations identify a different parameter from experimental ones. This is because a significant part of the treatment group X -support is discarded in order to avoid the ‘non-overlapping support’ type of bias.

Table 2.5: Bias decomposition of the simple differences estimator.

| Estimates for the JTPA study on the post-programme earnings. | | | | | | |
|--|---------------------------------|--------------------------------|-------------------------------------|----------------------------------|---------------------------------------|---|
| | Mean difference \hat{B} | Non- overlap \hat{B}_1 | Density weighting \hat{B}_2 | Selection bias \hat{B}_3 | Average bias \hat{B}_{common} | \hat{B}_{common} as % of treatment impact |
| <i>Experimental controls and eligible non-participants</i> | | | | | | |
| Adult males | -342 | 218 | -584 | 23 | 38 | 87% |
| Adult females | 33 | 80 | -78 | 31 | 38 | 129% |
| Male youth | 20 | 142 | -131 | 9 | 14 | 23% |
| Female youth | 42 | 74 | -67 | 35 | 49 | 7,239% |
| <i>Experimental controls and SIPP eligibles</i> | | | | | | |
| Adult males | -145 | 151 | -417 | 121 | 192 | 440% |
| Adult females | 47 | 97 | -172 | 122 | 198 | 676% |
| Male youth | -188 | 65 | -263 | 9 | 21 | 36% |
| Female youth | -88 | 83 | -168 | -3 | -13 | 1,969% |
| <i>Experimental controls and no-shows</i> | | | | | | |
| Adult males | 29 | -13 | 3 | 38 | 42 | 97% |
| Adult females | 9 | 1 | -9 | 18 | 20 | 68% |
| Male youth | 84 | 14 | -21 | 91 | 99 | 171% |
| Female youth | 18 | 3 | -31 | 46 | 51 | 7,441% |

Table 2.5 presents the empirical decomposition of the evaluation bias (see also Table 2 of Heckman et al., 97). Estimates of the bias are obtained from the comparison of

the experimental controls with three potential comparison groups: matched eligible non-participants, a group based on a different survey (SIPP) and a group of no-shows which includes controls and persons assigned to treatment that dropped out before receiving any service. The estimated bias results from a simple difference estimator of treatment impact.

It is clear that B_1 and B_2 account for the majority of the error in any case. Nonetheless, selection on unobservables (B_3) is a significant error as compared to the treatment impact, and is even greater when evaluating the bias on the common support (B_{common}). Another relevant point concerns with the usage of different data sets to construct the comparison group. The SIPP data panel includes information detailed enough to evaluate eligibility, but the precise location of respondents is unknown and the survey questions are not exactly the same. As a result, selection bias for estimates using this information is typically higher both in absolute and relative terms.

The group of no-shows is quite interesting as these persons are likely to be very similar to the treated. In fact, were enrolment random with respect to outcomes, they would be just like the experimental group. Most probably, however, this is not the case as they opted out of the programme, but the same matching methods as the ones used with eligible non-participants can be applied here to control for the differences. The third panel of table 2.5 shows that the bias is substantially lower using this group than eligible non-participants (except for male youth) but it is more heavily weighted toward the selection bias component, B_3 .

2.4.3 A Simulation Study

To further investigate the accuracy of non-experimental methods, Heckman and Smith (1998, see also Heckman, LaLonde and Smith, 1999) ran a fully controlled experiment based on simulated data. Data are created with an individual's model of participation and earnings and subsequently used to illustrate how biased the different methods are under different underlying hypothesis. Such approach requires a structural model of individual's decisions to be established *a priori*, and the results depend on the particular specification assumed. The model considered can be described as follows. Lets take an individual i

earning Y in period t , such that

$$Y_{it} = \beta + \alpha_i D_{it} + \theta_i + U_{it} \quad (2.17)$$

$$\text{where } U_{it} = \rho U_{i,t-1} + \epsilon_{it}$$

where β is the constant term in the earnings equation, the error term is composed of an individual fixed effect (θ) and an idiosyncratic autoregressive AR(1) process (U) with innovation ϵ ($E(\epsilon) = 0$). The individual-specific fixed effect is

$$\alpha_i = Y_{it}^1 - Y_{it}^0$$

where the superscript stands for the treatment status as before. The participation decision is based on the net gain from treatment, composed of the discounted gain from training (α_i/r where r stands for the interest rate), the foregone earnings at the participation period (Y_{ik}) and the direct cost of treatment (c_i),

$$D_{it} = \begin{cases} 1 & \text{iff } \frac{\alpha_i}{r} - Y_{ik} - c_i > 0 \text{ and } t > k \\ 0 & \text{otherwise} \end{cases} \quad (2.18)$$

Finally, an instrument Z is introduced by modelling the direct costs of treatment as follows,

$$c_i = \phi Z_i + V_i \quad (2.19)$$

where V stands for the error term.

The individual-specific treatment effect, α , is assumed to be independent from all the error components of the model and the instrumental variable, (θ, ϵ, V) . Perfect certainty is assumed except for one case where the individual-specific gains from training, $\alpha_i - E(\alpha)$, are not known at the moment of enrolling into treatment. This model reproduces the widely discussed Ashenfelter's dip.

Given a particular choice of the parameters, the model was used to simulate the behaviour of 1000 individuals over 10 periods (from $k-5$ to $k+4$, where k is the treatment period) 100 times under different assumptions. Table 2.6 displays the results of this simulation, showing the bias by type of estimator for unmatched and matched samples. In the present study, matching is based on earnings two periods before treatment ($k-2$). In

Table 2.6: Bias in non-experimental estimates of the TTE.

| | Heterogeneous treatment effect (α_i) | | | |
|---|--|---|--|---|
| | Homogeneous treatment effect (α) (1) | no selection on α_i (2) | selection on α_i | |
| | | | same variability of α (3) | increased variability of α (4) |
| | | | | |
| TTE: $E(\alpha D = 1)$ | 100 | 98.4 | 615.7 | 971.4 |
| <i>Unmatched sample</i> | | | | |
| Cross section | -494.7 | -494.7 | -98.5 | -60.5 |
| DID ($k - 3$ and $k + 3$) | -55.2 | -55.2 | -9.8 | -7.9 |
| IV: mean | -15.3 | -15.1 | 342.7 | -102.8 |
| IV: median | -5.8 | -5.7 | -146.0 | -247.7 |
| <i>Matched sample (on Y_{k-2})</i> | | | | |
| Cross section | -233.0 | -233.0 | -42.9 | -27.4 |
| DID ($k - 3$ and $k + 3$) | -243.2 | -243.2 | -43.3 | -28.1 |
| IV: mean | 27.3 | 28.2 | -305.0 | -427.5 |
| IV: median | 18.1 | 10.1 | -41.9 | 36.2 |

each case, three possible estimators are considered: the simple cross-sectional differences (CS), DID using periods $k - 3$ and $k + 3$ and IV. The magnitude of the bias is computed under four possible underlying hypothesis about the nature of the effect (homogeneous vs. heterogeneous), the amount of information available to the agent at the enrolling period and the magnitude of the variability of α . Each assumption corresponds to a different column in table 2.6.

Under the homogeneity assumption displayed in column 1, selection occurs on θ and U alone. Given that agents with lower values of θ are more likely to enrol (lower Y_{ik}), the CS estimator is severely downward biased. DID controls for the fixed effect and significantly reduces the bias, only being affected by the AR(1) process, U . It is, however, negatively affected by matching mainly due to the period matching takes place: controls are selected to reproduce treatments at $k - 2$ but they start differentiating immediately at $k - 1$ by recovering earlier in time from the characteristic dip in earnings. Finally, as expected, IV performs well and is consistent in this case.

Column 2 relaxes the homogeneous assumption but considers agents only know about

$E(\alpha)$ before taking treatment. The same conclusions can be drawn from these figures.

Column 3 presents the heterogeneous / perfect foresight case. CS and DID perform better given that selection now occurs largely on α_i . DID, in particular, shows remarkable small bias given that U accounts for a diminished share of the selection process. On the contrary, IV performs much worse now, a feature that is not unexpected since IV is not consistent for TTE under these conditions. It does, however, consistently estimate LATE since the model (2.17)-(2.19) satisfy Imbens and Angrist (1994) monotonic assumption. In the unmatched case, the LATE parameter is estimated to deviate 25% from the TTE when variation in Z is taken around the median. Finally, column (4) considers the same case as column (3) but with increased variability on α_i . Performance improves for CS and DID as participation decisions are more heavily based of α_i , but LATE does not seem to get closer to TTE.

2.4.4 Matching and Difference in Differences: An Area Based Evaluation of the British NDYP

The Blundell, Costa Dias, Meghir and Van Reenen (01) study investigates the impact of the New Deal for the Young People (NDYP) on employment in the first 18 months of the scheme (see the introduction to this thesis for a description of the programme). The identification strategy exploits the programme specific design features, including the fact that the programme was rolled out in certain pilot areas prior to the national roll out and the eligibility rules by which individuals older than 24 by the time they complete 6 months on the claimant count are excluded from the programme. Thus, two instruments were used to estimate the impact of treatment on the treated, namely age and area of residence. A before and after comparison can then be made using a regular DID estimator and improved by simultaneously applying matching as detailed in section 2.3.4. The next chapter details the estimation procedure and discusses the results obtained.

2.5 Conclusions

This chapter has presented an overview of alternative empirical methods for the evaluation of policy interventions at the microeconomic level. It has focused on social experiments,

natural experiments, matching methods, and instrumental variable methods. The idea has been to describe the assumptions and data requirements of each approach and to assess the parameters of interest that they are able to estimate. The appropriate choice of evaluation methods has been shown to depend on a combination of the data available and the policy parameter of interest. No one method dominates and all methods rest on heavy assumptions. Even social experiments rely on strong assumptions: they rule out spill over effects and are sensitive to non-random drop outs from the programme. Natural experiment methods, matching methods and instrumental variable methods all place tough requirements on the data and are fragile to untestable assumptions. Moreover, with heterogeneous response parameters, they each estimate different aspects of the programme impact. It is essential to have a clear understanding of the assumptions and data requirements involved in each method before undergoing an evaluation.

Chapter 3

Evaluating the employment impact of a mandatory job-search assistance programme

This chapter investigates the impact of the New Deal for the Young People (NDYP) on employment in the first 18 months of the scheme. As described in the introduction to this thesis, the NDYP is a recent labour market programme introduced in the UK which offers a number of alternative treatments to unemployed agents that enrol into it. It combines initial job search assistance followed by various subsidised options including wage subsidies to employers, temporary government jobs and full time education and training. The program is mandatory, meaning that all eligible agents are expected to participate or otherwise suffer potential benefits' sanctions. Eligibility is defined in terms of unemployment duration and age: participants are composed by agents aged between 18 and 24 years old claiming unemployment benefit for 6 months. The treatment starts with a period of up to 4 months of intense job-search assistance, denominated by "Gateway", and is followed by a subsidised option. Since the options take-up formally remove individuals from unemployment, the NDYP at least introduces an interval in the claiming spell.

These type of interventions have been quite difficult to evaluate in a robust and convincing way, the main problem being regarded as one of missing data: individuals are either participants or non-participants, but not both simultaneously. This explains why

the main focus of studies on this subject is on recovering the correct counterfactual (see the discussion in chapter 2 for an overview of the strategies adopted and some of the most commonly used methods). Such a task is facilitated by the use of randomised experiments, where the cleanest control groups can be obtained.¹ This kind of experiments, however, is not often available for labour market studies, and different paths need to be explored.² With non-experimental data, the choice of the *appropriate* evaluation method, together with the *appropriate* control group, is case-specific, depending on some criteria like the nature of the programme, the parameter of interest and the nature of the data available (see Blundell and Costa Dias, 00 and 02 and chapter 2 above).³ In the present case, the missing data problem is potentially even more serious given the compulsory nature of the programme and definition of the eligibility rules. Such a combination makes it impossible to find contemporaneous non-participants with the same characteristics as those of participants. If the variables defining eligibility are important in explaining the labour market behaviour of the agents, it may be very difficult to justify the adequacy of the control group.

Our approach consists of exploring sources of differential eligibility and different assumptions about the relationship between the outcome and the participation decision to identify the effects of the NDYP. On the *differential eligibility* side, we use two potential sources of identification. First, the fact that the programme is age-specific implies that using slightly older people of similar unemployment duration is a natural comparison group. Second we can exploit the fact that the programme was first piloted for 3 months (January to March 1998) in selected areas before being implemented nation-wide (the “National Roll Out” beginning in April 1998). This provides an additional dimension to explore on the construction of the control groups. Under a simple difference in differences setting, we show that the choice of the comparison group determines the parameter being

¹See Bassi (83 and 84) Hausman and Wise (85) and Burtless (95) on the advantages of experimental data.

²See Ashenfelter (78), Ashenfelter and Card (85) and Heckman and Robb (85 and 86) for original studies on the impact of social interventions.

³Experimental data may provide very useful information on the adequacy of non-experimental methods by assessing their reliability as in LaLonde (86), Heckman, Ichimura and Todd (97a) and Heckman, Smith and Clements (97).

estimated as various potential sources of biases are dealt with in different ways. We are especially concerned about substitution and equilibrium wage effects. Substitution occurs if participants take (some of) the jobs that non-participants would have got in the absence of treatment. Equilibrium wage effects may occur when the programme is wide enough to affect the wage pressure of eligible and ineligible individuals. While studying the Pilot period, we use a diversity of comparison groups who will be affected differentially by these types of indirect effects to obtain some indication on the importance of such biases.

We apply a number of different econometric techniques, all exploring the longitudinal characteristic of the data set being used but making different assumptions about the structure of the problem. A general set up is developed, where all estimators can be interpreted in the light of combined difference in differences and matching methodologies. The conditions under which each estimator identifies and estimates the impact of treatment on the treated are derived.

No evaluation methodology guarantees a complete solution to the missing data problem. In particular, the estimators being used in the present case rely on two critical assumptions: no selection bias and common time trends across groups.⁴ The former condition may not hold when individuals adjust their labour market behaviour in response to the introduction of the programme, anticipating or delaying their exit from unemployment. The latter assumption is broken if the treated and comparison groups react differently to macro shocks. To establish the reliability of the estimates being presented, we explore these possibilities evaluating their statistical significance.

We focus on the change in transitions from the unemployed claimant count to jobs during the Gateway period. We find that the outflow rate for men has risen by about 20% as a result of the NDYP during its National Roll Out (i.e. 5 per cent more men find jobs in the first four months of the NDYP above a pre-programme level of 25 per cent). Similar results show up from the use of different adopted estimators, independently of the amount or type of structure imposed, and they appear to be robust to pre-programme selectivity, changes in job quality and different cyclical effects. When focusing on the first three months the programme is introduced, there seems to be a large *programme introduction*

⁴See Heckman (79), Heckman and Robb (86), Blundell, Duncan and Meghir (98), Bell, Blundell and van Reenen (99) and Blundell and Dias (00 and 02) for precise descriptions of these conditions.

effect, whereby the impact is twice as large. Using data from the Pilot period, very similar estimates result from the use of different control groups. Such an outcome suggests that either wage and substitution effects are not very strong or they broadly cancel each other out. Despite being a reassuring result, it may not apply to the National Roll Out given the different proportions assumed by the programme and the spread of information about how it works. In fact we show that, based on data from the National Roll Out in the 2nd or 3rd quarter of the operation of the programme, the estimated impacts are lower; this may be an indication that in the longer term the early impacts of the NDYP are not sustainable.

However, there are reasons to expect that a programme such as the NDYP will have long-run sustainable effects. First, the programme is mandatory. Refusal to participate results in sanctions. Mandatory, sanction-enforced schemes have been found to be more effective than voluntary schemes.⁵ Second, the *disadvantaged youth* we consider are less disadvantaged than those typically studied, which compose the treated groups in typical US programs often found to be ineffective (e.g. ex-offenders). The only entry requirement is six months unemployment benefit claim, which is not so uncommon for those under 25 in Britain. Finally, recall that we are evaluating the effects of job search assistance and wage subsidies. Based on some U.S. evidence, such programs may be more effective.⁶

The structure of the chapter is as follows. We start in section 1 with a discussion of outcome of interest within the NDYP context. Section 2 presents the methodology we apply to estimate the effects of the NDYP Gateway. Within this environment, we discuss how the choice of the comparison group determines the parameter being identified along with the potential sources of bias in each case, and develop a combined difference in differences and matching set up where all the estimators being used can be interpreted. Section 3 describes the data and section 4 details the empirical results. We separate the analysis of the Pilot period of the programme, where more detail is possible given the additional instruments we are able to explore to construct the counterfactual. Males and females are also discussed separately. Finally, section 5 offers some concluding comments. The appendix to this chapter is presented on section 6, containing further information

⁵For example, Knab, Bos, Friedlander and Weissman (00) and Moffitt (96).

⁶On job-search assistance see the survey by Meyer (95) and on wage subsidies see Katz (98).

on the data and intermediate results, estimates for the effect of the ND during the Pilot period using different weighting schemes along with propensity score matching and details on the methods of estimation.

3.1 The NDYP and the choice of the outcome variables

As described in the introduction to this thesis, the NDYP is a recent UK labour market programme targeted at the 18 to 24 years old long-term unemployed and aimed at improving participants perspectives in the labour market by enhancing their skills and human capital. Individuals aged between 18 and 24 by the time of completion of the sixth month on Job Seekers' Allowance (JSA) are assigned to the programme and start receiving treatment.⁷ On assignment to the programme, the individual starts the first stage of the treatment called the Gateway. This is the part of the programme being evaluated in the present study. It lasts for up to 4 months and is composed of intensive job-search assistance and small basic skills' courses. The second stage is composed of four possible options, including subsidised employment, subsidised education, employment on the voluntary sector and employment on the environmental task force. If the treated agent returns to the claimant count at the end of the option period, the third stage of the programme is initiated, the Follow Through, which is composed of 13 weeks of job-search assistance.

Prior to its national release, the NDYP was piloted for a three months' period in 12 areas called the "Pathfinder Pilots" (see Anderson, Riley and Young, 1999). This happened between January to March 1998. Clearly, identification of the treatment effect under these conditions requires stronger assumptions than when an experiment is ran within regions using random assignment. As will be discussed, the problem relates with the fact that the counterfactual must either be drawn from a different labour market or from a group with different characteristics operating in the same labour market. Below we explore what we can identify under different assumptions.

⁷The JSA is the main form of unemployment benefit currently available in the UK. It is essentially a flat rate benefit paid every two weeks of about £40 a week. Past work experience is not a condition of receipt of JSA and although there is a requirement to *actively seek employment*, it is not time limited. See the introduction to this thesis for details.

Given that the programme has not been running for a long period, the focus of this chapter is on the evaluation of the Gateway. In particular, we are concerned with the degree to which enhanced job-search assistance has led to more outflows to jobs. The evaluation is based on data provided by the Pathfinder Pilot areas before the National Roll Out of the programme, as well as on data available following the National Roll Out. There are two main issues that need to be considered in evaluating the impact of the programme: the precise nature of the comparison group, and hence the definition of what is being measured, and the set of assumptions that underlie the interpretation of the parameter we estimate in each case. The clear understanding of these issues is an important input in an eventual cost-benefit analysis of the programme since they determine the outcome from the programme. There are some important aspects covered within this discussion. One of them concerns the extent to which we can estimate the overall impact of the programme on employment as opposed to the impact on the eligible individuals. Potential differences in the two outcomes may result from two main factors. First, the impact of the programme on eligible individuals may be at the expense of worsened labour market opportunities for similar but ineligible individuals. Second, the wider implementation of the programme and the opportunities it offers to participants may affect the equilibrium level of wages and employment, affecting all workers.

We study the impact of the programme on the proportion leaving unemployment within four months of entering the Gateway. The choice is mainly dictated by the desire to focus on the stated government targets and the paucity of data on individuals after they have finished the options.⁸ However an alternative outcome variable would have been the proportion leaving unemployment within, say, 8 or 10 months of entering the unemployment pool. This outcome variable would avoid the potential composition effects that may be induced by the anticipation of the programme for the eligible population. In particular, if the programme is perceived as being able to improve placements, then individuals close to the Gateway and eligible for the programme may reduce their search effort and wait for the programme. In this case, the average individual among eligibles would be more

⁸ Available data currently ends in July 99. Individuals entering the Gateway during April 98 and joining the one year long education and training option after four months of job-search assistance will only start searching for a job in August 99.

prone to leave unemployment than its counterpart in the comparison group, leading to increased exit rates for this group. However, we can test this hypothesis by estimating the proportion of those who left unemployment by the end of the sixth month in the eligible and ineligible group. Such a comparison will provide an idea of how important such compositional effects are likely to be.

We will pay special attention to the outflows into employment, but we also examine total outflows from unemployment to all destinations. To assess the importance of the estimated effects, we interpret them in an historical perspective. We provide some lower and upper bounds for the treatment effect by using our methodology during other pre-programme time periods. This can be done for total outflow for all years since 1982.

3.2 Identification and Estimation Methods

Our approach to estimate the impact of the NDYP programme relies on using information from the pilot period as well as information from the National Roll out.

The NDYP can affect employment of both eligible and ineligible individuals in a number of ways. First the eligible individuals receive job search assistance which may enhance their ability to find a job. Second, some of the individuals in the Gateway programme receive wage subsidies, reducing the cost of employing them for an initial period of six months. This wage subsidy will expand the employment of such workers but may also lead to a substitution of other workers for these cheaper ones. The extent to which this may happen will depend on a number of factors. If the subsidy just covers the deficit in productivity and the reservation wage of the workers as well as the costs of training, we would not expect any substitution; these workers are no cheaper than anyone else. Second, it will depend on the extent that these workers are substitutable in production for existing workers and on the extent that it is easy to churn workers. The latter is an important point, since the subsidy only lasts six months. Moreover the agencies implementing the NDYP are supposed to be monitoring the behaviour of firms using wage subsidies and employing individuals on the NDYP. Of course if job durations are generally short, firms will be able to use subsidised workers instead of the non-subsidised ones, without any extra effort.

An additional effect of the NDYP may be to decrease wage pressure through the increase in labour supply and through the presence of wage subsidies. This will tend to increase employment for all types of workers and will counteract the effects of substitution on the non-treatment group.

Assessing the importance of substitution and of general equilibrium (GE) effects through wages or other channels is of central importance. Using the comparison between the pilot and control areas as described below, and assuming these areas are sufficiently separate labour markets from each other, we will be able to assess the extent to which substitution and other GE effects combined are likely to be important *side-effects* of the programme, at least in the short run. Below we discuss the evaluation methodology, a central part of which is the choice of the comparison group. This choice is to a large extent governed by the issues discussed above.

Define by Y_{it}^1 the outcome for individual i in period t if exposed to the policy (treatment). The outcome for the same individual if not exposed to the policy is Y_{it}^0 . Consequently, the impact of the policy for the i -th individual at time t is $Y_{it}^1 - Y_{it}^0$. The average policy impact for those going through the NDYP is $E(Y_{it}^1 - Y_{it}^0 | ND = 1)$. This parameter will be the focus of our attention. Quite clearly, the evaluation problem relates to the missing data that would allow us to estimate $E(Y_{it}^0 | ND = 1)$ directly. In this section, we define a number of alternative comparison groups that will allow us to estimate this counterfactual mean. As will be pointed out, the definition of the estimated parameter will change in certain cases with the comparison group.

Consider first a contrast obtained by comparing employment growth in pilot areas to employment growth in control areas. Assume that

$$E(Y_{it}^0 | ND = 1, t = 1) - E(Y_{it}^0 | ND = 1, t = 0) = \tag{3.1}$$

$$E(Y_{it}^0 | ND = 0, t = 1) - E(Y_{it}^0 | ND = 0, t = 0)$$

where $ND = 1$ denotes the Pathfinder Pilot areas assigned to the NDYP pilot, $t = 0$ represents the period before implementation and $t = 1$ stands for the period after. This assumption means that the growth in employment in the Pathfinder Pilot areas would have been the same as in the non Pathfinder Pilot (control) areas in the absence of the

policy. In this case the missing counterfactual value can be replaced by

$$E(Y_{it}^0 | ND = 1, t = 1) = E(Y_{it}^0 | ND = 1, t = 0) + m_t$$

which is simply the employment level in the Pathfinder Pilot areas before the policy was implemented, adjusted for aggregate employment growth, given by

$$m_t = E(Y_{it}^0 | ND = 0, t = 1) - E(Y_{it}^0 | ND = 0, t = 0)$$

This gives rise to a straightforward difference in differences estimator. Under the assumption in 3.1, such a comparison of growth rates estimates the impact of the NDYP on individuals residing in a Pathfinder Pilot area, irrespective of whether they are eligible or not. Hence, this comparison estimates the effect of the programme including any impact of GE effects and is net of substitution.

However we can obtain an idea of the importance of indirect effects by comparing the growth of employment in pilot and control areas separately for eligible and ineligible individuals. Under assumption 3.1 applied separately to eligible and ineligible individuals, comparing the growth in the employment for the eligible individuals in the pilot and control areas will measure the combined impact of the treatment, substitution and GE due to wage changes. Comparing such estimate to that obtained using the ineligible individuals will net out the impact of substitution between the two groups, but will leave the effect of wage changes.⁹

The definition of the comparison group is of course central to the evaluation. The approach discussed above, used as comparison group the individuals in non-exposed areas during the pilot period. However, the pilot stage lasted three months only and it is possible that the impacts of the policy in this short first period are not generalisable, if anything because the administration of the programme would have been in its infancy. So, we next consider using data from the National Roll Out, the term referring to the national implementation. Suppose we start by assuming that assumption 3.1 is valid when $ND = 1$ refers to eligible individuals following the National implementation and $ND = 0$ refers to

⁹However, the indirect effect of wage changes will be multiplied by 2 under the assumption that all groups are similarly affected by such effect. If prices adjust slowly, the short-run estimates using such procedure will identify the overall impact of programme on the treated population.

similar but ineligible individuals, i.e. those unemployed over 6 months whose age is just above 24. The choice of this group makes it most likely that their overall characteristics and behaviour match that of the treatment group. Thus, the growth rate of employment for the two groups would be similar in the absence of the programme. Such an approach is similar to a regression discontinuity design.¹⁰ By making assumption 3.1 with respect to these two groups, we are ruling out any substitution effects or equilibrium wage effects that impact on the groups in a differential way. In this case, a comparison of the growth rates between eligible and ineligible individuals will provide an estimate of the impact of the programme on the eligible ones.

The virtue of the comparison group in terms of similarity to the treatment group may in fact be its greatest disadvantage. The substitution effects are likely to be much more severe the more substitutable the two groups are in production. In the event of substitution, the impact of the programme for the eligible group is biased upwards by the fact that the employment of the comparison group is decreasing. If such a decrease is, say, s , the net increase in employment is $2s$ lower than the estimated increase in employment. However the benefit in terms of employment for the target group would be s lower than our estimate. Within this framework of analysis, the only way we have of gauging the size of s is through the pilots, as discussed above. Alternatively a GE model would allow us to estimate s , at least in the long run, based on the substitutability of the two groups in production.

There are a number of additional issues that we need to address. First, there is the basic issue of whether we can assume that the two groups are subject to the same aggregate labour market trends. To the extent that the human capital of the two groups is perfectly substitutable and to the extent that preferences for work are the same, this assumption will be satisfied. However the latter in particular may not be the case. At that age, family formation takes place and the older individuals are more likely to be married, which may be associated with an increase in labour supply. We can address this issue by examining the trends in the exit rate from unemployment of the two groups for a number of years prior to the implementation of the NDYP. Over the preceding years there has been no major policy that explicitly discriminates between the two groups. This approach also suggests

¹⁰See Hahn, Todd and Van der Klaauw (01).

a method for bounding the impact of the policy using the historical trends in the two groups. In particular we can identify the pre-programme period within our data set which would maximize the impact of the policy (i.e. minimise m_t) and the period that would minimize it (i.e. maximize m_t). In the empirical section we show the historical trends for the two groups and we provide bounds for our estimates based on these fluctuations between the two groups.

The next important issue is whether the impact of the policy is heterogeneous with respect to observable characteristics, represented by X . If this is the case, we should interpret the estimate we obtain as an average impact across different effects but must make sure that a suitable comparison group exists. One way to address this problem is to use propensity score matching adapted for the case of difference in differences. In this case, there are two assignments that are non-random. One assignment is to the eligible population and the other assignment is to the relevant time period (before or after the reform). For the evaluation to make sense with heterogeneous treatments, we must guarantee that the distribution of the relevant observable characteristics is the same in the four cells defined by eligibility and time. One way of achieving this is to extend propensity score matching by defining two propensity scores - one for eligibility and one for time period. We then create a matched sample based on the two propensity scores. This approach ensures that the distribution of observed characteristics is balanced across all cells. In general, the assumption required to justify this approach is that

$$E(Y_{it}^0 | X, ND = 1, t = 1) - E(Y_{it}^0 | X, ND = 1, t = 0) = \\ E(Y_{it}^0 | X, ND = 0, t = 1) - E(Y_{it}^0 | X, ND = 0, t = 0)$$

where $ND = 1$ denotes eligibility and t is the time period. Following Dearden et al. (01), under this assumption it is possible to construct matched samples by conditioning on the propensity scores for eligibility, $P_{EX} = P(ND = 1 | X)$, and for being observed in time period $t = 1$, $P_{tX} = P(t = 1 | X)$

$$E(Y_{it}^0 | P_{EX}, P_{tX}, ND = 1, t = 1) - E(Y_{it}^0 | P_{EX}, P_{tX}, ND = 1, t = 0) = \quad (3.2) \\ E(Y_{it}^0 | P_{EX}, P_{tX}, ND = 0, t = 1) - E(Y_{it}^0 | P_{EX}, P_{tX}, ND = 0, t = 0)$$

Matching is performed on the observables in levels, which include information ob-

servable before enrolment only. This is because the X s aim at controlling for the eligibility status, which is evaluated determines participation. The observables we use include, among other things, labour market history. The approach can be implemented non-parametrically. In addition we compute simpler parametric methods that condition linearly on a number of observable characteristics. We discuss further these issues in the estimation section below.

Finally the discrete nature of our outcome variable may imply that the assumptions we make do not hold for the expectations (which are employment probabilities) but for some transformation thereof. In particular, one may be willing to apply the inverse of the probability function, which must be assumed known. The additional problem with discrete outcomes that is not usually addressed in the literature pertains the commonly used distributional assumptions for the discrete choice models, like the logit or probit. They do not support additively separable individual and time effects required by the difference in differences method. Instead, such assumptions should be imposed on the index rather than the probability itself. This is why it may be desirable to assume conditional independence on the indexes instead of the outcome.¹¹ In this case we assume that

$$f^{-1} [E (Y_{it}^0 | X, ND = 1, t = 1)] - f^{-1} [E (Y_{it}^0 | X, ND = 1, t = 0)] = \\ f^{-1} [E (Y_{it}^0 | X, ND = 0, t = 1)] - f^{-1} [E (Y_{it}^0 | X, ND = 0, t = 0)]$$

where f^{-1} is the inverse of the probability function (e.g. the inverse logistic).

Define by Y_{it} the employment indicator for individual i in period t . In the Pathfinder Pilot areas in period $t = 1$, this will represent the outcome under treatment. In all other cases it will represent an outcome under non-treatment. The impact of the policy can then be evaluated as

$$I(X) = E(Y_{it} | X, ND = 1, t = 1) - f[f^{-1}(E(Y_{it} | X, ND = 1, t = 1)) - \alpha(X)] \quad (3.3)$$

¹¹A recent study by Athey and Imbens (02) notices that DID applied to discrete choices has the undesirable property of potentially generating probabilities outside the $[0,1]$ range. This problem is avoided in the present case by imposing the conditional independence assumption on the indexes.

where

$$\alpha(X) = \{f^{-1}[E(Y_{it}|X, ND = 1, t = 1)] - f^{-1}[E(Y_{it}|X, ND = 1, t = 0)]\} - (3.4) \\ - \{f^{-1}[E(Y_{it}|X, ND = 0, t = 1)] - f^{-1}[E(Y_{it}|X, ND = 0, t = 0)]\}$$

3.2.1 Implementation

Given a particular choice of control group, all methods we apply have the same structure as implied by (3.3) and (3.4). They differ only in the way that the expectations in these expressions are computed.

In the linear matching difference in differences estimator, we run the following simple regression on the sample of control and treatment observations

$$Y_{it} = \theta_{ND} + d_t + \gamma'X + \alpha ND_{it} + \varepsilon_{it}$$

where Y_{it} is a discrete variable indicating whether the person is in employment or not, θ_{ND} is an eligibility specific intercept (may it be area or age defined or both, depending on the comparison group used), d_t reflects common/aggregate effects and X is included to correct for differences in observable characteristics between the areas. Alternatively, one can use a different parametric specification for the outcome as a function of the index presented above, $\theta_{ND} + d_t + \gamma'X + \alpha ND_{it} + \varepsilon_{it}$, and estimate the effect of treatment, α , under such transformation.

These procedures can be quite restrictive in a number of ways. First, they do not allow for α to depend on X . And second, they do not impose common support on the distribution of the X s across all four cells.

The first assumption can be relaxed under the parametric setting, and this is what we do within the non-linear logit specification. The effect of treatment is allowed to depend on the observable characteristics of the agents by applying the following estimation technique. A different relationship between the outcome and the observables is estimated by group of agents (treatment status * time). Such relationships entail the particular behaviour pattern of each group and the impact of treatment when it existed. By predicting the outcome of treated under the non-treated behavioural equation one obtains an estimate of how the treated would have been without the treatment would they belong to each of the

other groups and keep their observable characteristics. Applying difference in differences to such predictions produces an estimate of the expected impact of treatment among treated.

To relax both assumptions simultaneously, we supplement the above results by propensity score matching. As mentioned above, this involves matching on two propensity scores, which balances the distribution of the X characteristics in the treatment and control samples, before and after the reform. The matching method we use smoothes the counterfactual outcomes either with a Kernel based method or with splines (see, Heckman, Ichimura and Todd, 97 and Meghir and Palme, 01). We also present results based on nearest neighbour. These however turn out to be much less precise. We provide details on the estimation method in the appendix to this chapter (section 3.6.3).

3.2.2 Other estimation issues

The choice of the comparison group

As discussed above, the available options for the choice of the comparison group depend on the type of evaluation being performed. When assessing the programme from data on its National Roll Out, we are constrained to use ineligible individuals within the same area, for which we have chosen the age rule to define (in)eligibility. For the Pilot Study, however, the regional rule provides an additional instrument in the definition of the comparison group. We have used it in two ways, constructing two possible comparison groups: The first takes all eligible individuals living in all control areas; The second selects all eligible individuals in the set of control areas that most closely resemble the Pathfinder Pilot areas in a way detailed below.

The goal of a careful choice of the comparison group is to satisfy assumption (3.2) which requires that the time trend evolves in the same way for treatments and controls.

To have an idea of how similar any two groups are, we compare them in historical terms before the NDYP is introduced. The comparison was established on the outcome of interest, the conditional outflows from unemployment. The comparison groups are established only on the basis of the eligibility rules, not taking into account any other observable.

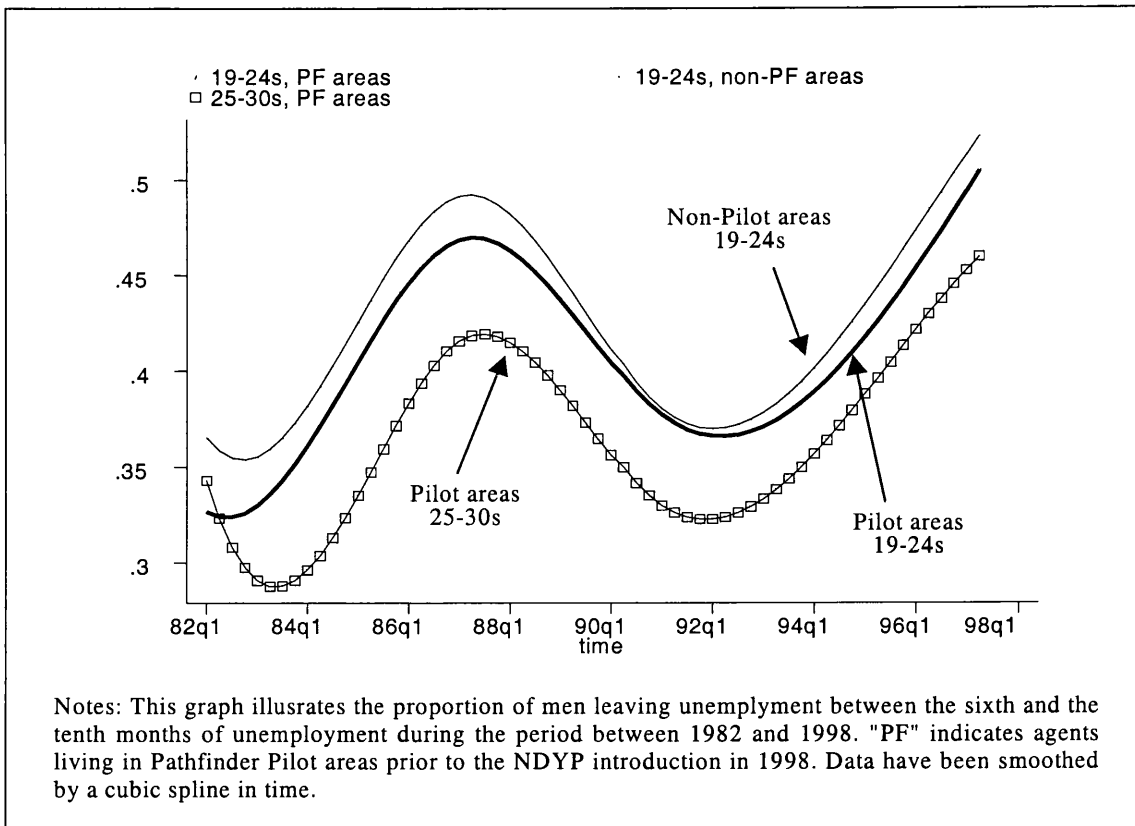


Figure 3.1: Outflows from the JSA conditional on completing 6 months on the claimant count. Effect by the end of the 10th month on JSA. Men only.

Figure 3.1 illustrates the evolution of the outcomes for men aged 19-24 years old and living in Pathfinder Pilot areas and in all control areas. It is clear that the Pathfinder Pilot areas have, on average, worse labour market conditions. However, for the purposes of evaluating the impact of the programme based on these two groups, what is important is that the difference between the two curves is kept nearly constant over time in order to guarantee that macro trends affect the two groups in similar ways. The older group aged 25-30 is also presented as a potential comparison. This group tends to have lower outflows than their younger counterparts. However, since 1990 the difference in the outflows over the cycle is similar. Nevertheless, this data shows that the size of the estimated impact can be sensitive to the choice of period for comparison and in the results section we are careful to test the sensitivity of the results to alternative timing assumptions.

Choosing comparable areas

When using all eligible individuals in control areas as a comparison group (or a matched sub-sample of them), it is being assumed that the two curves represented in Figure 3.1 are indeed parallel so that similar individuals are similarly affected by macro trends, independently of where they live. One can, however, choose the areas that more closely follow the cycle pattern identified for the Pathfinder Pilot areas. This can be done either within each of the matching procedures described above, or prior to them, selecting the areas where the comparisons are to be drawn from. We have chosen to adopt this latter option, matching the areas in a first step and applying all types of estimators comparing eligibles in different areas to the sub-samples obtained. In this procedure, we have used a completely non-parametric technique, as described below.

The aim of matching the areas is to achieve a match as close as possible with respect to labour market characteristics. The procedure followed to match on labour market characteristics makes use of a quarterly time-series of the outcome variable from 1982 to just before the introduction of the NDYP, in January 1998. A measure of distance was then computed for each possible pair of Pathfinder Pilot and control areas and the two nearest neighbours were chosen. Once the two nearest neighbouring areas have been chosen based on similarity of the labour market trends, we carry out the estimation procedure as described earlier.

Sensitivity of the results

The relative size of the estimated impact of the programme, when viewed in an historical perspective, can inform on how significant the result is. In order to do so, the series of year-by-year estimates of the impact of a fictitious programme has been computed.¹² Given the lack of data on “destination when leaving JSA” before August of 1996, we use

¹²This analysis is also informative on whether the assumptions on the comparability between any two groups being used are valid. In fact, before the introduction of the NDYP, the estimated impacts are expected to be zero given the absence of a policy that causes a differential behaviour between any two groups being compared. If, however, a large number of point estimates is found to be significantly different from zero, one might suspect that the assumptions on the comparability of the two groups being used are not valid.

information on “exits to all destinations” to perform this analysis.

Suppose, for instance, that the estimated effect of the NDYP Gateway lies within typical values of the historical estimates. This might be an indication that such result is determined by some random variation that is not being controlled for and is captured by the programme dummy. In such a case, doubts are raised on whether the estimated effect is actually capturing the causal effect of the programme alone. We can go further and bound the estimated impact of the Gateway using the distribution of year-by-year estimates to construct an upper and a lower bound to the estimated effect. This is done by taking the percentiles on the tail of the distribution - say, percentiles 5 and 95 or 10 and 90 - as being the expected value of the estimates in the absence of a programme, and using them to re-scale the estimated impact up or down accordingly.

Compositional changes in the treatment group

Such a large-scale programme may have compositional effects on the group of eligible individuals. Having learned about the eligibility rules, potential participants may change their behaviour in order to secure or avoid enrolment. If such a selection process is taking place, the estimated effects of the programme will be affected because the groups being compared are not what they would have been in the absence of the programme. We check for this selection bias by examining difference in difference estimates of individuals' probabilities of exiting unemployment in the pre-treatment period (i.e. in the months before reaching six months unemployment when the programme begins).

3.3 Data

The data are drawn from the publicly available 5 per cent longitudinal sample of the whole population claiming Job Seekers Allowance (JSA) in the UK from 1982 to June 1999 (the JUVOS database). This is an administrative database that includes individual information on spells on JSA, the unemployment benefit available in the UK, the main focus being the starting and ending dates of the spells. Individuals can be followed through all their JSA spells since the same group of the population is followed over time. However, although we know the length of time in non-JSA spells, we have no information on any

transitions between different jobs during these periods. Since 1996, however, the agencies have collected data on the destination when leaving the claimant count. There are 20 different destination codes, including exit to employment, training/education, other benefits, incarceration, etc. The JUVOS data set also includes a small number of other variables - age, gender, marital status, geographic location, previous occupation and sought occupation. Descriptive statistics on the treatment group and different comparison groups are presented in the appendix to this chapter (section 3.6.1).

We also make use of the NDYP Evaluation Dataset (NDED), an administrative data set that contains information on virtually all individuals that have gone through the NDYP, even if only briefly. For participants, very detailed information is available from the time they join the programme, including the types of treatment being administered and the timing of each intervention, letters being sent and interviews being made, a long list of socio-demographic variables and the destination when leaving the programme. Non-participants, however, are not included in the sample, which limits its use for evaluation purposes.

The use of the NDED is meant to complement the lack of information in JUVOS about the take-up of NDYP options. Since starting an option implies dropping from the JSA claimant count, there is a potentially large group that is being re-classified as non-unemployed while simply being driven through the programme according to its rules. Unfortunately, we are unable to securely identify these types of exits from the JUVOS data set.¹³ We use the NDED instead to know the proportion of participants that enrol in each type of option (in any given region-date) by length of the NDYP spell.

In drawing up the treatment groups we have used 19-24 year olds even though the NDYP also affects 18 year olds. This is because 18 year olds can still be in high school and high school is only compulsory up to the age of sixteen in the UK. Participation of 16 to 18 year olds in full time education grew rapidly over this period so we decided to avoid any time varying composition effects by dropping 18 year olds. In any case, inclusion

¹³There is a code in the JUVOS data which purports to have NDYP destinations but on investigation it proved very unreliable.

made no difference to the results.¹⁴

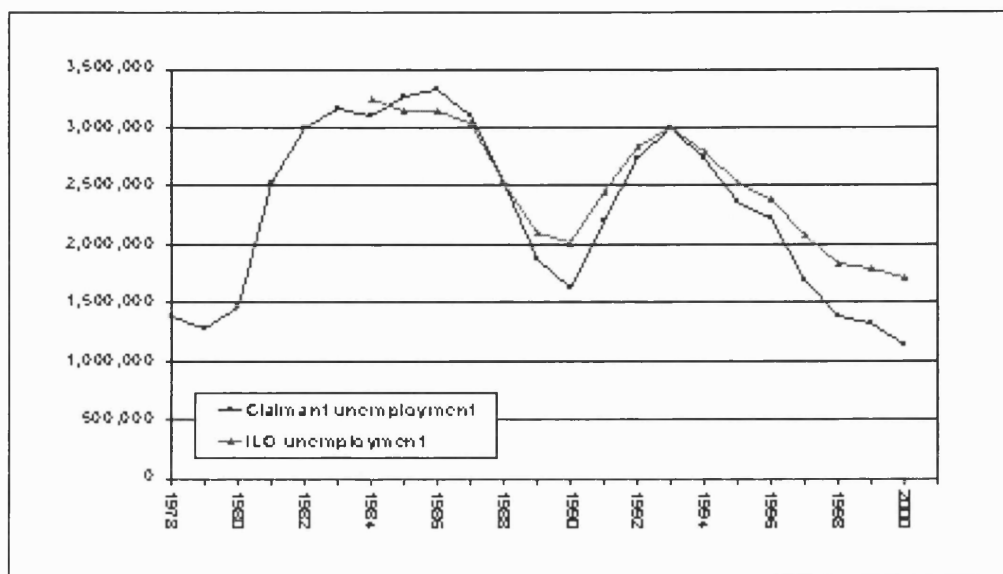
The historical period we are examining is partly dictated by the data. The current JUVOS data ends in July 1999. For the National Roll Out we consider all individuals who finished a 6-month JSA spell between April and December 1998 and then follow them up to four months later (so our end date is April 1999). We match this with the individuals who finished a 6-month JSA spell between April and December 1997. For the Pilot Study we compare individuals completing a 6-month JSA spell between the start of January and the end of March 1998 in the Pathfinder Pilot areas to the same group in January through March 1997. Ending the sample in April 1999 has the advantage that we avoid contaminating the NDYP effect with the introduction of the national minimum wage enforced from April 1999 onwards.¹⁵

Some information on the macro-economic climate is given in Figure 3.2. The NDYP was introduced at a favourable point of the business cycle by historical standards. There was no rapid improvement in the labour market between Spring 1998 and 1999, however, unlike the previous 12 months. The changing business cycle illustrates the reason why we have to select our comparison groups carefully in implementing our approach to ensure that these macro trends are “differenced out”.

Finally, it should also be pointed out that the effects of the programme in this favourable climate may not be easily applied to less favourable periods. First the pool of unemployed is likely to be of worse quality when the aggregate economy is booming. Opposing this is the fact that, in the presence of firing costs (formal or informal) hiring someone in boom may be less risky.

¹⁴One could also worry about 18-22 years old in college education. There is only a tiny fraction of this group in the unemployment pool, however.

¹⁵Britain had never had a national minimum wage before this date. There was a system of Wage Councils that set minima for certain groups of occupations in low wage industries. These only covered about 2 million of the 30 million UK workforce when they were abolished in 1993 (see Dickens, Machin and Manning, 99, for an analysis).



Notes: Data on ILO unemployment from Labor Force Survey, claimant count taken from Labor Market Trends, various issues.

Figure 3.2: Unemployment - claimant count and ILO measures

3.4 Results

This section presents estimates of the impact of the Gateway on the flows into employment. We analyse men and women separately given the different composition of the two groups and characteristics of their behaviour. We start by considering the men's case during the Pilot Period in subsection 1, and discuss the different possible estimates and respective underlying assumptions available. Subsection 2 presents the results obtained for men during the National Roll Out, establishing a comparison with what the estimates were for the Pilot Period and assessing the their robustness. Finally, the women's case is discussed in subsection 3.

3.4.1 Pilot Study: men's results

Tables 3.1 and 3.2 present the main estimates of the impact of the Gateway on eligible agents and the impact of the existence of the programme on ineligible, both groups being composed of men living in Pathfinder Pilot areas during the Pilot period. We consider a

Table 3.1: Effect of treatment on the treated men. Pilot period.
Gateway employment effects by the end of the 10th month.

| | Control group | Observations | Difference in differences combined with | | | |
|--|--|--------------|---|--------------------------|--|--|
| | | | (1) Linear matching | (2) Logit matching | (3) PS† matching | (4) PS†logit matching |
| <i>Comparing 19-24s living in all Pathfinder Pilot areas</i> | | | | | | |
| (1) | 19-24s in all control areas | 3,716 | 0.110** (0.039) | 0.098** (0.039) | 0.104** (0.046) <i>(0.024;0.182)</i> | 0.098** (0.044) <i>(0.015;0.176)</i> |
| (2) | 19-24s in matched control areas | 1,193 | 0.134** (0.053) | 0.073 (0.060) | 0.093 (0.073) <i>(-0.015;0.226)</i> | 0.080 (0.063) <i>(-0.018;0.190)</i> |
| (3) | 25-30s in all Pathfinder Pilot areas | 1,096 | 0.104* (0.055) | 0.091 (0.057) | 0.078 (0.079) <i>(-0.050;0.195)</i> | 0.074 (0.069) <i>(-0.068;0.182)</i> |
| (4) | 31-40s in all Pathfinder Pilot areas | 1,169 | 0.159** (0.050) | 0.096 (0.062) | 0.099* (0.078) <i>(-0.015;0.231)</i> | 0.082 (0.082) <i>(-0.063;0.205)</i> |
| | Outflow into the employment option | 4,486 | | | 0.057 | |

Notes: Estimates of the effects of the NDYP used the JUVOS 5% longitudinal sample of JSA claimants. Estimates of the outflows into the employment option used the NDED. Selected observations are those completing a 6 month spell on JSA over a predefined time interval - the present table considers the 1st quarters of 97 and 98. These individuals are then followed up to the end of the 10th month on JSA to check whether they have found a job. The eligible group (defined by the age and area criteria) is compared with the selected control group before and after the release of the programme to estimate its impact. All estimates from regressions controlling also for marital status, sought occupation, region and some information on the labour market history (comprising the number of JSA spells and the proportion of time on JSA over the 2 years that precede the start of the present spell). Age and the number of JSA spells since 1982 are also included when similar age groups are being compared. Propensity score matching is implemented over the same covariates as the other estimates and the outcomes for the comparison groups are smoothed using cubic splines on the two propensity scores to achieve higher precision. Standard errors in parentheses: estimates for non-linear matching method (column 2) used the delta method and estimates for the propensity score matching (columns 3 and 4) used bootstrapping with 200 replications. Bias-corrected 90% confidence intervals in italic - estimation used the same bootstrap results.

** = significant at 0.05 level. * = significant at 0.10 level. †Propensity score.

number of different possible comparison groups, providing some insight on the possible size of indirect effects. Each row in the table corresponds to a different comparison, including

Table 3.2: NDYP effect on the ineligible and the whole economy. Pilot period.
Gateway employment effects by the end of the 10th month. Men only.

| | Control group | Observations | Difference in differences combined with | | | |
|--|---------------------------------------|--------------|---|--------------------------|-------------------------------------|-------------------------------------|
| | | | (1) Linear matching | (2) Logit matching | (3) PS† matching | (4) PS†logit matching |
| <i>Comparing 25-30s living in all Pathfinder Pilot areas</i> | | | | | | |
| (1) | 25-30s in all control areas | 3,180 | 0.016 (0.042) | -0.012 (0.043) | 0.027 (0.049) (-0.058;0.107) | 0.031 (0.050) (-0.052;0.109) |
| (2) | 25-30s in matched control areas | 983 | 0.055 (0.058) | -0.027 (0.056) | -0.003 (0.066) (-0.107;0.112) | -0.018 (0.078) (-0.144;0.117) |
| <i>Comparing 19-30s living in all Pathfinder Pilot areas</i> | | | | | | |
| (3) | 19-30s in all control areas | 6,896 | 0.066** (0.029) | 0.052* (0.030) | 0.058* (0.034) (0.004;0.114) | 0.051 (0.034) (-0.004;0.109) |
| <i>Comparing 19-50s living in all Pathfinder Pilot areas</i> | | | | | | |
| (4) | 19-50s in all control areas | 12,749 | 0.036* (0.021) | 0.035* (0.021) | 0.044* (0.023) (0.004;0.080) | 0.042* (0.023) (0.004;0.078) |

Notes: see notes to table 3.1.

** = significant at 0.05 level. * = significant at 0.10 level. †Propensity score.

different estimates, obtained under different methods, of the effects of the Gateway on outflows to employment after 4 months of treatment.¹⁶ The first row in table 3.1 compares men aged 19 to 24 years old living in Pathfinder Pilot areas with a similar 19-24 years old group living in all control areas. After 4 months of treatment, it is estimated that the Gateway has improved participants' exits into employment very significantly - all the estimators point to an impact of about 10-11 percentage points. This effect is even more impressive if compared with the outflow rates reported in Table 3.3. In the pre-programme period only 24 per cent of individuals in the treatment group obtained employment over the similar four months period (compared to 33 per cent afterwards). Thus, the improved

¹⁶All regressions control for a set of other variables, including age when similar age groups are being compared, marital status, region, sought occupation and labour market variables. All computations have been performed excluding these covariates as well. Given the similarity of the results, however, we skip their presentation.

Table 3.3: Flows from the claimant count into employment.

Results by the end of the 10th month conditional on completing 6 months on JSA. Men only

| | Flows by the end of the 10th month on JSA. | | |
|---|--|---------------------|------------|
| | Before the programme | After the programme | Difference |
| <i>Pilot period</i> | | | |
| Treatments: 19-24s in PF†areas | 0.241 | 0.330 | +0.089 |
| Controls: 19-24s in all control areas | 0.271 | 0.250 | -0.021 |
| Controls: 19-24s in matched control areas | 0.228 | 0.233 | +0.005 |
| Controls: 25-30s in PF†areas | 0.276 | 0.260 | +0.016 |
| <i>National Roll Out</i> | | | |
| Treatments: 19-24s | 0.258 | 0.281 | +0.023 |
| Controls: 25-30s | 0.230 | 0.199 | -0.031 |

Notes: Estimates used the JUVOS 5% longitudinal sample of JSA claimants. Selected observations are those completing a 6 month spell on JSA over a predefined time interval. The present table considers the 2nd to 4th quarters of 1997 and 1998 for the National Roll Out estimates, and the 1st quarters of 1997 and 1998 for the Pilot period estimates. Individuals verifying this criterion are then followed up to the end of the 8th and 10th months on JSA to check whether they have found a job. The eligible group is compared with the selected control group.

†Pathfinder Pilot areas.

job-search assistance provided during the Gateway seems to have raised the probability of getting a job by about 42% (=10%/24%) after 4 months of treatment.

Of course, this result should be contrasted with the information from the NDED (NDYP Evaluation Database) concerning outflows into the employment option. It is estimated that the outflows into an employment option after 4 months of treatment sum up to 5.7 per cent of men joining the Gateway (see Table 3.1). Subtracting this off the overall NDYP effect would give a *pure* Gateway impact (on outflows to unsubsidised employment) of about 4 per cent. But this is likely to be a lower bound. The calculation assumes that there is essentially no deadweight of the employer subsidy. This happens under the assumption that participants can be split into groups according to their ability to find a job, and that subsidised jobs are being attributed to those in need of a subsidy to leave unemployment. If, on the other extreme, it is believed that the subsidised jobs are being allocated to the most employable participants, then the amount of scaling down required might be small. Furthermore, the NDED will tend to find larger job outflows because of fewer missing values. Thus 4 per cent is a lower bound for the pure Gateway

job assistance effect. The method used to estimate the impact of treatment does not seem to substantially influence the results, reflecting some robustness of the estimates to the functional form assumptions.¹⁷

The rest of the rows in table 3.1 present estimates for some of the other identifiable parameters discussed in section 3, also providing some clues about the robustness of the results. We start by restricting the comparison group to be composed of eligible men living in matched control areas in the second row. Depending on the method used, the estimated effect may rise or fall slightly, but not significantly so. This evidence supports the comparability of the two groups used in row 1.

The third row compares eligible and ineligible men aged 25 to 30 years old within the Pathfinder Pilot areas. Using an age-based eligibility criterion is our second main source of identification and is all that is available after the pilot period. The age-based point estimates of the 4-months effect are very close and insignificantly different from those in row 1 using different areas. The linear matching estimator, for example, suggests a treatment effect of 10.4 percentage points when 25-30 year olds are used as the comparison group (row 3) compared to 11 percentage points when 19-24 year olds in control areas are used as a comparison group (row 1). It was emphasized in section 3 that this estimate is based on different assumptions from the estimates in rows 1 and 2. In fact, it may suffer from substitution more acutely and it is not immune to local labour market wide wage effects. However, it is informative to know that the obtained results are very similar, independently of the procedure used. We cannot reject the simple null hypothesis of a model without substitution and GE wage effects. Alternatively, their effects may cancel out, the relative sizes of the substitution and wage effects being very similar. We further test for substitution using the older group of 31 to 40s living in Pathfinder Pilot areas as control. This group is expected to be less substitutable for 19-24 year olds than the younger 25-30 year old comparison group. Under this assumption, and given that substitution exacerbates the impact of the programme, we would expect this estimate to be lower than the one presented in row 3. But the fourth row presents an estimate of the 4 months effect

¹⁷The appendix to this chapter, section 3.6.1, presents some comparisons between treatments and controls with respect to some of the covariates being considered. A few checks on the quality of the propensity score matching are also included.

of the Gateway that, if anything, is higher than the previously presented results. This is not consistent with large substitution effects. In rows 1 and 2 of table 3.2 we compare ineligible individuals living in Pathfinder Pilot and control areas. If there were significant substitution effects or differential trends across regions we may find differences in outflows in the NDYP period. In fact, no significant effects of the Gateway are found.

Finally, rows 3 and 4 in table 3.2 contain estimates of the employment effect in the “whole market”. Men aged 19 to 30 and 19 to 50 years old and living in Pathfinder Pilot areas are compared with similar individuals living in control areas. The results only confirm what has been established before: that, during the Pilot period, the programme had a very significant positive impact on outflows to employment on the markets it has been implemented. The point estimates are smaller because 19-24 year olds are only a fraction of the larger age range. For example, just over half the 19-30 year old group are 19-24 year olds. The linear matching estimator in row 3 implies a NDYP effect of 6.6 percentage points - as expected just over half the magnitude of the effect in row 1 of table 3.1.

It is interesting to check how sensitive these results are to historical patterns. The lack of information about destinations when leaving the claimant count before 1996 imposes the use of a different variable, *outflows to all destinations*, to perform this analysis. Figure 3.3 considers different types of comparisons and plots the estimates of non-existent programs over time. The first panel in the graph compares eligible individuals living in Pathfinder Pilot areas with eligible individuals living in all control areas. The size of the Gateway effect, represented by the last point in the graph, is well above all other estimates for previous periods. This is just more evidence that the effects of the programme on participants during the Pilot period are very positive. Panel 2 compares participants with eligible individuals living in matched control areas. It shows a similar pattern but with a stronger effect of the Gateway, which may be a consequence of the higher volatility observed. Panels 3 and 4 also confirm the importance of the estimated impact of the Gateway by comparing participants with older groups.

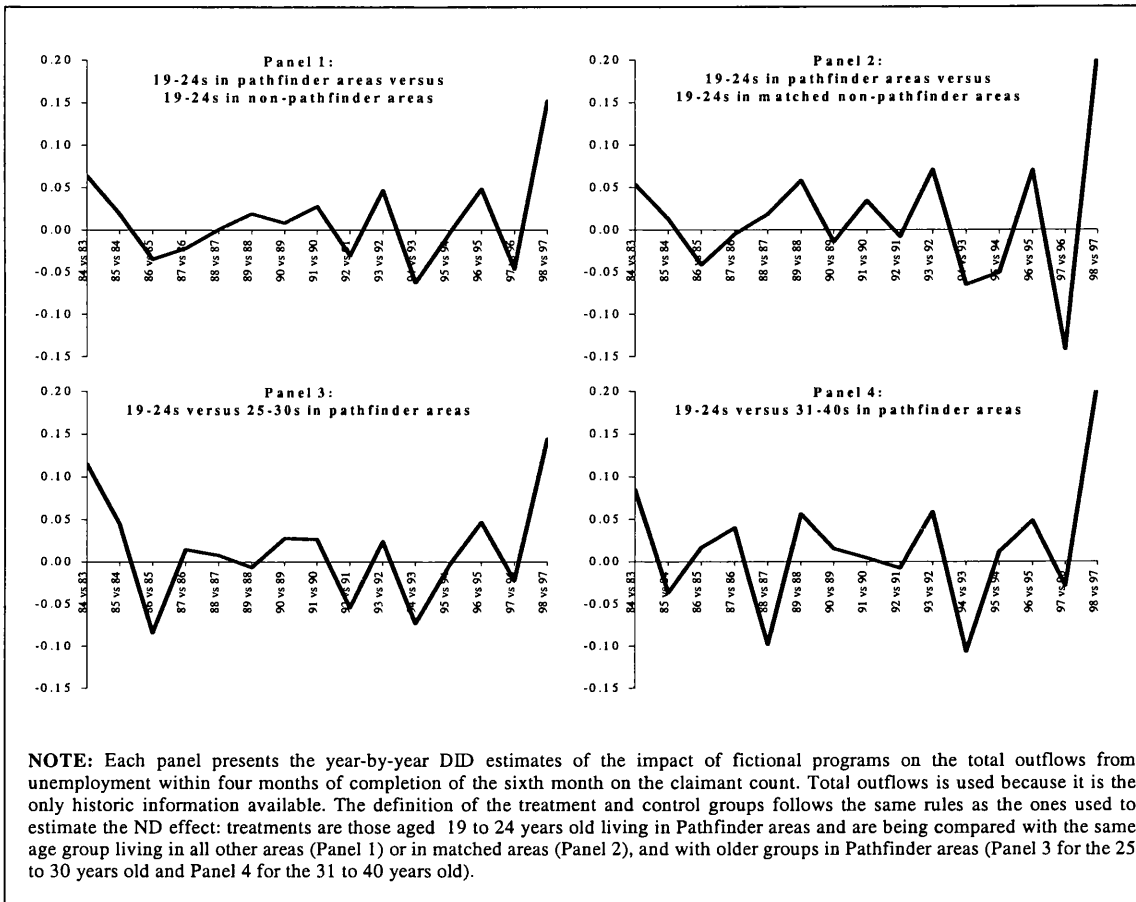


Figure 3.3: Difference in differences estimates over time. Outflows to all destinations. Men only.

3.4.2 National Roll Out: men’s results

Table 3.4 contains the main result from the National Roll Out. The first row shows an implied effect of around 5 per cent on a pre-programme base outflow (table 3.3) of 25.8 per cent, and once more, the method used does not seem to affect the result significantly. Although this is still a substantial impact, it is about half the magnitude estimated for the Pilot period. These differences in size can be accounted for by a *programme introduction* effect. In the first few months the programme is operating, a very large increase in the flows to employment is observed, which then falls as the programme matures. This is illustrated in the other rows of the table. The second and third rows report comparable estimates of the Gateway effect after 4 months of treatment for the first quarter the

Table 3.4: Effect of treatment on the treated men. Pilot and National Roll Out periods. Gateway employment effects by the end of the 10th month.

| Type of estimate | Observ. | Difference in differences combined with | | | |
|---|---------|---|--------------------------|-------------------------------------|--------------------------------------|
| | | (1) Linear matching | (2) Logit matching | (3) PS‡ matching | (4) PS‡logit matching |
| (1) TTE‡: Pilot period and the National Roll Out | 17,433 | 0.053** (0.013) | 0.044** (0.015) | 0.048** (0.015) (0.020;0.069) | 0.049** (0.016) (0.018;0.072) |
| Outflows to subsidised jobs | 55,051 | | | 0.039 | |
| (2) TTE‡: Pilot period (1st quarter in PF§areas) | 1,096 | 0.104* (0.055) | 0.091 (0.057) | 0.078 (0.079) (-0.050;0.195) | 0.074** (0.069) (-0.068;0.182) |
| Outflows to subsidised jobs | 4,486 | | | 0.057 | |
| (3) TTE‡: 1st quarter the NDYP operates in control areas | 5,169 | 0.088** (0.025) | 0.064** (0.027) | 0.078** (0.031) (0.021;0.131) | 0.075** (0.031) (0.015;0.128) |
| Outflows to subsidised jobs | 20,331 | | | 0.039 | |
| (4) TTE‡: 2nd and 3rd quarters the NDYP operates in all areas | 11,161 | 0.031* (0.016) | 0.023 (0.019) | 0.024 (0.019) (-0.008;0.051) | 0.027 (0.019) (-0.005;0.057) |
| Outflows to subsidised jobs | 30,234 | | | 0.036 | |

Notes: Estimates of the effects of the NDYP used the JUVOS 5% longitudinal sample of JSA claimants. Estimates of the outflows into employment option used the NDED. Selected observations are those completing a 6 month spell on JSA over a predefined time interval - the present table compares 1997 with 1998. These individuals are then followed up to the end of the 10th month on JSA to check whether they have found a job. The eligible group (defined by the age criterion) is compared with the control group before and after the release of the programme to estimate its impact. See notes to table 3.1 for further details on the estimation procedure.

** = significant at 0.05 level. * = significant at 0.10 level. †Treatment on treated effect. ‡Propensity score. §Pathfinder Pilot areas.

programme operates in the Pathfinder Pilot and control areas, respectively. As noticed before, estimates for the Pilot period (first quarter in Pathfinder Pilot areas) are about twice the size of the effect over the whole period. The same is also true if one considers the estimates for the first quarter the NDYP operates in control areas (see row 3). The fourth row presents estimates obtained using the following second and third quarters the programme is operating and these are comparatively much lower and less significant.

There are, of course, many possible explanations for this. One explanation is that

the agencies involved in delivering the programme are initially very enthusiastic, but this naturally erodes over time. Another possibility is that the programme diminishes fraud. This would have particularly important effects during the first few months after the release of the programme since potential participants are unlikely to be aware of the new claiming rules. Similar “cleaning up the register” effects have been noted of previous UK labour market reforms.¹⁸

There are many possible criticisms of the results. We shall now discuss some of the main ones - quality of job matches, selectivity and differential trends. How the programme affects the women will be discussed on the next section.

First, there is the issue of whether the quality of job matches has improved (or deteriorated) under the NDYP. One of the benefits from the NDYP is said to be that job matches are of higher quality due to greater job assistance and mentoring of the Personal advisor. For those who get onto the employer option there is a guarantee of one day a week training. On the other hand tougher monitoring may push claimants into low quality matches. Quality is difficult to measure without data on earnings and other job characteristics. One indicator of job match quality, however, is simply the longevity of a job. Following the government’s preferred measure, we define a “sustained” job as one that lasts at least thirteen weeks. The first row of Table 3.5 Panel A repeats the analysis but using the outflow to sustained jobs (instead of any job) as the outcome variable. The results are quite consistent with the earlier findings - the estimates point to an increase in the outflows to sustained jobs of 3-4%, which compares to estimates of around 5% for the outflows to all employment (first row of table 3.4).

Secondly, there is the issue of selectivity. It may be that the introduction of the NDYP has an effect on the (unobserved) quality of the inflow of individuals reaching 6 months of JSA. The most likely route for this is that claimants in the fifth or sixth months of JSA may alter their behaviour. If they believe the NDYP regime is *tougher* than the previous regime, they may be more likely to leave the unemployment rolls (this was one of the ways that RESTART, another job assistance programme introduced in 1986 was deemed to have worked). On the other hand, if the NDYP is seen as a desirable thing (e.g. because of subsidies to *good* jobs or training), then claimants may delay exit. If the main

¹⁸See Van Reenen (01) for a discussion of the RESTART programme and the introduction of the JSA.

Table 3.5: Robustness of the results.
Comparing 19-24s with 25-30s. Men only.

| | | Difference in differences combined with | | | | |
|---|--|---|--------------------|--------------------|-------------------------------------|-------------------------------------|
| | | (1) | (2) | (3) | (4) | |
| | | Linear | Logit | PS† | PS†logit | |
| Observations | | matching | matching | matching | matching | |
| <i>Panel A: Outflows to sustained jobs (conditional on being on JSA for 6 months).</i> | | | | | | |
| (1) | Estimates | 17,433 | 0.045** (0.011) | 0.031** (0.013) | 0.035** (0.013) (0.013;0.055) | 0.033** (0.016) (0.005;0.054) |
| | Outflows to sustained subsidised job | 55,051 | | 0.031 | | |
| <i>Panel B: Outflows to employment before the start of the Gateway (conditional on being on JSA for 4 or 5 months).</i> | | | | | | |
| (2) | Effect between months 5 and 6 of JSA | 20,957 | 0.004 (0.008) | 0.000 (0.010) | 0.004 (0.009) (-0.011;0.019) | 0.003 (0.010) (-0.013;0.020) |
| (3) | Effect between months 4 and 6 of JSA | 25,510 | 0.009 (0.010) | 0.001 (0.011) | 0.009 (0.011) (-0.011;0.026) | 0.009 (0.011) (-0.010;0.027) |
| <i>Panel C: Outflows to all destinations (conditional on being on JSA for 6 months).</i> | | | | | | |
| (4) | Estimates | 17,433 | 0.108** (0.015) | 0.093** (0.016) | 0.095** (0.018) (0.061;0.124) | 0.095** (0.018) (0.060;0.123) |
| (5) | Lower bound | | 0.084** (0.019) | 0.062** (0.020) | 0.048** (0.023) (0.010;0.087) | 0.046** (0.022) (0.010;0.084) |
| (6) | Upper bound | | 0.143** (0.019) | 0.119** (0.020) | 0.126** (0.024) (0.087;0.164) | 0.133** (0.026) (0.091;0.175) |
| | Outflows to all NDYP options | 55,051 | | 0.137 | | |

Notes: Estimates of the effects of the NDYP used the JUVOS 5% longitudinal sample of JSA claimants. Estimates of the outflows into employment option used the NDED. All estimates compare the eligible group (defined by the age criterion) with the selected control group before (1997) and after (1998) the release of the programme to estimate its impact. See notes to table 3.1 for further details on the estimation procedure.

Panel A refers to the stock of individuals completing a 6 month spell on JSA and follows them up to the end of the 10th month on JSA to check whether they have found a sustained job. A sustained job is one that lasts for more than 13 weeks. Panel B uses the stock of individuals completing either a 4 or a 5 month spell on JSA and follows them up to the end of the 6th month on JSA to check whether they have found a job. Panel C uses the stock of individuals completing 6 months of unemployment and follows them up to the end of the 10th month on JSA to check whether they have left at all. Upper and lower bounds are presented in Panel C using historical series of a similar parameter (see text for details).

** = significant at 0.05 level. * = significant at 0.10 level. †Propensity score.

effect is increased toughness, then we may underestimate the positive effects of the NDYP as there has been a decline in the unobserved quality of the stock (assuming the most

job ready decide to leap into jobs before they are pushed off the unemployment rolls). If the NDYP is perceived as more attractive than the previous regime (as the qualitative evidence suggests) then we may actually be overestimating the effects of the Gateway period as the more job ready actually delay their exits prior to entering the Gateway.

To investigate these selectivity problems we examine outflows to employment during the fourth and fifth month of JSA, using the same methodology as before. The results are presented in rows 2 and 3 of Table 3.5, Panel B. The introduction of the NDYP had no significant impact on the outflows to employment prior to six months duration. All the estimates are small and insignificant at conventional levels.

Thirdly, we have not controlled for differential trends. Using the same method as before (see section 3.4.1) we calculate upper and lower bounds for the NDYP effect on outflow rates. The average effect is again smaller than the estimates for the Pilot period (see rows 5 and 6 of Table 3.5, Panel C). Nevertheless, even at the lower bound there is a significant effect of the programme on the outflow rates to all destinations.

3.4.3 The impact of the programme on women

Finally, note that we have focused our results on male job outflow rates. Three quarters of all participants in the NDYP are men, but clearly the impact on women is also of great interest. The results for women are not as clear cut as those for men. This is mainly because there is a systematic trend in the labour market behaviour of older (25-30) compared to younger (19-24) women. The main problem, therefore, resides on the choice of the appropriate comparison group.

Figure 3.4 illustrates the difficulties encountered by plotting the conditional exits to all destinations against time for treatments and different possible control groups. It is apparent from the upper panel of Figure 5 that an estimator based on different age groups can be severely contaminated by differential trends. Compared to the younger age groups, the older age group seems to have systematically improved its position in the labour market over the 1982-99 period. If this trend extends to the treatment period, it is expected that such comparison under-estimates the impact of treatment on the treated. On the other hand, the lower panel of the graph suggests that the macro shocks seem to affect younger age groups living in different geographic regions much more similarly,

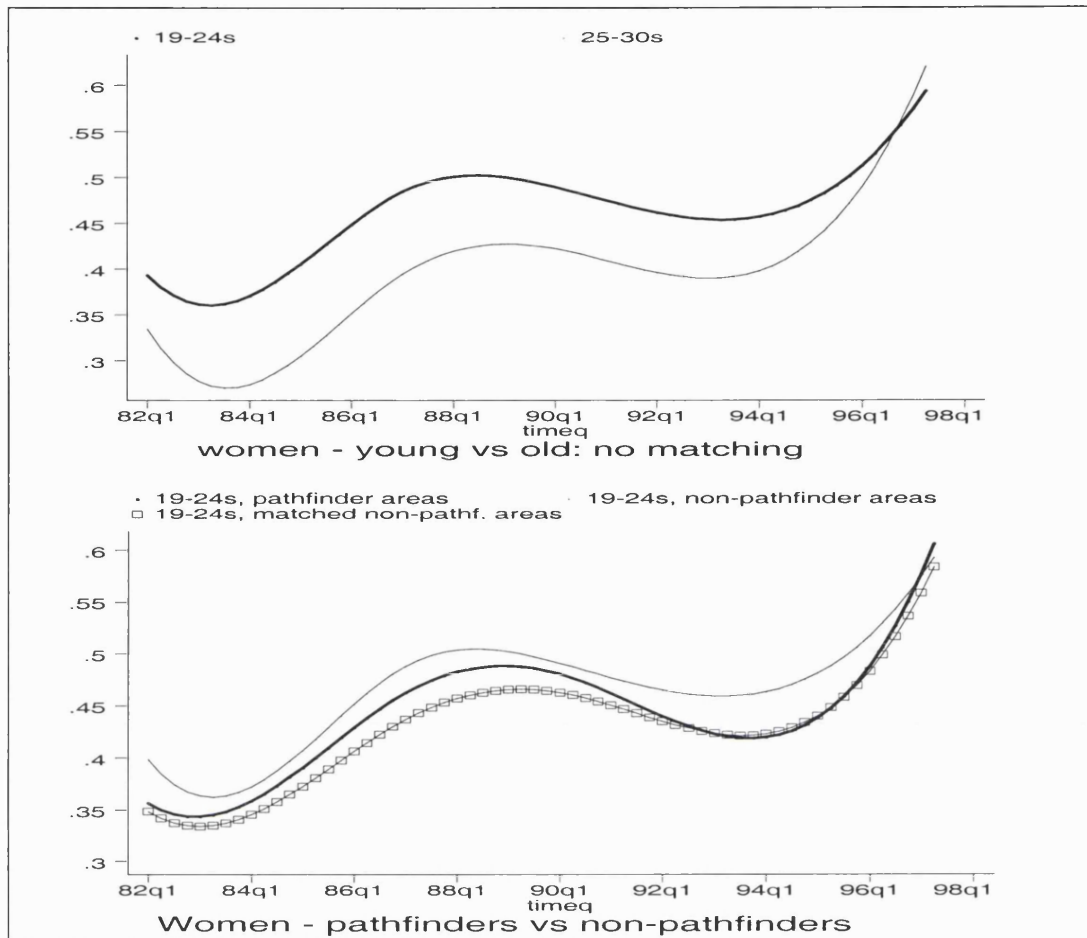


Figure 3.4: Women outflows from the JSA conditional on completing 6 months. Effect by the end of month 10.

making the Pathfinder - non-Pathfinder 19-24 year old groups comparable. Matching on regions improves the pattern, the two curves for treatment and comparisons being closer both in levels and in slopes. The upshot of this is that using older women as a comparison group is not valid, and we should focus on the pathfinder data to evaluate the effect of the NDYP for women.

Table 3.6 presents some estimates of the impact of the programme on treated individuals using different comparison groups and estimation techniques. All estimates resulting from the comparison of similar age groups point to a positive effect of the programme on the outflows to employment (see rows 1 and 2). These estimates are much less precise,

Table 3.6: Effect of treatment on the treated women. Pilot period.
Gateway employment effects by the end of the 10th month.

| | Control group | Observations | Difference in differences combined with | | | |
|--|--|--------------|---|--------------------------|-------------------------------------|-------------------------------------|
| | | | (1) Linear matching | (2) Logit matching | (3) PS† matching | (4) PS†logit matching |
| <i>Comparing 19-24s living in all Pathfinder Pilot areas</i> | | | | | | |
| (1) | 19-24s in all control areas | 1,592 | 0.061 (0.058) | 0.026 (0.060) | 0.057 (0.084) (-0.073;0.219) | 0.051 (0.083) (-0.096;0.190) |
| (2) | 19-24s in matched control areas | 596 | 0.025 (0.071) | 0.013 (0.077) | 0.136 (0.151) (-0.106;0.374) | 0.113 (0.149) (-0.162;0.334) |
| (3) | 25-30s in all Pathfinder Pilot areas | 400 | -0.047 (0.100) | -0.057 (0.101) | -0.053 (0.213) (-0.447;0.270) | -0.080 (0.193) (-0.449;0.219) |
| | Outflow into the employment option | 1,693 | | | 0.048 | |

Notes: See notes to table 3.1.

** = significant at 0.05 level. * = significant at 0.10 level. †Propensity score.

more sensitive to the estimation technique used and generally smaller, but do not seem to reject the conclusions drawn for men. For example, the linear matching estimator in row 1 suggests an impact effect of 6.1 per cent compared to 11.0 per cent for men. The lack of precision is likely to be a consequence of the smaller sample sizes. Notice that the increased job taking-up rate seems to be mainly accounted for by the employment option, which ensured a job to almost 5 per cent of the treated during this period. As expected, comparing different age groups changes the results drastically and in the predicted direction (see row 3): despite remaining statistically insignificant, the estimates are actually negative. Together with the pattern depicted in figure 3.4, this explains why the women's case is not explored during the National Roll Out of the programme. The only group we can draw comparisons from is composed of individuals older than the participants, and these are subject to very differential trends.

3.5 Conclusions

This paper has examined the labour market impact of the British NDYP programme. The NDYP is a compulsory programme affecting all young people claiming unemployment benefit for at least six months. The programme offers a combination of treatments, particularly job assistance for four months and a wage subsidy paid to employers. Two sources of identification are used to construct comparison groups in order to make inferences on the impact of the NDYP: a comparison between Pathfinder Pilot and control areas and an age-related eligibility criteria. Our results suggest similar quantitative effects whichever comparison group is chosen.

Based on the Pilot period of the programme, we find an economically and statistically significant effect of the programme on outflows to employment among men. The programme appears to have caused an increase in the probability of young men (who had been unemployed for 6 months) finding a job in the next four months. On average, this increase is about 5 percentage points (relative to a pre-programme baseline of 26 per cent). Part of this overall effect is the job subsidy element and part is a pure “Gateway” element (enhanced job search). We estimate that at least 1 percentage point of the 5 percentage points is due to the Gateway services, such as job search assistance. We also found that the treatment impact is much larger in the first quarter of introduction. This puts in question whether the effects of this aspect of the programme will be sustained in the long run. Our findings are robust to a large number of experiments, including a number of different comparison groups.

3.6 Appendix to chapter 3

3.6.1 Data

Table 3.7 compares the mean values of some of the independent variables used in the analysis before and after matching on the propensity scores.¹⁹

It can be observed that similar age groups are much more alike, at least with respect to the considered characteristics (compare columns 1 and 2 with 5 and 6). Moreover,

¹⁹Other comparisons are available and can be provided under request.

Table 3.7: Descriptive statistics for different treatment and control groups.
Men only.

| | 19-24s in PF† vs. 19-24s in control areas. Pilot period. | | | | 19-24s vs. 25-30s in all areas. 1st 3 quarters. | | | |
|---|--|-------|-----------------|-----|---|-------|-----------------|-------|
| | No matching | | Matching on PS§ | | No matching | | Matching on PS§ | |
| | T† | C† | T | C | T | C | T | C |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Observations | 273 | 1,306 | 264 | 264 | 4,377 | 4,086 | 4,359 | 4,359 |
| Married | .08 | .10 | .06 | .08 | .10 | .23* | .10 | .10 |
| Time unemployed over the last 2 years | | | | | | | | |
| Less than 6 months | .46 | .48 | .46 | .49 | .43 | .33* | .43 | .46* |
| Less than 12 months | .64 | .66 | .64 | .66 | .64 | .51* | .64 | .64 |
| Number of unemployment spells over the last 2 years | | | | | | | | |
| 0 | .29 | .26 | .29 | .31 | .20 | .16* | .20 | .20 |
| 1 to 2 | .59 | .56 | .58 | .57 | .58 | .64* | .58 | .59 |
| 3 to 5 | .12 | .17* | .12 | .10 | .21 | .18* | .21 | .19* |
| 6 or more | .01 | .01 | .01 | .01 | .01 | .01 | .01 | .01 |
| Sought occupation | | | | | | | | |
| Manager | .03 | .02 | .03 | .04 | .02 | .02* | .02 | .02 |
| Professional | .01 | .02 | .01 | .02 | .01 | .03* | .01 | .01 |
| Technical | .07 | .07 | .07 | .11 | .05 | .07* | .05 | .06 |
| Clerical | .12 | .17* | .12 | .15 | .18 | .13* | .18 | .18 |
| Craft | .19 | .12* | .19 | .17 | .14 | .17* | .14 | .14 |
| Personal services | .11 | .08 | .11 | .09 | .07 | .06* | .07 | .06 |
| Sales | .10 | .10 | .10 | .10 | .11 | .07* | .10 | .10 |
| Machine operator | .07 | .09 | .07 | .07 | .10 | .14* | .10 | .11 |
| Other | .29 | .31 | .28 | .24 | .31 | .29* | .31 | .32 |
| Region | | | | | | | | |
| South East | .19 | .26* | .19 | .19 | .24 | .30* | .24 | .26* |
| East Anglia | .00 | .00 | .00 | .00 | .02 | .02 | .02 | .02 |
| South West | .08 | .06* | .08 | .09 | .05 | .05 | .05 | .05 |
| West Midlands | .17 | .09* | .17 | .19 | .10 | .08* | .10 | .09 |
| East Midlands | .04 | .06* | .04 | .04 | .07 | .07 | .07 | .07 |
| York | .12 | .11 | .12 | .13 | .12 | .11 | .12 | .11 |
| North West | .07 | .17* | .06 | .04 | .15 | .14 | .15 | .15 |
| North | .16 | .07* | .16 | .11 | .08 | .07 | .08 | .08 |
| Wales | .13 | .06* | .13 | .16 | .06 | .05 | .06 | .05 |
| Scotland | .05 | .11* | .05 | .05 | .12 | .11* | .12 | .11 |

*Estimated mean for treatments and controls are significantly different at 5% level.

†T and C stand for the treatment and control groups, respectively. ‡Pathfinder Pilot areas. §Propensity score.

matching on the propensity scores significantly improves the similarity between the groups (compare columns 3-4 with 1-2 or columns 7-8 with 5-6).

A more detailed diagnosis of the quality of the propensity score matching is presented

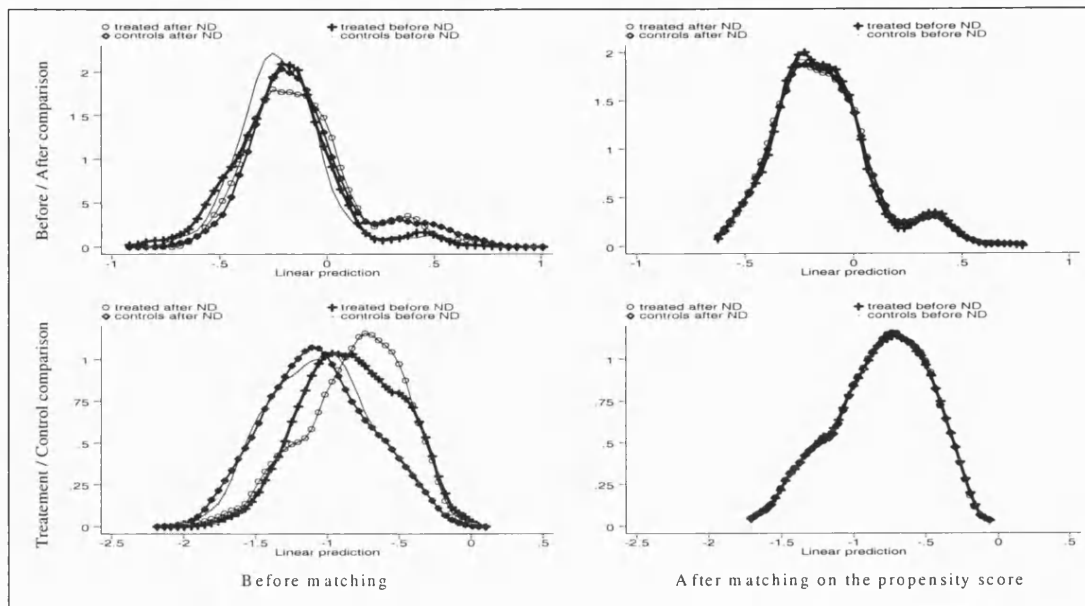


Figure 3.5: Distribution of the PS before and after matching: men aged 19-24 in Pathfinder Pilot and control areas.

in figures 3.5 to 3.8. These plots represent the distribution of the two propensity scores used in the matching process over the entire population and over specific subgroups. We compare 19 to 24 years old living in Pathfinder Pilot areas with 19 to 24 years old in all control areas during the pilot period. All groups being included in the analysis are plotted: treatments and controls, before and after the release of the NDYP. As expected, matching significantly improves the similarity between the curves - it can be observed that the curves on the right hand side of figure 3.5 overlap almost precisely. Moreover, nearly all the initial support is maintained after matching. Figures 3.6 to 3.8 give some indications of how identical the distributions of the propensity scores are over sub-groups of the population. It is apparent that matching worked well even over sub-populations, making the distributions quite similar. Very similar results were obtained when using other groups and are available under request.

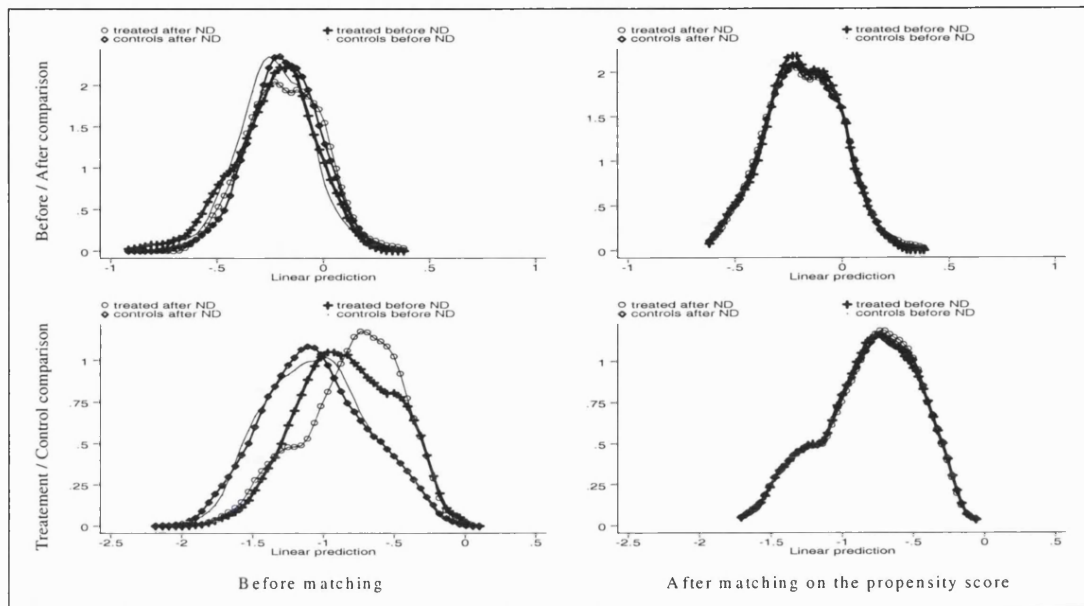


Figure 3.6: Distribution of the PS before and after matching: single men aged 19-24 in Pathfinder Pilot and control areas.

3.6.2 Gateway employment effects under different propensity score matching techniques

Tables 3.8 and 3.9 present estimates for the employment effects of the Gateway among men during the Pilot period using three possible variations of the propensity score matching method under the linear specification assumption. Columns (1) to (3) present propensity score matching estimates of the parameters presented in tables 3.1 and 3.2 in the main text. Column (1) displays the estimates for the standard nearest neighbour propensity score method, where only one observation from each comparison group is chosen to match each observation in the treatment group - the closest one from the perspective of the two propensity scores at use. Column (2) uses the same method as in column (1) but smoothes the outcome of the comparison group. The same comparisons are chosen but the smoothed outcome is used to estimate the impact of the programme. Column (3) uses kernel weights to select the counterfactual for each treatment observation: controls that are relatively near the treatment observation in terms of the propensity scores are given a weight depending on how close they are. These estimates used an Epanechnikov

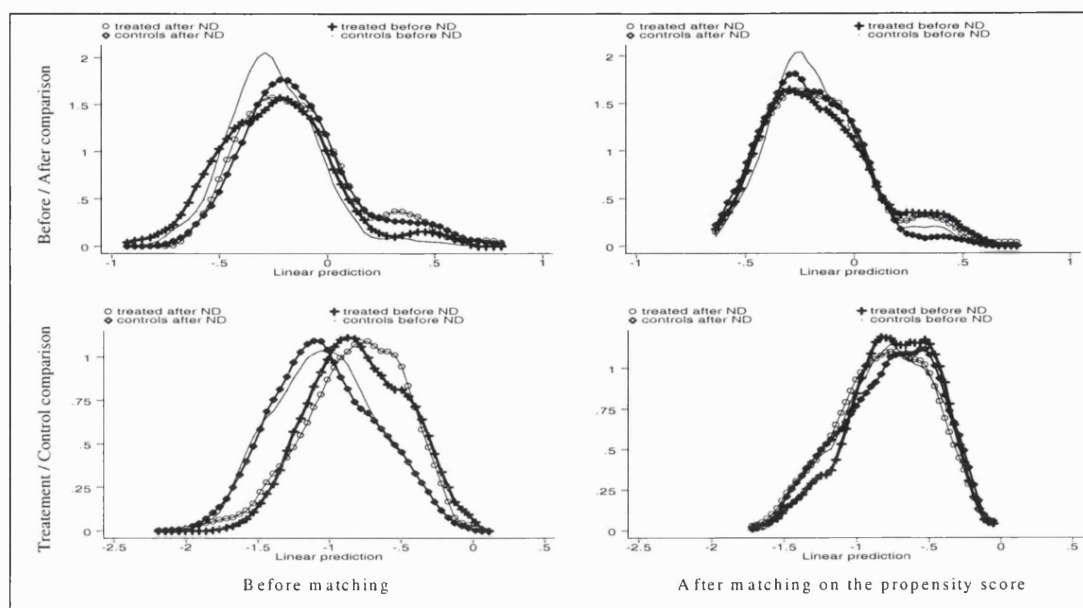


Figure 3.7: Distribution of the PS before and after matching: men w/ less than 6 months on unemployment in the last 2 years in Pathfinder Pilot and control areas.

function with a diagonal matrix of bandwidths. The main result from tables 3.8 and 3.9 is that all methods produce similar estimates, and this remains true when comparing with the numbers in tables 3.1 and 3.2 in the main text. However, the precision of the estimates does change from method to method. The estimated standard errors presented in column (1) are much higher than similar estimates produced by other methods. The strong variation resulting from the fact that only one observation is being chosen as a control for each treated individual is in part to blame. The standard errors presented in column (3) are significantly lower but still too high to sustain a definitive conclusion. Estimates in column (2), however, are generally more precise, the result being due to the smoothing of the counterfactual outcomes.

3.6.3 Estimation methods

The practical implementation of the completely parametric methods is discussed in the main text, and so we omit it here.

We use propensity score matching based on two dimensions, time and eligibility, and

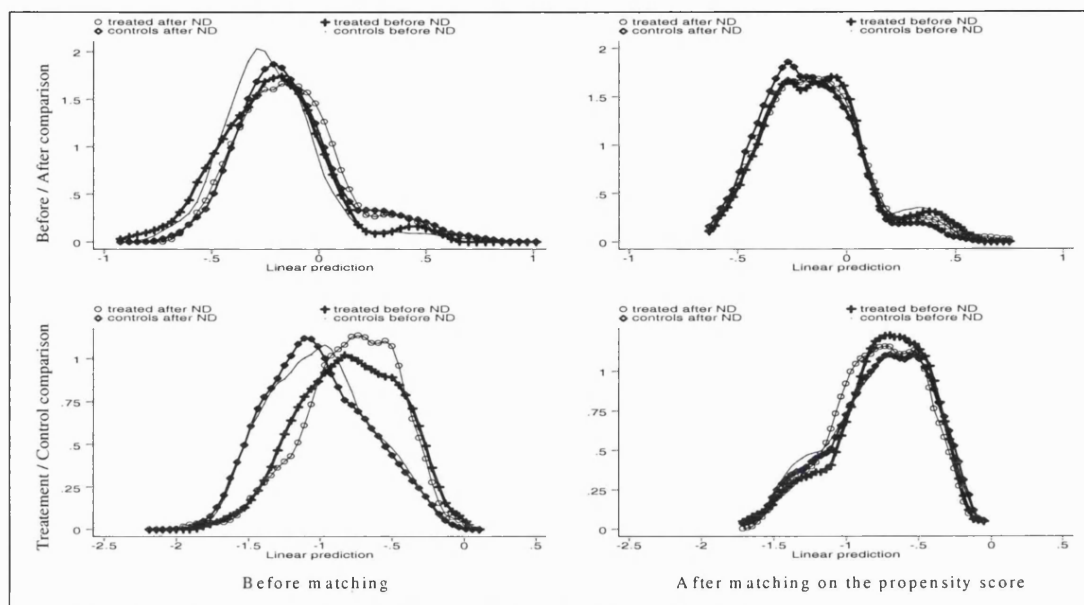


Figure 3.8: Distribution of the PS before and after matching: men with less than 2 unemployment spells in the last 2 years in Pathfinder Pilot and control areas.

using either the nearest neighbour method or smoothing the outcomes applying splines or kernel weights. With the same set of observables used in the completely parametric estimates, we compute the two propensity scores, $P_{EX} = P(ND = 1 | X)$ and $P_{tX} = P(t = 1 | X)$.

In the *nearest neighbour* case, each treated individual is paired with one observation from each of the three control groups, the one that minimizes the Euclidean distance with respect to the two propensity scores conditional on two maximum distance restrictions, one for each dimension. Matching is done with replacement, meaning that each control may be chosen more than once and is weighted accordingly.

Under the *smoothing splines* method, we run a regression of the outcome of interest on a cubic polynomial of the two propensity scores for the control groups. Predictions of the outcome under the three non-treatment cases for each of the matched treated observations under the nearest neighbour method are then computed and used to estimate the impact of treatment.

The use of *kernel weights* to select each of the three control groups is based on the

Table 3.8: Effect of treatment on the treated men. Pilot period.
Gateway employment effects by the end of the 10th month.

| | | Difference in differences combined with propensity score matching | | | |
|--|------------------------------------|--|------------------------------------|-------------------------------------|------------------------------------|
| | | (1) | (2) | (3) | |
| Control group | Observations | nearest neighbour | nearest neighbour w/ smoothing | kernel weights | |
| <i>Comparing 19-24s living in all Pathfinder Pilot areas</i> | | | | | |
| (1) | 19-24s in all control areas | 3,716 | 0.110 (0.083) (-0.028,0.238) | 0.104** (0.046) (0.024;0.182) | 0.078 (0.056) (-0.010;0.170) |
| (2) | 19-24s in matched control areas | 1,193 | 0.084 (0.100) (-0.076,0.245) | 0.093 (0.073) (-0.015;0.226) | 0.070 (0.068) (-0.043;0.183) |
| (3) | 25-30s in all control areas | 1,096 | 0.069 (0.112) (-0.117,0.248) | 0.078 (0.079) (-0.050;0.195) | 0.054 (0.081) (-0.083;0.191) |
| (4) | 31-40s in all control areas | 1,169 | 0.089 (0.129) (-0.116,0.307) | 0.099* (0.078) (-0.015;0.231) | 0.094 (0.078) (-0.034;0.227) |

Notes: See notes to table 3.1.

** = significant at 0.05 level. * = significant at 0.10 level.

Epanechnikov function and a diagonal matrix of (constant) bandwidths, each element of the diagonal being given by $1.06\sigma_x n^{-1/5}$.

Having constructed the three counterfactuals, the simple difference in difference method is applied to estimate the effect of the programme under the assumption of separable additivity of the group and time effects. We also transform the outcome applying the logit transformation, as shown in equations 3.4 and 3.3, to estimate the impact of the ND under a non-linear specification.

Table 3.9: NDYP effect on the ineligible and the whole economy. Pilot period.
Gateway employment effects by the end of the 10th month. Men only.

| | Control group | Observations | Difference in differences combined with propensity score matching | | |
|--|---------------------------------|--------------|---|---------------------------------------|-------------------------------------|
| | | | (1) nearest neighbour | (2) nearest neighbour w/ smoothing | (3) kernel weights |
| <i>Comparing 25-30s living in all Pathfinder Pilot areas</i> | | | | | |
| (1) | 25-30s in all control areas | 3,180 | 0.016 (0.092) (-0.149,0.164) | 0.027 (0.049) (-0.058;0.107) | 0.015 (0.063) (-0.079;0.130) |
| (2) | 25-30s in matched control areas | 983 | -0.016 (0.126) (-0.220,0.185) | -0.003 (0.066) (-0.107;0.112) | -0.028 (0.081) (-0.167;0.105) |
| <i>Comparing 19-30s living in all Pathfinder Pilot areas</i> | | | | | |
| (3) | 19-30s in all control areas | 6,896 | 0.033 (0.058) (-0.058,0.132) | 0.058* (0.034) (0.004;0.114) | 0.051 (0.041) (-0.019;0.118) |
| <i>Comparing 19-50s living in all Pathfinder Pilot areas</i> | | | | | |
| (4) | 19-30s in all control areas | 12,749 | 0.025 (0.042) (-0.053,0.094) | 0.044* (0.023) (0.004;0.080) | 0.023 (0.026) (-0.025;0.063) |

Notes: See notes to table 3.1.

** = significant at 0.05 level. * = significant at 0.10 level.

Chapter 4

The dynamics of earnings in a life-cycle model of labour supply

Job earnings play a central role on the evaluation of labour market programmes, being responsible for many of the individuals' labour market choices as whether or not to invest in education, look for a job or enrol into training. The aim of most of these programmes is to improve working perspectives of participants, reducing welfare dependence. This is usually attempted by the provision of training or educational courses or by the introduction of a direct stimulus to take up a job such as job-search assistance, sanctions or wage subsidies. Such interventions are expected to work by increasing productivity and, therefore, potential earnings among the treated, making future employment more desirable. Thus, by unveiling important information about the structure of the decision process, earnings are fundamental to explain labour supply behaviour and form an accurate idea of the possible impact of social interventions.¹

However, to correctly assess the value of a given treatment one should look at a long period of working decisions and earnings. The lack of empirical studies that accomplish such analysis just reflects the inexistence of appropriate data and the public pressure to produce policy assessments soon after their release. There are, however, different reasons

¹See, for example, the extensive review of the labour supply literature by Blundell and MaCurdy (99). The authors analyse aspects of labour supply modelling in close interlink with policy reforms, and devote a great deal of attention to the structural modelling of working decisions.

for why very short-run evaluations may be misleading. On the one hand, differences in labour market experiences may be of large importance soon after treatment but have minor impact in the long-run. For instance, a given period on a job may be enough to overcome the differential in human capital resulting from a given treatment or a previous stigmatising unemployment experience. Treatment, on the other hand, may open opportunities to participants as they improve skills and human capital, creating the possibility for future improvements and gains not observed in the short-run.²

Under uninsurable uncertainty, medium to long run evaluations continue to be required. In such environment, agents labour market decisions are likely to be influenced by the amount of risk involved and the agent's taste for it. Within the human capital theory, the payoff to labour market investments is expected to occur with some time delay as skills and human capital build up. But then, so do the returns to treatment that (partially) insure against this sort of risks. That is, it takes time for the options being compared to be revealed. The reverse implies that measuring the amount of risk involved in certain labour market options must make use of individual information covering large periods of life.

It has been noticed before that uncertainty about future income strongly influences some of the agents' economic decisions. This relationship has been particularly explored in the literature concerning consumption, the nature of savings and the importance of precautionary savings (see Kimball, 90; Carroll, 94; Attanasio, 99; Banks, Blundell and Brugavini, 99). Studies on the variability of earnings over the life-cycle and the dynamics of earnings variance have been frequently motivated by these applications (see, for example, Meghir and Pistaferri, 01). There are, however, some analysis concerning labour market decisions. An early example is Eckstein and Wolpin (89), who estimate a dynamic model of labour market participation with working experience. However, this is done under the assumption that uncertainty does not influence participation. More recently, Altug and Miller (98) present a life-cycle model of labour supply to analyse how experience and working hours affect current wages and employment within complete markets. Low (99)

²Sianesi (02) looks at the effects of policies in Sweden for a period of 5 years after the programme takes place. The identified impact for sub-groups of the population on a number of outcomes which include the employment probabilities are shown to change with the time distance from treatment.

studies the effects of non-market insurance on life-cycle decisions about labour supply and savings, but does not attempt to model human capital or earnings.

In this study, we develop and estimate a life-cycle model of labour supply, human capital and earnings. The model is sufficiently simple to be workable in a more complete setup of the economy, which is the subject of the next chapter together with the evaluation of the overall impact of labour market programmes. It is assumed that heterogeneous individuals decide about working in a learning-by-doing technology for human capital, consistent with results presented in Cossa, Heckman and Lochner, 99.³ Agents are endowed with different skills, possibly not perfectly substitutable in production and commanding jobs with different learning contents. It is explicitly acknowledged that labour market participation decisions are taken under uncertainty, potentially preventing/inducing working decisions.⁴ In each time period, an idiosyncratic productivity shock arrives, which determines how much effort workers are to put in work if they decide to do so. This is a transitory wage shock, which has long been acknowledged as an important component of the treatment selection process: the so-called Ashenfelter dip is a recurrently observed characteristic of the participants' wages data (see Ashenfelter and Card, 85, Bassi, 83, 84, Smith, 97, Heckman LaLonde and Smith, 99 and Heckman and Smith, 99). Notwithstanding, it is allowed to have a permanent component through human capital accumulation, consistent with the learning-by-doing setup where working is productive for the learning process as well. As such, the permanent impact of the productivity shock reflects the fact that the more the agent works, the more he/she is able to learn. Consequently, enrolment into a potential programme is primarily determined by the agent's idiosyncratic labour market history, which explains his/her skill-specific knowledge. This is consistent with the evidence presented in Heckman and Smith, 99, on the importance of unemployment dynamics

³The authors find this approach to be more consistent with the data than the rivalrous activities that compose the on-the-job training model.

⁴The importance of uncertainty in explaining labour market dynamics and its dependence on characteristics such as ability, age, skills and knowledge has been noticed before (see Carroll and Samwick, 97). The relationship between the amount of risk and the individual characteristics is explicitly considered in the present analysis as human capital provides insurance against the risk. Moreover, taking risk into account is also consistent with the considerable mobility observed in the labour market over the whole business cycle and affecting all groups (see Davis, Haltiwanger and Schuh, 96).

in explaining participation.

The one being presented here is a much simpler model than others developed in the literature of earnings' dynamics (recent examples are presented in Hubbard, Skinner and Zeldes, 94; Carroll and Samwick, 97 and Meghir and Pistaferri, 01). We do not attempt to model the dynamics in earnings volatility or allow for heterogeneous earnings' variance. Contrary to most of these studies, however, the main focus of our model is on human capital accumulation, which is endogenously modelled and different from experience. Conditional on individual's characteristics and together with the idiosyncratic innovation disclosed in each period, human capital determines participation.

The rest of this chapter goes as follows. It starts by presenting the individual's dynamic model of labour supply with skills and human capital in section 1. It will be clear from the discussion that many of the design decisions are closely linked to the primer goal of this thesis, namely to assess the impact of labour market interventions of the New Deal for the Young People (NDYP) type. Section 2 discusses the characteristics of the model that will shape the estimation procedure discussed in section 3. The estimation strategy adopted here heavily explores the structural model in order to extract the conditions required to identify the parameters. As will be shown, such procedure avoids computationally expensive evaluations of complicated likelihood functions required by some other approaches (see Blundell and MaCurdy, 99). Section 4 presents the data used in the estimation procedure and section 5 discusses the obtained estimates, attempting some sensitivity analysis. Section 6 draws some conclusions and the appendix to this chapter is in section 7.

4.1 A human capital model of labour supply

Given the stated goal of evaluating labour market policies, we develop a model of labour supply and human capital formation with heterogeneous skills. The structure of the model goes as follows. Agents are assumed to live for a fixed number of time periods, A , evaluating at each period the state of the nature and deciding optimally about activity and consumption. Labour market activity comprises working and staying at home, the two

options being mutually exclusive.⁵ Education is the source of different skills, potentially non-substitutable in production and commanding jobs with different learning contents. Education attainment is not endogenous in this version of the model, which looks solely at the working decisions. Agents are assumed to enrol in the labour market with a certain level of education which is kept unchanged throughout their working life. This assumption is relaxed in the next chapter, where schooling is an additional labour market option available to the agents. While at work, the agents accumulate human capital through learning by doing. The accumulated human capital is productive in future periods in jobs of the agent's specific skill only. Consumption is assumed to be of an homogeneous good and is where utility is derived from. A dynamic decision process is considered, where the dynamics comes both through savings and human capital accumulation. Finally, the agent is assumed to be rational, making decisions based on an inter-temporal utility function.

We consider different heterogeneity dimensions. First, agents are born different with respect to their ability to learn in each type of job: some individuals may have a comparative advantage from working in low(high)-skilled jobs while others may be equally suited to all types of jobs.⁶ Second, agents are also born different with respect to the type of skills owned, characterised by the educational attainment. Third, markets are incomplete by the existence of idiosyncratic uninsurable uncertainty. We consider an *idiosyncratic* productivity shock arising in each period of the agent's life and affecting current earnings and future levels of human capital would he/she decide to work.⁷ It is modelled as a transitory wage shock, which has long been acknowledged as an important component of the treatment selection process: the so-called Ashenfelter dip is a recurrently observed characteristic in wages' data among participants in labour market programmes (Heckman and Smith, 99). Notwithstanding, the transitory wage shocks being considered here have

⁵A discrete decision process for the labour market activity is consistent with evidence describing the participation status as explaining the largest share of the variability in individual hours worked over time and most of the shape of labour supply as a function of wages (Pencavel, 86; Blundell and MaCurdy, 99; Browning, Hansen and Heckman, 99).

⁶Sattinger (93) shows evidence that the earnings' distribution changes with skills. See also Heckman, Lochner and Taber (98a).

⁷In what follows, "productivity shock" always refers to the individual-level innovation as no aggregate uncertainty is considered in the present model.

a permanent impact through human capital accumulation. This is consistent with the learning by doing structure of the model, where working is productive for the learning process as well.⁸ Thus, we denominate this shock as productivity shock and its permanent impact reflects the fact that the more effort is put in work, the larger the outcome in terms of learning. Working is, therefore, a potential risky decision, and labour market choices are made under risk aversion.⁹ On the job accumulated knowledge insures against future shocks, making labour market experience an important determinant of participation both in the labour market and in eventual social programmes.¹⁰ At the aggregate level, however, no uncertainty is considered and there is perfect foresight of future prices.

The agent's state and decisions can be summarised by a set of variables. As discussed, the two decision variables are consumption and labour market activity, and choices over them will determine the evolution of some of the state variables, namely assets and human capital. Experience affects the amount of human capital through learning by doing but has no effect on its type so that the resulting human capital is perfectly substitutable in production for the previous one. The agent is only able to supply one type of human capital determined by the type of skill he/she is endowed with. This is characterised by the educational level, which is assumed non-changeable throughout individuals' life.

4.1.1 The decision rule

Lets assume the agent maximises an inter-temporally separable lifetime utility function. Let u denote each period utility function, dependent on contemporaneous consumption only (c). Under the present setting, c is a continuous decision variable and the utility function is assumed to be strictly increasing and concave in c . On its turn, the working status (d) is modelled by a discrete decision variable as only the participation decision is being modelled. It takes the value 1 whenever the agent decides to take up a job and is 0

⁸Cossa, Heckman and Lochner (99) find this approach to be more consistent with the data than the rivalrous activities that compose the on-the-job training model.

⁹Empirical evidence suggests that a large proportion of the population is very risk averse (Barsky, Juster, Kimball and Shapiro, 97) and precautionary savings have been found to constitute the largest component of total savings (Carroll, 94; Attanasio, 99; Banks, Blundell and Brugavini, 01).

¹⁰This is consistent with recent evidence on the determinants of programme participation, which draws attention for the importance of past individual labour market dynamics (Heckman and Smith, 99).

otherwise. By excluding leisure from the utility function, we are outruling wealth effects as, *ceteris paribus*, richer agents or those more productive will have no incentive to “buy” extra leisure time. Under homogeneous preferences, such assumption affects mainly the decisions of agents at the top of the wealth distribution, the group of least interest when studying the impact of interventions targeted at the most disadvantaged. As described below, fixed costs from working and uncertainty about future returns from such investment are the causes of unemployment in this model. Thus, incentives to move unemployed back to work should deal with the fixed costs from taking up a job.

Decisions are based on the state of the nature the agent faces. In each time period t the state space faced by an agent aged a ¹¹ is described by the ability type of the agent (θ), the amount of assets (k), the level of schooling (s) which determines the type of human capital (skills) and is assumed unchangeable throughout individuals’ working life,¹² the amount of human capital (h) and the idiosyncratic productivity shock (π). The latter is the only random component of the model and is assumed to be iid. Together, they represent the state space faced by agent i and are denominated by X ,

$$X_{ita} = (\theta_i, s_i, k_{ita}, h_{ita}, \pi_{ita})$$

The individual’s problem at age 1 for an agent i of type θ with schooling level s and starting his/her working life at time t' can, therefore, be written in the following way,

$$\max_{\{c,d\}_{a=1}^A} E \left\{ \sum_{a=1}^A \beta^a u(c_{ita} | X_{ita}) \middle| (R_t, W_{st}, B_t, \tau_t)_{t=t'}^{t'+A-1} \right\} \quad (4.1)$$

where β is the discount factor, and $(R_t, W_{st}, B_t, \tau_t)_{t=t'}^{t'+A}$ represents the sequence of prices faced over the agent’s entire life - R_t stands for the interest rate, W_{st} is the wage rate of skill s , B_t is the unemployment insurance and τ_t is the tax rate imposed on earnings. All prices are expressed in real terms and known by the beginning of the agents life as perfect foresight is being assumed. Equation (4.1) establishes the optimisation problem: faced with a state of the nature, the risk averse agent makes optimal decisions about consumption and working. The expectations refer solely to the productivity shock, π ,

¹¹ Age a and time period t differ by a constant for each individual i .

¹²Such assumption is to be relaxed in the next chapter, where individual educational decisions are made endogenous.

which affects the current productivity and the speed of human capital accumulation as described below.

Risk may affect the working decisions when relatively large fixed costs apply. This will be clear from the following description of the dynamics of the model as working is an investment with future payoffs in terms of accumulated human capital. Thus, even when the contemporaneous earnings do not pay for the fixed costs of taking up a job, future returns to the investment in terms of improved working perspectives and earnings might.

4.1.2 The dynamics of the state variables

To complete the set up of the problem, one must establish how the state variables evolve over time. Lets start by the level of assets. Individual's wealth accumulation follows the rule,

$$k_{i,t+1,a+1} = (1 + R_t)k_{ita} + d_{ita}h_{ita}\pi_{ita}W_{st}(1 - \tau_t) + (1 - d_{ita})B_t - c_{ita} \quad (4.2)$$

where the first term on the rhs represents accumulated savings, the second term stands for net earnings, the third term is the income if not working and the fourth term is consumption. The earnings of a working agent in each period depend on the type of skills and amount of human capital, s and h respectively, on the idiosyncratic transitory wage shock, π , and on the tax rate, τ . W_s is, therefore, the price of one unit of human capital of type s . An unemployed agent is paid an unemployment benefit, B . As usual, total savings equals total income net of total consumption. The budget constraint corresponds to the restriction $k_{i,t'+A,A+1} \geq 0$.

Human capital is accumulated through experience. The agent's working life starts with a given level of skills, corresponding to the individual's educational attainment, and an initial amount of human capital which is assumed to be ability- and skill-specific. That is, human capital is assumed to be characterised by the level of education, which defines different skills assumed to be potentially not perfectly substitutable in production. Moreover, individual's skills are assumed to be unchangeable throughout the working life. Accumulation of human capital occurs while working at a rate that depends on the agent's level of ability, type of skills and previously accumulated human capital. Given that a learning by doing technology is being assumed, the rate of human capital accumulation

is adjusted by the individual- and time- specific productivity shock, π . Such feature establishes a dependence of the learning outcome on the effort put in work. Moreover, it also allows for a permanent effect of the productivity shock despite such innovation being assumed uncorrelated over time. The permanent effect arises through human capital accumulation. Finally, no depreciation of human capital is assumed if the agent decides to stay at home. Thus, the amount of human capital evolves as follows,

$$\begin{aligned} h_{i,t+1,a+1} &= h_{ita} (1 + \nu(\theta, s, h_{ita}) \pi_{ita}) && \text{if } d_{ita} = 1 \\ h_{i,t+1,a+1} &= h_{ita} && \text{if } d_{ita} = 0 \end{aligned} \quad (4.3)$$

where the parameter $\nu(\theta, h, s)$ stands for the agent's ability to learn the s -type of skills.

Finally, all individuals are equal at birth apart from the characteristic θ which determines their on-the-job learning ability and the level of skills, s : they all live for the same number of periods, A , being endowed with the same amount of human capital within (θ, s) -group, and the same amount of assets, $k_0 = 0$. The conditions at birth can be described as,

$$k_{it'1} = 0 \quad (4.4)$$

$$h_{it'1} = \underline{h}^{\theta s} \quad (4.5)$$

Under the dynamics just discussed, the recursive version of the problem can now be written as

$$V_{ita}^{\theta s}(k_{ita}, h_{ita}, \pi_{ita}) = \max_{c_{ita}, d_{ita}} \left\{ u(c_{ita} | X_{ita}) + \beta E_{\pi} V_{i,t+1,a+1}^{\theta s}(k_{i,t+1,a+1}, h_{i,t+1,a+1}, \pi_{i,t+1,a+1}) \right\}$$

where $V_{ita}^{\theta s}$ represents the value function of agent i of type θ with education s when aged a at time t and $E_{\pi} V_{ita}^{\theta}$ represents the expected value function while the information about contemporaneous shocks has not been disclosed.

4.2 Properties of the model

The main focus of the present study is placed on the working decision and the rule of human capital accumulation. Some characteristics of the presented model will prove very

useful for the identification strategy and we therefore discuss them in this section. We start with the existence of a solution to the presented problem. Magnac and Robin (1991) use a setting with similar characteristics to study the choice between self-employment and wage-work. They establish a number of results, the first being the existence of a solution for as long as the shocks have bounded support. Their result can be directly applied to the model being presently discussed.

However, to develop a simple estimation strategy and to implement the solution computationally, which will be attempted in the next chapter, stronger properties are required. The next feature to be discussed relates to the specification of the working decision: the working status is decided on a reservation rule principle. In other words, the agent decides to work whenever the particular realisation of π exceeds the optimally determined reservation value denoted by $\pi_{ita}^R \left(X_{ita} \mid (R, W_s, B, \tau)_t^{t+A-a} \right)$.

In all that follows, a compact space is considered for the state variables, X . Denote by $X^{-\pi}$ the set of state variables excluding the productivity shock. For simplicity of notation, the individual subscript i will be omitted from now onwards.

The working decision is characterised in Lemma 1,

Lemma 1 *Given the particular conditions faced by an individual of a certain age at a given time period, the working decision can be described by a productivity shock reservation policy when π is uncorrelated over time: the agent prefers working to staying at home whenever $\pi > \pi_{ta}^R \left(X_{ta}^{-\pi} \mid (R, W_s, B, \tau)_t^{t+A-a} \right)$.*

Proof. Suppose that, for a given $X^{-\pi}$ at some period t when the agent is aged a , working is preferred to staying at home when $\pi = \pi'$. Thus,

$$V_{ta}^{\theta s} (k_{ta}, h_{ta}, \pi'_{ta} \mid d_{ta} = 1) > V_{ta}^{\theta s} (k_{ta}, h_{ta}, \pi'_{ta} \mid d_{ta} = 0)$$

Taking any $\pi'' > \pi'$ we obtain,

$$\begin{aligned} V_{ta}^{\theta s} (k_{ta}, h_{ta}, \pi''_{ta} \mid d_{ta} = 1) &> V_{ta}^{\theta s} (k_{ta}, h_{ta}, \pi'_{ta} \mid d_{ta} = 1) > \\ V_{ta}^{\theta s} (k_{ta}, h_{ta}, \pi'_{ta} \mid d_{ta} = 0) &= V_{ta}^{\theta s} (k_{ta}, h_{ta}, \pi''_{ta} \mid d_{ta} = 0) \end{aligned} \quad (4.6)$$

meaning that the agent prefers to work at any π larger than π' . We can always find a sufficiently low π , say π^* , such that the agent will prefer to stay at home since the

contemporaneous and future working income is relatively low - for instance, this always happens at $\pi^* = 0$ as outcomes from working both in terms of earnings and human capital are nil.¹³ But then, the agent will still prefer to stay at home for any other $\pi^{**} < \pi^*$ or otherwise working would make the agent better off at π^* given condition (4.6). Thus, there will be a threshold, π^R , dependent on $X^{-\pi}$ that completely characterises the decision between working and staying at home. In extreme cases, this threshold may be set very low, meaning that the agent is willing to work in practically any circumstances. ■

Lemma 1 provides a workable representation of the value function that simplifies the following analysis. The next result relates with continuity, differentiability and concavity of the value function. It can be stated that,

Lemma 2 *Assume that utility is a function of consumption alone. On a bounded and convex (for the continuous variables) state space, if the density function of the idiosyncratic productivity shock, π , is continuously differentiable (C^1), and the coefficient of absolute risk aversion is decreasing but not “too much” so, then the expected value function $E_{\pi}V^{\theta_s}(k, h, \pi)$ is a strictly increasing, C^2 and concave function of k .*

Proof. See appendix to this chapter (section 4.7.1). ■

The first requirement in Lemma 2 dictates that wealth effects are out ruled from this analysis. Contrary to risk aversion, wealth effects make working less likely as the beginning of the period level of assets increases. This matter has been discussed before, but it is now worth noting that including a taste for leisure creates additional complications in the analysis by making it possible for the conditional (on the working status) value functions to cross more than once with respect to k .¹⁴ Given our main group of interest and the homogeneity in preferences, such assumption is unlikely to be a serious drawback.

Since the expected value function, EV , is a weighted mean of conditional value functions for each possible labour market option, the second requirement in Lemma 2 just means that continuity and differentiability of the weights are needed.

¹³Such assertion is true for any $B > 0$, which is being assumed throughout this chapter as B stands for the unemployment benefit. More generally, if $B \leq 0$, it will hold for $\pi^* = B / (W_s h (1 - \tau))$.

¹⁴In this case, concavity would require a relatively low density of π at the reservation value, π^R . But this is an undesirable request as π^R is endogenously chosen by the agent.

The third requirement, however, is a bit more cumbersome: it relates to how fast risk aversion declines as the agent becomes wealthier due to changes in the beginning of the period endowment. An agent with decreasing absolute risk aversion is more willing to take risks the wealthier he/she is. In terms of our model, this means that the working reservation policy is a function of k : $\partial\pi_a^R/\partial k_a \leq 0$. Risk aversion implies that, relative to the value of the non-risky option, the value of the risky option increases with k . That is, the relative value of a given potential gain is larger the wealthier the agent is. If this effect is heavily weighted, which happens when $\partial\pi_a^R/\partial k_a$ is large in absolute terms, it may locally overtake the concavity of the utility function, making the value function non-concave. Under constant absolute risk aversion, the agent willingness to take up a risk is independent of the level of assets. Thus, the preferred option does not change with k , being always either the risky or the non-risky. In such case, concavity follows straightforwardly. The more decreasing is risk aversion with the level of beginning of the period wealth, the more pronounced becomes the kink in the value function caused by the crossing of the two conditional (on the working status) value functions, making it more difficult to smooth out. A more detailed analytical explanation is provided in the appendix to this chapter.

Under Lemma 2, the Euler equation is a necessary and sufficient condition for the optimal consumption decision given the working status. It can be written,

$$\left. \frac{\partial u}{\partial c_{ta}} \right|_{d_{ta}} = E \left[\beta (1 + R_t) \left. \frac{\partial u}{\partial c_{t+1,a+1}} \right|_{d_{ta}} \right]$$

The next result characterises the consumption decision rule. It makes it clear that income alone is not enough to identify the current optimal consumption, the level of schooling and the amount of human capital being also required.

Lemma 3 *Assume that utility is a function of consumption alone. If $EV(k, \cdot)$ and $u(c | \cdot)$ are strictly increasing, concave and C^2 in k and c respectively, then consumption and savings are normal goods for fixed working decisions.*

Proof. See appendix to this chapter (section 4.7.2). ■

Notice that in a more general specification of this type of model that includes leisure in the utility function, consumption will not usually be normal even when conditional on the discrete labour market decision. An exception occurs if leisure and consumption are

perfectly substitutable.¹⁵ This is because the dynamic setting of the model dictates that, at a certain level of consumption, a current increase in income may be used to purchase additional leisure in the future instead of contemporaneous consumption.

4.3 Estimation process

This section presents the fundamental aspects related with the identification of the parameters of the model. As discussed below, the structural set-up will be heavily explored to provide conditions for identification. However, some additional assumptions are needed at places, and this will be made clear from what follows. The data requirements will be discussed together with the methodological issues given the strong interlink between the appropriate methodology and the data at hand. It will be argued that information on future labour market behaviour can be quite informative about the present performance. A discussion about the specific data set being used in estimation, however, is postponed to the next section.

¹⁵Notice that the adopted specification is equivalent to assume that consumption and leisure are perfect substitutes. To make it clear, suppose that working is associated with some disutility in the following way

$$u_{ta}(c, d) = f(c + \alpha_d(1 - d))$$

If a transformation of the consumption variable is now considered,

$$\bar{c} = c + \alpha_d(1 - d)$$

the consumers problem may be re-written as

$$\begin{aligned} V_{ta}^{\theta s}(k_{ta}, h_{ta}, \pi_{ta}) &= \max_{\bar{c}, d} \left\{ u_{ta}(\bar{c}_{ta}) + \beta EV_{t+1, a+1}^{\theta s}(k_{t+1, a+1}, h_{t+1, a+1}) \right\} \\ & \text{s.t.} \\ k_{t+1, a+1} &= (1 + R_t) k_{ta} + d_{ta} h_{ta} \pi_{ta} W_{st} + (1 - d_{ta})(B_t + \alpha_d) - \bar{c}_{ta} \\ h_{t+1, a+1} &= h_{ta} (1 + \nu(\theta, s, h_{ta}) \pi_{ta}) & \text{if } d_{ta} = 1 \\ h_{t+1, a+1} &= h_{ta} & \text{if } d_{ta} = 0 \end{aligned}$$

which is precisely the same problem if the additional transformation is performed,

$$\bar{B}_t = B_t + \alpha_d$$

That is, as long as consumption and leisure are perfectly substitutable, there is a monetary equivalent to the utility cost of working that is independent of the shape of the utility function, units or level of utility.

4.3.1 Identification using the structural model

The selection problem

Lets start by establishing the (net) earnings equation of an agent of type θ with educational level s at any time t when aged a . It follows from equation (4.2) that the logarithm of net earnings can be written as follows,

$$\ln E_{ta}^{\theta s} = \ln(W_{st}(1 - \tau_t)) + \ln h_{ta}^{\theta s}(\pi(a - 1)) + \ln \pi_{ta} \quad (4.7)$$

where $E_{ta}^{\theta s}$ is the net earnings of an (θ, s) agent when aged a at time t and $h_{ta}^{\theta s}$ is the corresponding level of human capital. It should be reminded that h depends on the past labour market history of the agents, as detailed in equation (4.3). Conditional on (t, θ, s, a) , past labour market history is completely characterised by the sequence of idiosyncratic productivity shocks experienced up to age $a - 1$, here denoted by $\pi(a - 1)$.

On the other hand, it has been previously established (see lemma 1) that the working status is decided on a reservation rule principle: the agent chooses to work whenever the particular realisation of the productivity shock π exceeds the optimally determined reservation rule. As before, denote by $X_{ta}^{-\pi}$ the state space faced by an agent aged a at time t excluding the productivity shock, π . The reservation policy takes the following aspect,

$$d_{ta}^{\theta s} = 1 \quad \text{iff} \quad \pi_{ta} \geq \pi_{ta}^R(X_{ta}^{-\pi} | (R, W_s, B, \tau)_t^{t+A-a}) \quad (4.8)$$

All the variables explaining π^R are indeed important: t and a determine the aggregate prices faced by the agent during the rest of his/her life, θ and s characterise the speed of human capital accumulation, h and s determine the income from working and k informs on how risk aversion may affect working decisions when not enough human capital has been accumulated to make working a riskless choice.

For as long as one is willing to impose some additional structure, the earnings selection model (4.7) and (4.8) discloses information about the individual's reservation policies and the variability of the idiosyncratic productivity shock. We proceed parametrically by postulating a functional form for the reservation policy, π^R , and a given distribution for the idiosyncratic productivity/efficiency shock, π . The specific parametric assumptions chosen are discussed below when presenting the practical estimation issues.

Under the assumption on homogeneity of initial levels of human capital (equation (4.5)), this model is particularly informative about the distribution of the productivity shock when applied to young agents at the start of their working lives. Thus, we import this assumption from the theoretical setting, which is needed to overcome the task of establishing the individual's amount of human capital. Homogeneity of initial levels of human capital is critical for the estimation of equation (4.7) since π , the random component of the model, and h , the individual-specific amount of human capital, are expected to be negatively correlated among workers. The assumption is imposed within education-ability groups as follows,

- A1 The initial level of s -specific human capital is homogeneous among agents of type θ , $h_{1i}^{\theta s} = \underline{h}^{\theta s}$, where the subscript i indexes individuals and 1 locates human capital at the initial period of life.

The discussed setting makes a sample of youngsters at a given time the best source of information to identify the variance of the productivity shock and the aggregate wage rates by educational level. On the other hand, individual-specific participation rules can be estimated at different stages of the life-cycle and be used to identify the growth rate of wages.¹⁶ This is the subject of the next discussion.

Growth rate of Earnings

Later in the lifetime, experience alone is not enough to explain the individual's level of human capital: the ability/skill/human capital specific rate of human capital accumulation, $\nu(\theta, s, h)$, and the idiosyncratic history of productivity shocks, $\pi(a)$, determine human capital and earnings as well.¹⁷ Identification of the earnings' growth rate is possible under the presence of rich longitudinal data by making use of past and future information.

¹⁶Though the amount of human capital is a structural component of π^R , one could think of including other measures that partially control for differences in skills on the top of experience. For instance, past health conditions, births or accumulated wealth may contain information about previous shocks that affect the present level of human capital. This is less of a problem in the estimation of the decision rule than in the wage equation itself.

¹⁷For instance, an agent that receives a good shock, meaning that he/she is well fitted for the job, is likely to dedicate more effort, earn more and improve his/her skills at a faster pace.

To make things more precise, let's take a given cohort, to whom we can observe a number of characteristics including ability and educational levels, working status, wealth and earnings growth from time t to $t + 1$. For agents of ability type θ and educational level s that choose to work in period t , the human capital growth rate can be formalised as,

$$\ln \left[h_{t+1,a+1}^{\theta s} (\pi(a)) \right] - \ln \left[h_{ta}^{\theta s} (\pi(a-1)) \right] = \nu(\theta, s, h) \pi_{at} \quad (4.9)$$

Since human capital is a continuous variable, we do now assume a particular specification for the shape of $\nu(\theta, s, h)$.

A2 The rate of human capital accumulation changes at a constant rate with the level of accumulated human capital.

It can therefore be written that $\nu(\theta, s, h) = \nu(\theta, s) r(\theta, s)^{h_{ta}^{\theta s} - h^{\theta s}}$, where $\nu(\theta, s)$ is the initial rate of human capital accumulation and $r(\theta, s)$ stands for the rate of adjustment of the human capital growth rate or the rate at which $\nu(\theta, h, s)$ changes with human capital. Both parameters are ability- and skill-specific. This is a simple and parsimonious specification of the human capital growth rate, and that justifies the choice. Its only constraint is the requirement for $\nu(\theta, s, h)$ to decrease at a constant rate with h , namely r (it is being assumed that $r < 1$, which will be confirmed empirically). Other specifications allowing for faster or slower movements in $\nu(\theta, s, h)$ can also be tried out. However, at this stage we lack the required data to distinguish between specifications differing on the speed of change of the human capital growth rate.¹⁸

Under (A2), equation (4.9) can be replaced by

$$\ln \left[h_{t+1,a+1}^{\theta s} (\pi(a)) \right] - \ln \left[h_{ta}^{\theta s} (\pi(a-1)) \right] = \nu(\theta, s) r(\theta, s)^{h_{ta}^{\theta s} - h^{\theta s}} \pi_{at} \quad (4.10)$$

By replacing (4.10) in (4.7) one obtains,

$$\begin{aligned} \ln E_{t+1,a+1}^{\theta s} - \ln E_{ta}^{\theta s} = & \quad (4.11) \\ & [\ln W_{s,t+1} - \ln W_{st}] + \nu(\theta, s) r(\theta, s)^{h_{ta}^{\theta s} - h^{\theta s}} \pi_{ta} + [\ln \pi_{t+1,a+1} - \ln \pi_{ta}] \end{aligned}$$

¹⁸As discussed below when presenting the data issues, only 3 periods of the agents' working life are available.

Equation (4.11) decomposes the earnings growth rate in three parts: the aggregate share $(\ln W_{s,t+1} - \ln W_{st})$, the idiosyncratic accumulation of human capital $(\nu(\theta, s) r(\theta, s)^{h_{ta}^{\theta s} - h^{\theta s}} \pi_{ta})$ and the idiosyncratic between periods differential in productivity $(\ln \pi_{t+1,a+1} - \ln \pi_{ta})$. Estimation of $\nu(\theta, s)$ and $r(\theta, s)$ require the selection rule (4.8) to be applied to the two periods, t and $t+1$, in order to predict the reservation policies. Working agents are those experiencing productivity shocks above the estimated threshold. In addition, we know from the theoretical setting that the state variables follow a Markov process, meaning that $X_{t+1,a+1}^{-\pi} = g(X_{ta}^{-\pi}, \pi_{ta} | (R, W_s, B, \tau)_t^{t+A-a})$. But then the reservation rule at time $t+1$ is a function of the previous period state of the nature. Conditional on the price path, $\pi_{t+1,a+1}^R(X_{t+1,a+1}^{-\pi}) = \pi_{t+1,a+1}^R(g(X_{ta}^{-\pi}, \pi_{ta})) = \pi_{t+1}^R(X_{ta}^{-\pi}, \pi_{ta})$. Hence (omitting the age subscript for simplicity of notation),

$$\begin{aligned}
E(\ln \pi_t | d_t = 1, d_{t+1} = 1, X_t^{-\pi}) &= \tag{4.12} \\
&\int_{\ln \pi_t^R(X_t^{-\pi})}^{+\infty} \int_{\ln \pi_{t+1}^R(X_t^{-\pi}, \pi_t)}^{+\infty} \ln \pi_t \frac{f(\ln \pi_t, \ln \pi_{t+1})}{\rho(X_t^{-\pi})} d \ln \pi_{t+1} d \ln \pi_t = \\
&\frac{1}{\rho(X_t^{-\pi})} \int_{\ln \pi_t^R(X_t^{-\pi})}^{+\infty} \ln \pi_t f^m(\ln \pi_t) \int_{\pi_{t+1}^R(X_t^{-\pi}, \pi_t)}^{+\infty} f^m(\ln \pi_{t+1}) d \ln \pi_{t+1} d \ln \pi_t = \\
&\frac{1}{\rho(X_t^{-\pi})} \int_{\ln \pi_t^R(X_t^{-\pi})}^{+\infty} \ln \pi_t f^m(\ln \pi_t) [1 - F^m(\ln \pi_{t+1}^R(X_t^{-\pi}, \pi_t))] d \ln \pi_t
\end{aligned}$$

where f and F stand for the density and cumulative joint distribution of the productivity shocks and f^m and F^m stand for the respective marginal distribution, which is the same for both variables given that the shocks are iid across agents and over time. The term $\rho(X_t^{-\pi})$ can be expressed as,

$$\begin{aligned}
\rho(X_t^{-\pi}) &= \int_{\ln \pi_t^R(X_t^{-\pi})}^{+\infty} \int_{\ln \pi_{t+1}^R(X_t^{-\pi}, \pi_t)}^{+\infty} f(\ln \pi_t, \ln \pi_{t+1}) d \ln \pi_{t+1} d \ln \pi_t = \\
&\int_{\ln \pi_t^R(X_t^{-\pi})}^{+\infty} f^m(\ln \pi_t) [1 - F^m(\ln \pi_{t+1}^R(X_t^{-\pi}, \pi_t))] d \ln \pi_t = \\
&F^m(\ln \pi_t^R(X_t^{-\pi})) E[F^m(\ln \pi_{t+1}^R(X_t^{-\pi}, \pi_t)) | \pi_t \geq \pi_t^R]
\end{aligned}$$

Equation (4.12) makes it clear that the reservation policy $\pi_{t+1,a+1}^R$ is informative about

the size of past shocks, π_{ta} , conditional on the state of the nature $X_{ta}^{-\pi}$. To use such result in full, we make the following claim,

Lemma 4 *Conditional on $X_{ta}^{-\pi}$ and the prices $(R, W_s, B, \tau)_t^{t+A-a}$, the reservation policy $\pi_{t+1, a+1}^R$ is a monotonic (strictly decreasing) function of the shock π_{ta} among workers aged a at time t .*

Proof. The proof is straightforward. Take a given $X_{ta}^{-\pi}$ at a moment in time. Among workers, the shock π_{ta} has two distinct effects: it changes the contemporaneous returns from working and affects the speed of human capital accumulation. Agents enjoying from a larger shock π_{ta} save more (see Lemma 4.3) and accumulate more human capital. But then they are wealthier next period, making risk aversion less of a problem in their working decisions (see the proof of Lemma 4.2), and working becomes a better option given their accumulated skills. That is, working becomes more likely, which is to say that $\pi_{t+1, a+1}^R$ decreases with π_{ta} among time t workers aged a conditional on X_{ta} . ■

The result described in Lemma 4 is a consequence of the exclusion of wealth effects from the working decision process as leisure is not considered in the utility function (or, which is the same, is considered to be perfectly substitutable for consumption). The importance of such hypothesis in the context of evaluating labour market interventions has been discussed before. Its benefits for the analysis can now be seen: it provides a strong tool by allowing the distribution of $\pi_{t+1, a+1}^R | X_{ta}^{-\pi}, d_{ta} = 1$ to be used to characterise the distribution of $\pi_{ta} | X_{ta}^{-\pi}, \pi_{ta} \geq \pi_{ta}^R$. We use the information on the predicted reservation rule, $\widehat{\pi_{t+1, a+1}^R}$, to define the predicted shock at time t , $\widehat{\pi}_{ta}$, to be given by

$$\widehat{\pi}_{ta} : P\left(\pi_{t+1, a+1}^R \leq \widehat{\pi_{t+1, a+1}^R} | X_{ta}, d_{ta} = 1\right) = P\left(\pi_{ta} \leq \widehat{\pi}_{ta} | X_{ta}, \pi_{ta} \geq \pi_{ta}^R\right) \quad (4.13)$$

Notice that in all this and following discussions, “predicted” applies to the econometrician predictions, not to the agents’ forecasts.

As in what concerns to period $t + 1$, the information on whether the agent has worked before is not relevant given the particular realisation of $X_{t+1, a+1}^{-\pi}$. This is a direct consequence of the theoretical setting of the model, where it is established that the present state of the nature is enough to completely determine the optimal choices without making use of any additional past information.

Establishing the bounds for the productivity shock, π

The characterisation of the human capital production function still requires the bounds for the productivity shock to be determined. Such parameters are needed to implement the numerical solution of the problem as attempted in the next chapter and ensure that the state space is bounded as required to establish the properties of the value function (Lemma 2).

To determine the distribution of the random component of the model, we first restate the i.i.d. assumption and then assume that the bounds are symmetric in logarithms.

A3 The support the productivity shock, π , is $[\underline{\pi}, \bar{\pi}]$ where $\ln \underline{\pi} = -\ln \bar{\pi}$.

We now notice that the maximum possible earnings growth rate for any (θ, s) -group is attained when $\pi_{t+1,a+1}$ is $\bar{\pi}$ and π_{ta} is either $\bar{\pi}$ or $\underline{\pi}$, where $\bar{\pi}$ and $\underline{\pi}$ stand, respectively, for the upper and lower bounds of the distribution of π . This results from equation (4.11). The two possible values the productivity shock π_{ta} might assume are a consequence of its dual effect on the earnings growth rate. On the one hand, it affects the growth rate of earnings positively by increasing the rate of human capital accumulation. On the other hand, it has a negative impact on the growth rate of earnings by increasing the level of earnings in period t .

Therefore, $\bar{\pi}$ is either characterised by,

$$\bar{\pi} = \frac{\overline{(\ln E_{t+1,a+1}^{\theta s} - \ln E_{ta}^{\theta s})} - (\ln W_{s,t+1} - \ln W_{st})}{\nu(\theta, s, h)} \quad (4.14)$$

or

$$\nu(\theta, s, h) \exp(-\ln \bar{\pi}) + 2 \ln \bar{\pi} = \overline{(\ln E_{t+1,a+a}^{\theta s} - \ln E_{ta}^{\theta s})} - (\ln W_{s,t+1} - \ln W_{st}) \quad (4.15)$$

depending on whether the maximum is achieved at $\pi_{ta} = \bar{\pi}$ or $\pi_{ta} = \underline{\pi}$, respectively

By using the maximum observed earnings growth rate conditional on the level of human capital, $\overline{(\ln E_{t+1,a+1}^{\theta s} - \ln E_{ta}^{\theta s})}$, together with the predicted growth rate of wages, $(\ln W_{s,t+1} - \ln W_{st})$, one can approximate both measures of $\bar{\pi}$ and obtain an estimate of the upper bound by choosing the minimum.¹⁹ As the same interval is assumed to affect all

¹⁹The choice of the minimum is explained by the following argument. Lets denote by $\bar{\pi}^m$ and $\bar{\pi}^M$ the

agents, independently of their ability, schooling or level of human capital, the maximum group estimate should be used as an estimate of $\bar{\pi}$. Quite obviously, the lower bound is then set to be $\underline{\pi} = \exp(-\ln \bar{\pi})$. As the bounds are used to approximate the individuals' period t productivity shocks, an iterative procedure is required at this stage.

4.3.2 Practical estimation procedure

The estimation procedure can be summarised as follows. We start by making a parametric assumption about the distribution of the productivity shock, π .

A4 Let $\ln \pi$ follow a truncated normal distribution $N(0, \sigma_\pi)$ in $[\underline{\pi}, \bar{\pi}]$.

Given (A4), and the symmetry of the support of $\ln \pi$, (A3), the distribution of the productivity shocks becomes perfectly known when σ_π and $\bar{\pi}$ are identified. The aim of *Step 1* of the estimation method is to identify σ_π . To do so, we assume that the support of $\ln \pi$ is unbounded. Such assumption is to be relaxed in the following estimation steps but is unlikely to strongly affect the estimates at this stage.²⁰

We use the selection model of earnings, (4.7) and (4.8), to estimate σ_π . This requires the some further parametric assumptions.

A5 The reservation policy is a linear function of the characteristics, Z ,

$$\ln \pi_{ita}^R = Z_{ita}\gamma \quad (4.16)$$

where Z includes innate ability (θ), education (s), human capital (h), wealth (k) and information on fixed costs from working (B),²¹ and γ is the corresponding vector of parameters.

minimum and maximum estimates for the upper bound of the distribution of π , respectively. Suppose instead that $\bar{\pi}^M$ would be chosen and that it happens under hypothesis (4.14). What would happen to agents experiencing a shock in $t + 1$, $\pi_{t+1,a+1}$, between $\bar{\pi}^m$ and $\bar{\pi}^M$ and a shock in t equal to $\underline{\pi}$? Their earnings growth rate would exceed the maximum observed in the population as that is achieved at $\pi_{t+1,a+1} = \bar{\pi}^m$. A similar argument can be developed for hypothesis (4.15).

²⁰In fact, empirical results presented below show that only the very thin tails of the distribution of $\ln \pi$ are truncated.

²¹As will be seen below, some heterogeneity of fixed costs from working is allowed for in the practical empirical analysis.

A6 The working decision rule, d_{ita} , is specified as

$$\begin{aligned} y_{ita} &= -Z_{ita}\gamma + \ln \pi_{ita} \\ d_{ita} &= 1 [y_{ita} > 0] \end{aligned} \quad (4.17)$$

where π is the productivity shock.

Assumptions (A5) and (A6) establish the participation rule. Assumption (A7) below does the same to the earnings equation,

A7 The earnings equation is specified as

$$\ln E_{ita} = W_{ita}\beta + \ln \pi_{ita} + \varepsilon_{ita} \quad (4.18)$$

where W includes ability (θ), education (s) and human capital (h), and β is the corresponding vector of parameters. ε is an error term assumed to be iid and uncorrelated with π .

Several comments are due at this stage. First, the error term ε includes no permanent effects. All the permanent differences in earnings not controlled for by the variables in W are assumed to arise through the productivity shock, π , which has long-lasting effects on human capital. Second, conditional on the other characteristics, information on wealth and fixed costs from working is assumed to affect the participation decision but not earnings. This is the exclusion restriction dictated by the theoretical framework. And Third, Z and W are assumed to be observable. A more detailed discussion of the variables included in the model is postponed to the subsections describing the data and the empirical results. However, it should now be noticed that no direct measure of the amount of human capital is possibly available in the data. The solution advanced here is to use young individuals at the beginning of their working life, for whom homogeneity of human capital is assumed.

Under the above discussed conditions, a two-stage selection estimator can be applied to model (4.17) and (4.18) among youngsters to identify σ_π (see Heckman, 78 and 79).²²

Step 2 uses the selection model (4.17) at different stages of the life cycle to produce predictions of the reservation policies, π^R , as specified in equation (4.16). Controlling for

²²An additional outcome from such analysis is an estimate of the wage rates for different types of human capital. Such information will be used the next chapter to approximate the aggregate production function.

differences in human capital is now more difficult as it depends on the agents' idiosyncratic labour market experience and is not fully measured by the number of working periods. For simplicity, this version of the analysis uses working years to control for human capital. However, a better solution uses an iterative procedure to bring information on the actual unobserved levels of human capital as predicted in the next steps back to this step. The following steps describe the procedure that is now implemented to simultaneously uncover information about the idiosyncratic levels of human capital and estimate the human capital growth rates. Its extension to include an actual measure of human capital in this step seems straightforward but computationally heavier.

Step 3 establishes the first guess for the bounds of the productivity shock support.

Step 4 aims at predicting the actual productivity shocks. Predictions are only required for working agents as the idiosyncratic productivity shock has no impact on unemployed agents. We use two conditions to predict π among working agents. The first uses the computations performed in step 2: the reservation policies, which impose a lower bound on the set of possible values of π among workers. The second derives from the distribution of next period reservation policies conditional on today's working status and characteristics, Z . Lemma 4 establishes a monotonicity relationship between tomorrow's reservation policy and today's productivity shock for working agents conditional on all other characteristics. Thus, the distribution of the following period reservation policy can be used to infer the distribution of the present period productivity shock within the bounds established by the contemporaneous reservation value, π^R , and the upper bound for the support of π . This makes use of condition (4.13). To do so, observations within each $(Z, d = 1)$ -group were first ranked by their next period reservation policy. The predicted value for the productivity shock corresponds to the respective centile in the (truncated by π^R and $\bar{\pi}$) distribution of today's productivity shock. Alternatively, we also used the expected value within a small interval around the respective centile but no significant differences were found.

Knowledge about the individual productivity shocks can then be used to predict the individual-specific levels of human capital throughout the working life, which is different from experience given the heterogeneity in employment experiences across individuals and over time within each individual life. This is done in *Step 5*, together with the estimation of

the rates of human capital accumulation by (θ, s) -group. Given that neither human capital or human capital growth rate are observable, the two factors must be approximated in simultaneous. Identification requires information on at least three consecutive periods at the start of the agents' working life. It uses the earnings' growth rate equation (4.11) applied to ages 1 to 2 and 2 to 3. For an agent aged 1 at time t , the regression equations can be written as

$$\begin{aligned} \ln E_{t+1,2}^{\theta s} - \ln E_{t1}^{\theta s} = & \quad (4.19) \\ & [\ln W_{s,t+1} - \ln W_{st}] + \nu(\theta, s)\hat{\pi}_{t1} + [\ln \hat{\pi}_{t+1,2} - \ln \hat{\pi}_{t1}] + [\varepsilon_{t+1,2} - \varepsilon_{t1}] \end{aligned}$$

for the transition between ages 1 and 2, and

$$\begin{aligned} \ln E_{t+2,3}^{\theta s} - \ln E_{t+12}^{\theta s} = & \quad (4.20) \\ & [\ln W_{s,t+2} - \ln W_{s,t+1}] + \\ & \nu(\theta, s)r(\theta, s)^{h_{t+1,2}^{\theta s} - \hat{h}^{\theta s}} \hat{\pi}_{t+1,2} + [\ln \hat{\pi}_{t+2,3} - \ln \hat{\pi}_{t+1,2}] + [\varepsilon_{t+1,3} - \varepsilon_{t+1,2}] \end{aligned}$$

for the transition between ages 2 and 3. Equations (4.19) and (4.20) show that identification of the initial rate of human capital accumulation makes use of the first working period and the adjustment rate uses the following periods. Estimation is performed using non-linear least squares.

Step 6 computes the new bounds for the productivity shock. The two possibilities presented in equations (4.14) and (4.15) are computed using the 95th centile of the earnings growth rate to estimate the maximum growth rate. Other parameters required for these computations were obtained in Step 5 and the Newton's method was used to compute the zero of (4.15). The new upper bound corresponds to the minimum of the two estimates.

The last stage of the estimation procedure, *Step 7*, compares the new estimate of $\bar{\pi}$ with the old one. If convergence has been achieved as defined by a maximum acceptable distance defined *a priori*, the iterative process stops and the new set of estimates is accepted. Otherwise, we return to Step 4 and repeat the process of estimating the human capital growth rates.

4.4 The data

The estimation of the human capital production function used the National Child Development Study (NCDS58), a UK longitudinal data set containing information on individuals

born in one week of 1958. Several waves of interviews were performed at different stages of individuals' life-cycle, namely when they were aged 0, 7, 11, 16, 23, 33 and 42 years old. Detailed data was collected at each developmental stage, including information on health, cognitive development, educational attainment, household composition, parents socio-economic status and education, housing and other forms of wealth, employment status, employment history and income.

A major advantage of the NCDS58 over other datasets relates to its reporting of ability measures. At ages 7 to 16, the sampled individuals were tested on their reading and maths ability. Interviews at 7 and 11 also include the teachers' qualitative assessment of each child development along different dimensions covering oral, awareness, reading, creative and numerical abilities for the 7 years olds and oral, reading, maths and general knowledge for the 11 years olds. We have used this information to construct a measure of the ability level for each individual, averaging the reported test scores when available and otherwise resorting to the teacher's assessments. Three ability groups are then formed according to the assessed ability.²³

The education information is also very detailed in the NCDS. All the waves include questions about the highest educational level the individual ever obtained and possible educational or training courses completed between interviews. Again, we considered three levels of education, corresponding to, respectively, the basic level, achieved by 16 years of age and equivalent to less than 5 O-levels ($s = 1$), completion of secondary school or A-levels ($s = 2$) and graduation from college ($s = 3$). The NDYP design was driving such choice. In fact, programme participants have been characterised as generally belonging to the least educated group of the population (Bell, Blundell and Van Reenen, 99) and the treatment being offered is not sufficiently long and intense to directly transform unskilled in the most skilled labour. Indirectly this may happen, as once a first step is overcome, participants become more prone to continue investing in education. Thus, subsidies may affect unskilled agents at the margin of investing in education but do not directly make them high skilled workers.

Data used in estimation correspond mainly to the interviews at 23, 33 and 42 years of

²³The several measures of ability are highly correlated, making the composition of the groups only mildly affected by the choice of the measure.

age. As described in step four of the estimation procedure presented in section 4.3.2, the rates of human capital accumulation are computed using observations on working individuals that have been continuously employed between observations. Included observations consist of agents working up to the age of 33 and agents working up to the age of 42. The former are used to identify the initial rate of human capital accumulation and the latter are used to identify the adjustment rate. The working requirement is explained by the need to observe earnings and to guarantee that the productivity shocks impact on all agents for a similar amount of time. Finally, the observational period corresponds to the women's fertility age, where the dynamics of family matters as marriage and childbearing may strongly affect labour market decisions. Such analysis is clearly outside the scope of this study. However, since the inclusion of women did not change the nature of the results, we have opted for considering as many observations as possible, keeping them in the analysis. The appendix to this chapter (section 4.7.3) contains a description of the observations used in the analysis.

4.5 Estimation results

4.5.1 Earnings model at the beginning of the working life.

Estimation is based on data on real net hourly wages at 23, 33 and 42 years old, along with information on ability, education, experience, household composition, health and wealth. We start by computing the variance of the productivity shock at 23 years of age, which is the closest observation to the start of the agents working life. This is done using the model described by assumptions (A6) and (A7). As described in (A6), the selection process is assumed to depend on ability, education, human capital, wealth and fixed costs from working. Ability is directly measured by the information on cognitive development described in the data section above. We also included family inheritances (conditional on other wealth measures) to account for other potentially important background factors not measured by the cognitive ability variable constructed. The educational variable included has also been described before, within the data section. Human capital is not directly observable in the data, and this is the reason to use the beginning of the working life information at this stage given the assumption on homogeneity of initial levels of human

capital. However, since our first working age observation is for 23 years old, we also include information on past working time. Wealth is measured by variables on the possession of own accommodation, financial savings, having a partner and working partner. It is being assumed that, at the start of the working life, wealth is expected to influence working decisions more than to be already a consequence of past successful working experiences, which could signal permanent differences between agents not measured by the ability variables. Finally, differential fixed costs from working are controlled for by gender, partner and the existence of children under 4. The earnings selection model places the exclusion restriction on wealth and fixed costs from working, so that related variables are excluded from the earnings regression.

The overall results from this analysis are presented in tables 4.1 and 4.2.

A number of specifications have been tried out, the results being very similar. The four possible specifications presented reveal the main results. The first two consider both men and women while the third and fourth ones are applied to the reduced sample of men only. The second and fourth distinguish between initial levels of human capital by ability*education while the first and third assume that the initial levels of human capital are homogeneous within ability groups independently of the educational level. The essence of the results is similarly reproduced in each model. The first specification is our preferred one for a reason made clear in what follows. Notice that the columns numbering in tables 4.1 through 4.3 is consistent, referring always to the same models.

Table 4.1 contains the figures for the selection process. It is common to the four specifications that higher levels of education and ability promote the working status. When restricting the sample to men, only the group with the highest level of ability seems to behave differently (column (3)), especially when in low to medium skilled jobs (column (4)). Experience was controlled for mainly because the three levels of education require quite distinct investments in terms of time out of work. Having worked for a reasonable period (more than 1 year) seems to be quite important, especially if occurring in the recent past (*worked past 2 yrs*). But once this is controlled for, only experience in the labour market of more than 5 years impacts on the probabilities of being observed as an employee among men. It mainly distinguishes unskilled agents that have working experience, but not other type of workers as A-levels (medium skills) are only completed at 18 years of age, precisely

Table 4.1: Selection model of earnings.
Selection regression for agents aged 23.

| | Men and Women | | | | Men only | | | |
|-----------------------|---------------|-------|----------|-------|----------|-------|----------|-------|
| | (1) | | (2) | | (3) | | (4) | |
| | Coef. | SE | Coef. | SE | Coef. | SE | Coef. | SE |
| female | -0.163** | 0.040 | -0.157** | 0.039 | | | | |
| medium skills | 0.887** | 0.053 | 0.925** | 0.091 | 0.924** | 0.075 | 0.916 | 0.112 |
| high skills | 1.581** | 0.093 | 1.592** | 0.162 | 1.776** | 0.136 | 1.846** | 0.237 |
| ability level 2 | 0.137** | 0.045 | | | 0.064 | 0.064 | | |
| ability level 3 | 0.272** | 0.052 | | | 0.302** | 0.074 | | |
| ab. 2, low skills | | | 0.113** | 0.053 | | | 0.022 | 0.079 |
| ab. 3, low skills | | | 0.310** | 0.081 | | | 0.384** | 0.122 |
| ab. 2, med. skills | | | 0.092 | 0.098 | | | 0.049 | 0.121 |
| ab. 3, med. skills | | | 0.211** | 0.095 | | | 0.318** | 0.119 |
| ab. 2, high skills | | | 0.160 | 0.162 | | | 0.139 | 0.240 |
| ab. 3, high skills | | | 0.240* | 0.148 | | | 0.175 | 0.216 |
| exper. \geq 12 mths | 0.611** | 0.092 | 0.608** | 0.091 | 0.739** | 0.139 | 0.716** | 0.137 |
| exper. \geq 24 mths | 0.480** | 0.082 | 0.468** | 0.083 | 0.174 | 0.127 | 0.195 | 0.126 |
| exper. \geq 36 mths | -0.083 | 0.104 | -0.055 | 0.102 | 0.104 | 0.154 | 0.156 | 0.151 |
| exper. \geq 48 mths | 0.229** | 0.083 | 0.214** | 0.082 | 0.056 | 0.124 | 0.044 | 0.122 |
| exper. \geq 60 mths | 0.560** | 0.058 | 0.551** | 0.058 | 0.636** | 0.083 | 0.630** | 0.083 |
| exper. \geq 72 mths | 0.440** | 0.055 | 0.438** | 0.055 | 0.573** | 0.079 | 0.572** | 0.079 |
| exper. \geq 84 mths | 1.655** | 0.107 | 1.656** | 0.107 | 1.807** | 0.204 | 1.817** | 0.203 |
| worked past 2 yrs | 2.155** | 0.116 | 2.166** | 0.115 | 2.185** | 0.163 | 2.203** | 0.162 |
| accommodation | 0.010 | 0.052 | 0.013 | 0.052 | 0.352** | 0.090 | 0.335** | 0.088 |
| financial savings | 0.454** | 0.044 | 0.449** | 0.044 | 0.585** | 0.056 | 0.574** | 0.055 |
| partner in FTJ | -0.614** | 0.067 | -0.612** | 0.066 | 0.342** | 0.112 | 0.365** | 0.111 |
| past inheritances | 0.047 | 0.055 | 0.044 | 0.054 | 0.012 | 0.086 | 0.011 | 0.085 |
| Children under 4 | -1.218** | 0.054 | -1.202** | 0.053 | -0.004 | 0.106 | 0.014 | 0.104 |
| partner | 0.460** | 0.067 | 0.450** | 0.067 | -0.002 | 0.098 | -0.008 | 0.097 |
| constant | -3.714** | 0.136 | -3.704** | 0.135 | -4.121** | 0.192 | -4.095** | 0.190 |

Notes: Estimation of the working decision equation used data from the NCDS58 for agents aged 23 years of age.

These estimates are part of a maximum likelihood estimation of the earnings selection model. Estimates of the earnings equation are presented in table 4.2. Estimation used 11,347 observations when men and women are taken together and 5,473 observations when only men are being considered. From these, 7,762 and 4,397 include information on earnings as well (not censored).

** = significant at 0.05 level. * = significant at 0.10 level.

Table 4.2: Selection model of earnings.

Earnings regression for agents aged 23.

| | Men and Women | | | | Men only | | | |
|-----------------------|---------------|-------|----------|--------|----------|-------|----------|-------|
| | (1) | | (2) | | (3) | | (4) | |
| | Coef. | SE | Coef. | SE | Coef. | SE | Coef. | SE |
| female | -0.159** | 0.006 | -0.157** | 0.006 | | | | |
| medium skills | 0.152** | 0.014 | 0.193** | 0.019 | 0.064** | 0.020 | 0.096** | 0.025 |
| high skills | 0.338** | 0.027 | 0.336** | 0.035 | 0.164** | 0.038 | 0.119** | 0.050 |
| ability level 2 | 0.056** | 0.008 | | | 0.043** | 0.010 | | |
| ability level 3 | 0.098** | 0.009 | | | 0.072** | 0.012 | | |
| ab. 2, low skills | | | 0.057** | 0.009 | | | 0.043** | 0.013 |
| ab. 3, low skills | | | 0.123** | 0.013 | | | 0.093** | 0.018 |
| ab. 2, med. skills | | | 0.027* | 0.016 | | | 0.021 | 0.020 |
| ab. 3, med. skills | | | 0.044** | 0.016 | | | 0.042** | 0.020 |
| ab. 2, high skills | | | 0.044* | 0.027 | | | 0.072* | 0.039 |
| ab. 3, high skills | | | 0.100** | 0.025 | | | 0.099** | 0.035 |
| exper. \geq 12 mths | 0.118** | 0.021 | 0.116** | 0.021 | 0.074** | 0.031 | 0.070** | 0.030 |
| exper. \geq 24 mths | 0.048** | 0.014 | 0.051** | 0.0145 | 0.089** | 0.022 | 0.092** | 0.022 |
| exper. \geq 36 mths | 0.028 | 0.027 | 0.025 | 0.027 | -0.024 | 0.041 | -0.028 | 0.040 |
| exper. \geq 48 mths | 0.052* | 0.031 | 0.046* | 0.026 | 0.057* | 0.031 | 0.058* | 0.030 |
| exper. \geq 60 mths | 0.055** | 0.013 | 0.052** | 0.013 | 0.031 | 0.019 | 0.031 | 0.019 |
| exper. \geq 72 mths | 0.025 | 0.014 | 0.021 | 0.014 | -0.014 | 0.020 | -0.014 | 0.020 |
| exper. \geq 84 mths | 0.070** | 0.011 | 0.069** | 0.011 | 0.014 | 0.014 | 0.013 | 0.014 |
| worked past 2 yrs | 0.008 | 0.066 | 0.012 | 0.066 | -0.105 | 0.083 | -0.104 | 0.082 |
| constant | 0.945** | 0.080 | 0.945** | 0.080 | 1.453** | 0.099 | 1.469** | 0.098 |
| lambda | 0.049** | 0.016 | 0.050** | 0.016 | -0.136** | 0.018 | -0.140** | 0.018 |

Notes: See notes to table 4.2.

** = significant at 0.05 level. * = significant at 0.10 level.

5 years before the interview. When women are also considered, however, past labour market experience becomes more important to explain participation (columns (1) and (2)), reflecting women stronger dependence on past conditions when deciding about participation. Working is also more likely among agents with higher levels of wealth, consistent with the hypothesis that risk may affect working decisions. Finally, the composition of the

household is only important if women are being considered, suggesting that differential fixed cost of working are particular important among them.

Table 4.2 shows the results for the earnings equations. The skill premiums are stronger when considering women into the analysis, consistent with results reported in Blundell, Dearden, Goodman and Reed (97) on the returns to education by gender. The initial levels of human capital differ mainly across ability groups, not so much by education within ability levels. This is also observed in the selection process (see table 4.1) and together with the fact that the main outcome does not change with the model being used explains our preference for the specification in column (1). Thus, the initial level of human capital is assumed to be homogeneous within ability group, independently of the level of education, $h^{\theta s} = h^{\theta}$, and the wage rates are defined in terms of initial levels human capital among agents of ability type 1. Once education and ability are controlled for, experience does not seem to play a very important role in determining men's earnings at this stage of their life (see columns (3) and (4)). The numbers suggest that the important distinction is that between the start of the working life, a potentially apprenticeship and experimental phase, and the following stages. Women earnings appear to be more sensitive to previous experience but since this does not affect the other results we have chosen to include them in the following analysis (columns (1) and (2)). Having worked in the recent past, however, is systematically estimated to have no impact in current earnings, independently of the type of model being adopted.²⁴

Table 4.3: Standard error for the productivity shock.

| | | Men and Women | | | | Men only | | | |
|--------------|--|---------------|-------|-------|-------|----------|-------|-------|-------|
| | | (1) | | (2) | | (3) | | (4) | |
| | | Coef. | SE | Coef. | SE | Coef. | SE | Coef. | SE |
| SE for π | | 0.260 | 0.002 | 0.259 | 0.002 | 0.276 | 0.004 | 0.276 | 0.004 |

Finally, the estimates for the standard error of the productivity shock are presented in table 4.3. Notice that the figures for this parameter, the one to be used on the following analysis, appear to be robust to the specification chosen.

²⁴This result is consistent with the non-depreciation assumption.

4.5.2 The rates of human accumulation

Three levels of ability were considered in the estimation and simulation process. Tables 4.4 and 4.5 show, respectively, the estimates for ν and r by ability and education levels.²⁵ Table 4.6 characterises the distribution of the productivity shock. Estimates are based on the model described by equations (4.19) and (4.20) and the upper bound for the distribution of π has been estimated by the iterative procedure described in Steps 3 to 7 of the estimation procedure.

Table 4.4: Initial rate of human capital accumulation.

| | By ability and education. | | |
|-----------------|-------------------------------------|-------------------------------------|-------------------------------------|
| | Low skills | Medium skills | High skills |
| Ability level 1 | 0.271** (0.050) (0.171,0.346) | 0.281** (0.118) (0.098,0.494) | 0.195** (0.077) (0.088,0.331) |
| Ability level 2 | 0.298** (0.059) (0.179,0.388) | 0.259** (0.082) (0.112,0.390) | 0.239** (0.045) (0.160,0.314) |
| Ability level 3 | 0.242** (0.065) (0.123,0.364) | 0.281** (0.078) (0.136,0.394) | 0.295** (0.047) (0.219,0.378) |

Notes: Estimation of the rates of human capital accumulation used data from the NCDS58 for agents aged 23, 33 and 42 years of age. This is part of a maximum likelihood estimation of the earnings selection model. Estimation used 4,148 observations of men and women. A table in the appendix to this chapter splits the observations by cells. Standard errors in parenthesis: estimates used bootstrapping with 150 replications. Bias-corrected 90percentile below the standard errors: estimation used the same bootstrap results.

** = significant at 0.05 level. * = significant at 0.10 level.

The values displayed in table 4.4 stand for the initial 5 years rate of human capital accumulation and table 4.5 presents the 5 years adjustment rates. These are central values, for agents experiencing an average productivity/efficiency shock. Individual specific rates depend on the particular productivity shock experienced. For instance, the figure presented on the top left of Table 4.4 is the rate of human capital accumulation for the first 5 years of an agent's working life when supplying unskilled labour and endowed with ability

²⁵The total number of observations used to compute these parameters is 4,148. The number of observations by cell is detailed in the appendix to this chapter.

Table 4.5: Adjustment rate of the rate of human capital accumulation.

| By ability and education. | | | |
|---------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| | Low skills | Medium skills | High skills |
| Ability level 1 | 0.662** (0.152) (0.366,0.874) | 0.379 (0.355) (0.070,0.724) | 0.636 (0.401) (0.084,1.372) |
| Ability level 2 | 0.636** (0.147) (0.384,0.866) | 0.576** (0.158) (0.285,0.798) | 0.687** (0.208) (0.339,0.986) |
| Ability level 3 | 0.682** (0.181) (0.414,0.968) | 0.628** (0.142) (0.370) | 0.803** (0.135) (0.544,0.978) |

Notes: See notes to table 4.4.

** = significant at 0.05 level. * = significant at 0.10 level.

Table 4.6: Distribution of the productivity shock.

| Estimates | |
|-------------|------------------------------------|
| Upper bound | 3.11** (0.599) (2.417,4.326) |

level 1. Together with table 4.5, the results suggest that low ability workers accumulate less human capital than other ability groups, and the speed of accumulation decreases with education. Medium ability individuals accumulate human capital at similar rates independently of the level of skills. As expected, high ability individuals are the most efficient in terms of human capital accumulation in high skilled jobs.

The visualisation of the rates of human capital accumulation facilitates the interpretation of the presented coefficients and their comparison. The plot of human capital accumulation rate by (θ, s) -groups is displayed in figure 4.1. It shows the relationship between the human capital accumulation rate (y axis) and the level of human capital (x axis) by ability and education. Plotted values correspond to five years accumulation rates for agents experiencing an average productivity shock.

All the curves exhibit similar shapes, dictated by the fact that all estimates for $r(\theta, s)$ are lower than 1. However, some curves clearly dominate others in terms of the speed

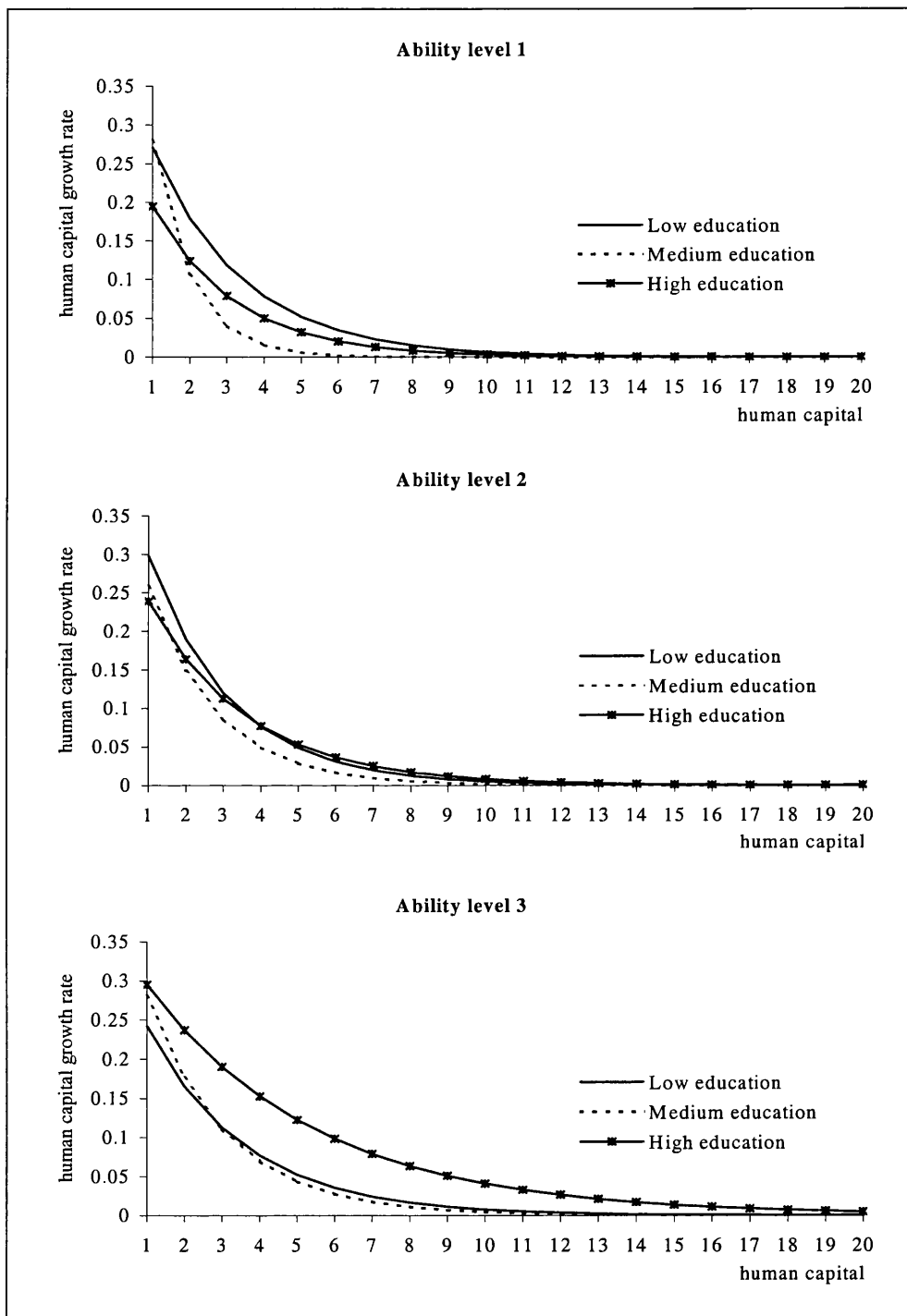


Figure 4.1: Rates of human capital accumulation by ability and educational levels.

of human capital accumulation. For instance, ability level 1 performs always better in unskilled jobs. On the other hand, ability level 2 seems to be almost equally productive in all types of jobs, doing only slightly better in unskilled jobs. On the contrary, ability level 3 clearly performs much better in high skilled jobs.

The bounds computed for the productivity shock are displayed in table 4.6. When compared to the standard error presented above, they are large enough to impose only mild truncation on the distribution of the shocks.

4.5.3 Sensitivity analysis

There are two data reasons for the estimates of the rates of human capital accumulation presented above to be biased. First, data observations are ten years apart while estimates refer to five years' periods. The reason for such difference pertains the requirements of the simulations performed in the next chapter. This estimation procedure is carried out under the goal of producing a global model of the economy that reflects the main empirical facts surrounding the implementation of the NDYP. The length of the time periods adopted in the model is of 5 years, requiring an adjustment to make the estimates consistent with the structural specification. And Second, ten years is a long time and the rate of human capital accumulation is expected to adjust throughout such period as agents become more knowledgeable. We study the consequences of such facts in what follows and present estimates that help to bound the true human capital accumulation rates.

Sensitivity test 1: non-matching estimation and data periods

Lets start by focusing on the consequences of using a different time period in estimation from what is supplied by the data. Since the former time span is the double of the latter, one can write the accumulation of human capital equation that has actually been estimated (results in tables 4.4 and 4.5) as follows,

$$\begin{aligned} \ln E_{t+2,a+2}^{\theta s} - \ln E_{ta}^{\theta s} &= [\ln W_{s,t+2} - \ln W_{st}] + \\ &\nu(\theta, s)r(\theta, s)^{h_{ta}^{\theta s} - h^{\theta s}} \pi_{ta} + \nu(\theta, s)r(\theta, s)^{h_{t+1,a+1}^{\theta s} - h^{\theta s}} \pi_{t+1,a+1} + \\ &[\ln \pi_{t+2,a+2} - \ln \pi_{ta}] \end{aligned} \quad (4.21)$$

The data does not directly reveal either h or π . Thus, an additional identification assumption is required to overcome the additional missing data problem of not observing the agents behaviour in every period of the analysis. The identification assumption used to estimate equation (4.21) is the following,

A8 The unobserved productivity shocks for periods t and $t + 1$ are assumed to take the same value, that is, $\pi_{ta} = \pi_{t+1,a+1}$.

Now notice that assumption (A8) is expected to cause the estimates of the human capital accumulation rate to be downward biased. This is because the observations being used in the estimation procedure are those of workers for the whole period. For these agents, one expects their reservation policy to decrease with time as they accumulate human capital. But this means they will be willing to stay at work later in life even if experiencing a relatively low productivity shock. On average, therefore, experienced productivity shocks are expected to fall among workers as they age. But then, by setting the predicted $\pi_{t+1,a+1}$ equal to the predicted π_{ta} one expects to incur in an upward bias on average. To compensate for that, \hat{r} will be downward biased. Notice that ν is unlikely to be strongly affected since it is estimated with the first observations for each individual, relying only on the predictions for π at the start of the working life.

Table 4.7: Sensitivity test 1.

| Human capital accumulation rates by ability and education. | | | |
|--|------------|---------------|-------------|
| | Low skills | Medium skills | High skills |
| Initial human capital accumulation rate $\nu(\theta, s)$ | | | |
| Ability level 1 | 0.205 | 0.321 | 0.204 |
| Ability level 2 | 0.222 | 0.283 | 0.229 |
| Ability level 3 | 0.185 | 0.275 | 0.286 |
| Adjustment rate $r(\theta, s)$ | | | |
| Ability level 1 | 0.892 | 0.675 | 0.824 |
| Ability level 2 | 0.863 | 0.841 | 0.811 |
| Ability level 3 | 0.845 | 0.862 | 0.876 |

Notes: See notes to table 4.4 .

Table 4.7 presents estimates under an alternative identification assumption,

A4' The unobserved productivity shocks for period $t + 1$ and $t + 2$ are assumed to take the same values, that is, $\pi_{t+1,a+1} = \pi_{t+2,a+2}$.

Contrary to the case established by A4, assumption A4' creates an upward bias in the estimates of the human capital accumulation rate since the predicted shocks between observations, $\pi_{t+1,a+1}$, are downward biased on average. The numbers in table 4.7 confirm this guess. Figure 4.2 confirms that the starting level for the rates of human capital accumulation are close to what has been previously established but the slope becomes less pronounced. Nevertheless, the relative pattern is kept unchanged.

Sensitivity test 2: long time periods

A ten year distance between data points is a rather long time period and human capital evolves gradually over time. Since the rate of human capital accumulation changes with the level of human capital, one expects it to adjust slowly even if all the rest is kept unchanged. On the extreme case, the rate of human capital accumulation adjusts instantaneously as the agent accumulates knowledge, becoming

$$h' = h \left\{ 1 + \pi \nu(\theta, s) \int_h^{h'} r(\theta, s)^{u - \frac{h\theta s}{h'}} du \right\}$$

where h' is the next instant level of human capital when the agent starts with h and receives a shock π . The observed growth rate of earnings then becomes,

$$\begin{aligned} \ln E_{t+2,a+2}^{\theta s} - \ln E_{ta}^{\theta s} = & \quad (4.22) \\ [\ln W_{s,t+2} - \ln W_{st}] + \pi_{ta} \nu(\theta, s) \int_{h_{ta}}^{h_{t+2,a+2}} r(\theta, s)^{u - \frac{h\theta s}{h_{t+2,a+2}}} du + [\ln \pi_{t+2,a+2} - \ln \pi_{ta}] \end{aligned}$$

Again, such observation raises suspicions about the estimates presented in tables 4.4 and 4.5. Once more, the figures presented for r are expected to be downward biased to compensate for the fact that a smaller h is being considered, namely the beginning of the period value. An alternative upper bound for the rates of human capital accumulation is presented in table 4.8, which uses the following specification for the estimation of the

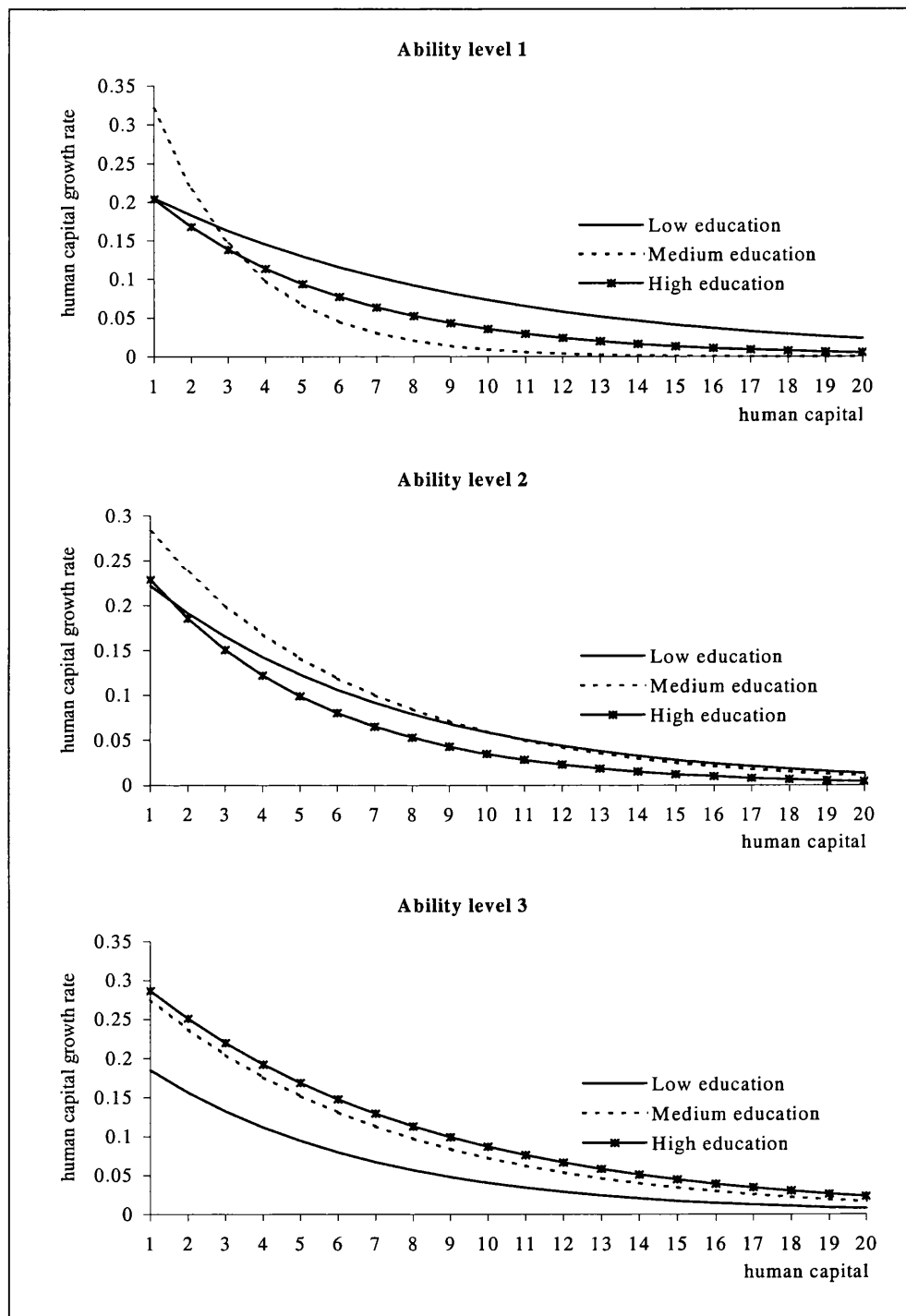


Figure 4.2: Sensitivity test 1 - Upper bound to the rates of human capital accumulation by ability and education.

rates of human capital accumulation,

$$\begin{aligned} \ln E_{t+2,a+2}^{\theta s} - \ln E_{ta}^{\theta s} &= [\ln W_{s,t+2} - \ln W_{st}] + \\ &\quad \nu(\theta, s)r(\theta, s)^{h_{t+1,a+1}^{\theta s} - h^{\theta s}} \pi_{ta} + \nu(\theta, s)r(\theta, s)^{h_{t+2,a+2}^{\theta s} - h^{\theta s}} \pi_{t+1,a+1} + \\ &\quad [\ln \pi_{t+2,a+2} - \ln \pi_{ta}] \end{aligned}$$

That is, instead of using the beginning of the period level of human capital, the end of the period value is adopted to adjust the growth rate, supposedly an higher level than the ones in between observations.

Table 4.8: Sensitivity test 2.

| Human capital accumulation rates by ability and education. | | | |
|--|------------|---------------|-------------|
| | Low skills | Medium skills | High skills |
| Initial human capital accumulation rate $\nu(\theta, s)$ | | | |
| Ability level 1 | 0.252 | 0.433 | 0.270 |
| Ability level 2 | 0.277 | 0.360 | 0.335 |
| Ability level 3 | 0.238 | 0.348 | 0.359 |
| Adjustment rate $r(\theta, s)$ | | | |
| Ability level 1 | 0.929 | 0.901 | 0.913 |
| Ability level 2 | 0.918 | 0.924 | 0.914 |
| Ability level 3 | 0.901 | 0.940 | 0.940 |

Notes: See notes to table 4.4.

The results to this test resemble less what was obtained in the original estimation procedure though the curves depicted in figure 4.3 show the same pattern as before. The initial levels, however, are set at generally higher values and the slopes are significantly less pronounced than before. However, a stronger response to this test is expected since taking the end of the period level of human capital significantly slows the learning process requiring a significant change in the accumulation rates.

4.6 Conclusions

The present chapter discusses the estimation of earnings equations in a labour supply framework with endogenous human capital formation. A structural model of the agents

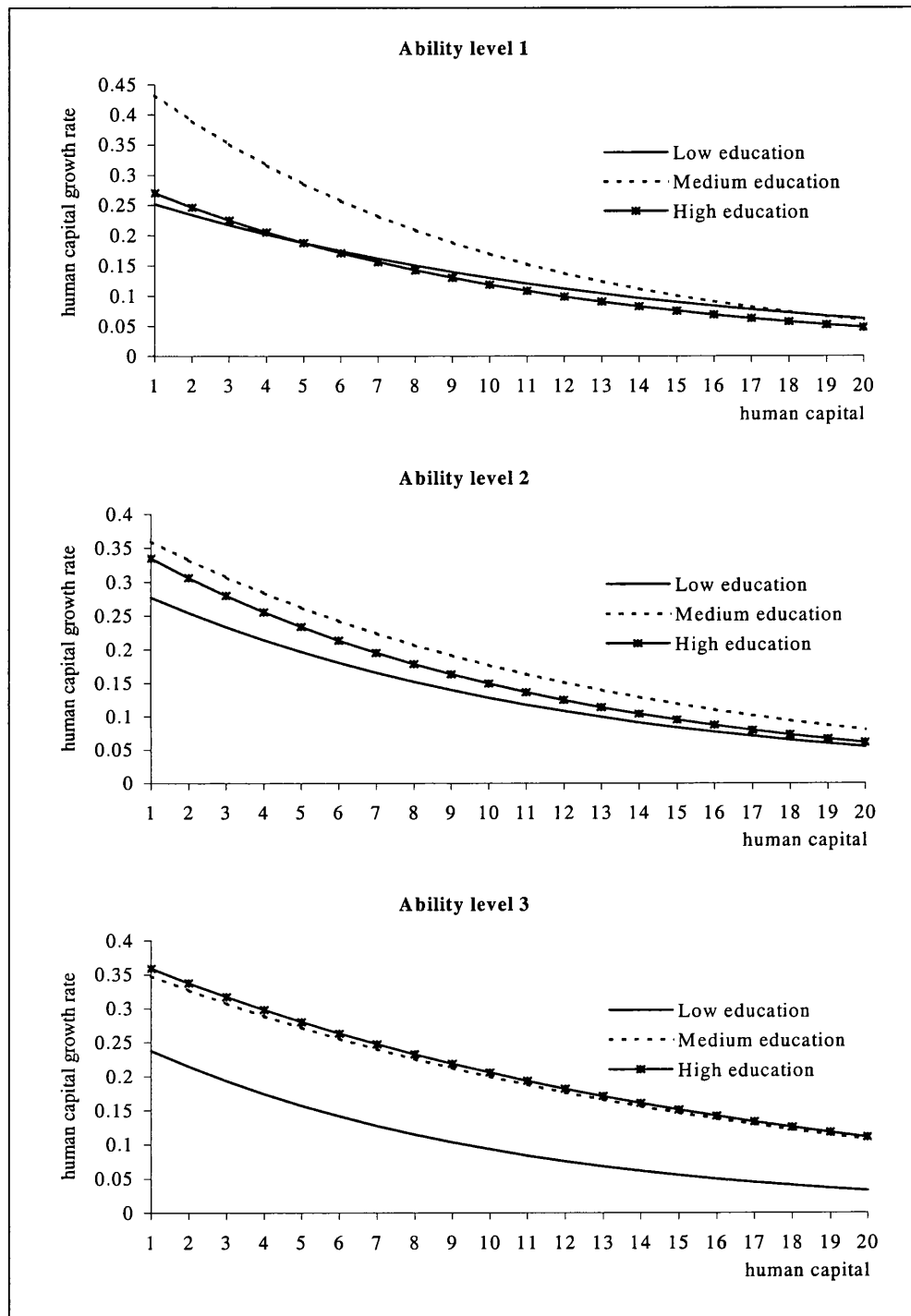


Figure 4.3: Sensitivity test 2 - Upper bound to the rates of human capital accumulation by ability and education.

working life is proposed and explored to develop a simple estimation technique that avoids complicated and time consuming likelihood evaluations. The main strategy is to use past, present and future information about labour market behaviour to infer about the present working experience. The model is estimated using data on the NCDS58, a British cohort data set of individuals born in one week of 1958. The results suggest that the adopted measure of ability does influence agents' learning performance on the job. Moreover, accumulated knowledge makes it more difficult to continue accumulating additional human capital. This is reflected on the shape of the curves depicting the rates of human capital accumulation by level of previously accumulated human capital. The qualitative results are not too sensitive to the unmatched time periods in data and in the model or to the duration of the time periods being considered. The latter, however, does impact on the measured size of the rates of human capital accumulation. Future work with different data sets may shed some light to the actual importance of this issue. The results from this chapter are to be used as an input in the analysis performed in the next one to simulate the reactions of individuals to the introduction of labour market policies.

4.7 Appendix to chapter 4

4.7.1 Proof of Lemma 2

The proof uses induction. To simplify the notation, we omit the time and age indexes from the expressions whenever it is clear what is meant.

Last period of life: age A

At this age the agent maximises his/her present utility, which implies following the consumption rule,

$$c_{tA}^* = (1 + R_t) k_{tA} + d_{tA} \pi_{tA} h_{tA} W_{st} + (1 - d_{tA}) B$$

Whether the agent decides to work or not depends on the realisation of π for any given X . Thus, the value function is,

$$\begin{aligned} V_{tA}^{\theta s}(k_{tA}, h_{tA}, \pi_{tA}) &= u((1 + R_t) k_{tA} + \pi_{tA} h_{tA} W_{st}) && \text{if } d_{tA} = 1 \\ V_{tA}^{\theta s}(k_{tA}, h_{tA}, \pi_{tA}) &= u((1 + R_t) k_{tA} + B_t) && \text{if } d_{tA} = 0 \end{aligned}$$

and is, therefore, strictly increasing, C^2 and concave. The expected value function is,

$$EV_{tA}^{\theta s}(k_{tA}, h_{tA}) = V_{tA}^{\theta s}(\cdot | d_{tA} = 0) P(d_{tA} = 0) + V_{tA}^{\theta s}(\cdot | d_{tA} = 1) P(d_{tA} = 1) = \\ \int_{\underline{\pi}}^{\pi^R} V_{tA}^{\theta s}(\cdot | d_{tA} = 0) dF(\pi) + \int_{\pi^R}^{\bar{\pi}} V_{tA}^{\theta s}(\cdot | d_{tA} = 1) dF(\pi)$$

where $F(\pi)$ is the cdf of the productivity shock, which is assumed to be independent of time and cohort, and $V(\cdot | d)$ represents the conditional value function. But since both $V_{tA}^{\theta s}(\cdot | d_{tA} = 0)$ and $V_{tA}^{\theta s}(\cdot | d_{tA} = 1)$ are strictly increasing, C^2 and concave in k and π^R does not depend on k , the expected value function is also strictly increasing, C^2 and concave in k .

Age between 1 and $A - 1$

Suppose that $EV_{t+1, a+1}^{\theta s}(k, h)$ is strictly increasing, C^2 and concave in k . We want to prove that $EV_{ta}(k, h)$ also enjoys from the same properties. The proof follows in steps.

1 *The conditional value functions $V_{ta}^{\theta s}(\cdot | d)$ are increasing, C^2 and concave in k .*

Given that u and $EV_{t+1, a+1}^{\theta s}$ are C^2 , strictly increasing and concave in c_{ta} and $k_{t+1, a+1}$, standard analysis shows that $V_{ta}^{\theta s}(k, h, \pi | d)$ is strictly increasing, C^2 and concave in k .²⁶ One can now apply the envelope theorem to yield,

$$\frac{\partial V_{ta}^{\theta s}(k, h, \pi | d)}{\partial k} = \frac{\partial u(c)}{\partial k} = (1 + R) \frac{\partial u(c)}{\partial c}$$

2 *The reservation value for the productivity shock π_{ta} is continuous in k .*

The characteristic of the reservation value is to solve the equality between value functions for different working choices. Therefore, given the continuous differentiability of this conditional value functions, the implicit function theorem may be applied to ensure that the reservation policies are continuously differentiable functions of k_{ta} .

3 *The value function $EV_{ta}^{\theta s}(k, h)$ is a C^2 function of k_{ta} .*

Given the reservation policies stated in Lemma 1, the expected value function at age

²⁶See Stokey and Lucas (1989, chapter 9) for a detailed discussion.

a can be written as follows,

$$\begin{aligned} EV_{ta}^{\theta s}(k, h) &= V_{ta}^{\theta s}(\cdot | d_{ta} = 1) P(d_{ta} = 1) + V_{ta}^{\theta s}(\cdot | d_{ta} = 0) P(d_{ta} = 0) \\ &= \int_{\pi^R}^{\bar{\pi}} V_{ta}^{\theta s}(\cdot | d = 1) dF(\pi) + \int_{\underline{\pi}}^{\pi^R} V_{ta}^{\theta s}(\cdot | d = 0) dF(\pi) \end{aligned}$$

Given the continuity and continuous differentiability of the joint density function of the shocks and of the reservation values as a function of k , it is assured that the expected value function is C^2 .

4 The value function $EV_{ta}^{\theta s}(k, h)$ is an increasing and concave function of k .

The first derivative of $EV_{ta}^{\theta s}(k, h)$ is given by,

$$\begin{aligned} \frac{\partial EV_{ta}^{\theta s}(k, h)}{\partial k_{ta}} &= \int_{\pi^R}^{\bar{\pi}} \frac{\partial V_{ta}^{\theta s}(\cdot | d = 1)}{\partial k_{ta}} dF(\pi) + \int_{\underline{\pi}}^{\pi^R} \frac{\partial V_{ta}^{\theta s}(\cdot | d = 0)}{\partial k_{ta}} dF(\pi) + \\ &\quad \frac{\partial \pi_{ta}^R}{\partial k_{ta}} \left(V_{ta}^{\theta s}(\cdot | d = 0) - V_{ta}^{\theta s}(\cdot | d = 1) \right) dF(\pi_{ta}^R) = \\ &= \int_{\pi^R}^{\bar{\pi}} \frac{\partial V_{ta}^{\theta s}(\cdot | d = 1)}{\partial k_{ta}} dF(\pi) + \int_{\underline{\pi}}^{\pi^R} \frac{\partial V_{ta}^{\theta s}(\cdot | d = 0)}{\partial k_{ta}} dF(\pi) \end{aligned}$$

which is always positive since the conditional value functions are strictly increasing.

Thus, EV is an increasing function of k .

The second derivative is given by,

$$\begin{aligned} \frac{\partial^2 EV_{at}(k, h, s)}{\partial k^2} &= \int_{\pi^R}^{\bar{\pi}} \frac{\partial^2 V_{ta}^{\theta s}(\cdot | d = 1)}{\partial k^2} dF(\pi) + \int_{\underline{\pi}}^{\pi^R} \frac{\partial^2 V_{ta}^{\theta s}(\cdot | d = 0)}{\partial k^2} dF(\pi) + \\ &\quad \frac{\partial \pi^R}{\partial k} \left(\frac{\partial V_{ta}^{\theta s}(\cdot | d = 0)}{\partial k} - \frac{\partial V_{ta}^{\theta s}(\cdot | d = 1)}{\partial k} \right) dF(\pi^R) \end{aligned}$$

The two first terms are guaranteed to be negative given the concavity of the conditional value functions. However, for an agent with decreasing absolute risk aversion, the last term is potentially positive. To see why, let's consider the third term a bit more in detail. Under decreasing absolute risk aversion, the derivative of the reservation value π^R with respect to k is negative: the wealthier the individual, the more willing he/she is to take the risk. For precisely the same reason, the derivative of

$\partial V_{ta}^{\theta s}(\cdot | d = 1)$ with respect to k is larger than that of $\partial V_{ta}^{\theta s}(\cdot | d = 1)$. Therefore, this term is positive. Now notice that the less responsive the reservation policies are in respect to k , the less important the last three terms are. Concavity is guaranteed when considering an utility function with a constant coefficient of absolute risk aversion. On the contrary, very responsive policies to changes in k make the kink in the value function more difficult to smooth out. ■

4.7.2 Proof of Lemma 3

Given the properties of the expected value function, the values for the state variables at age a and the working and schooling decisions, the optimal condition for c is,

$$\frac{\partial u(c_{ta})}{\partial c_{ta}} = \beta(1 + R_t) \frac{\partial EV_{t+1,a+1}^{\theta s}(k_{t+1,a+1}, h_{t+1,a+1} | h_{t+1,a+1})}{\partial k_{t+1,a+1}}$$

Let $I_{ta} = (1 + R_t)k_{ta} + d_{ta}\pi_{ta}h_{ta}W_{st} + (1 - d_{ta})B_t$ be the total net income at age a and time t and, keeping d constant, suppose it increases. The law of motion for k (equation (4.2)) implies that either c_a or k_{a+1} increase. But they both must increase given the concavity of $EV_{t+1,a+1}$ and u in k and c , respectively. ■

If there is a change in d_a , however, consumption may cease to be normal. The following example clarifies why this is so. Suppose that for a small change in π the agent changes his/her mind about working. Every other state variables staying the same, there is a change in I_a . Suppose $I_a(d_a = 1) < I_a(d_a = 0)$. But the total expected income in the future must be larger under the working option (or otherwise the agent would not take the risk) overcoming the monetary cost of investing in human capital today. In such case, consumption under the non-investment may be lower than under the investment case at the same time that the present net income is higher.

4.7.3 The NCDS58 data

Tables 4.9 describes the attrition in the panel study NCDS58 by gender and determines the number of observations used to compute the rates of human capital accumulation.

The distribution of the observations used in the final stage of the estimation is presented in table 4.10. It is evident from the figures that individuals characterised by the Ability Level 1 concentrate very strongly on unskilled labour, explaining the lack of precision

Table 4.9: Observations in the NCDS58.

| | Men | Women | Total |
|--------------------------------|-------|-------|--------|
| At birth | 9,593 | 8,960 | 18,553 |
| With a measure of ability | 8,625 | 8,133 | 16,758 |
| At 23 years of age | 6,267 | 6,270 | 12,537 |
| At 33 years of age | 5,605 | 5,795 | 11,400 |
| At 42 years of age | 5,628 | 5,789 | 11,417 |
| Continuously employed up to 33 | 2,226 | 592 | 2,818 |
| Continuously employed up to 42 | 1,922 | 398 | 2,320 |

Table 4.10: Observations used to compute the rate of human capital accumulation.

| | Low skills | Medium skills | High skills | Total |
|---|------------|---------------|-------------|-------|
| Continuously employed up to 33 years of age | | | | |
| Ability level 1 | 437 | 94 | 31 | 562 |
| Ability level 2 | 362 | 244 | 149 | 755 |
| Ability level 3 | 156 | 309 | 444 | 909 |
| Total | 955 | 647 | 624 | 2,226 |
| Continuously employed up to 42 years of age | | | | |
| Ability level 1 | 357 | 83 | 32 | 472 |
| Ability level 2 | 304 | 222 | 125 | 651 |
| Ability level 3 | 128 | 272 | 399 | 799 |
| Total | 789 | 577 | 556 | 1,922 |

found for estimates of the rate of human capital accumulation among skilled agents of that group.

Chapter 5

The overall evaluation of labour market interventions

Most of the literature on labour market programmes' evaluation, from the seminal papers by Ashenfelter (78), Ashenfelter and Card (85) and Heckman and Robb (85 and 86) to the more recent surveys by Heckman, LaLonde and Smith (98) on methodological issues and Katz (96) on the evaluation of wage subsidy policies, has been focusing on the direct effects of the policies. These are commonly understood as the direct effects on participants or the direct effect on a randomly chosen individual. Individual information is used to assess the benefits of the intervention, the main challenge being that of identifying the correct counterfactual. A crucial assumption of these methods is that non-participants outcomes are not affected by the existence of the programme, exhibiting the same outcome that would have been observed were there no programme taking place. That is, the existence of indirect effects is ruled out. The importance of such assumption depends mainly on the characteristics of the programme being evaluated. A small, very focused programme, or one characterised by very loose rules, is unlikely to have strong implications on the wide economy. On the other hand, a programme with global implementation, defined by precise rules and effective treatments is expected to have broader effects through substitution, displacement and prices.

The present study focuses on the evaluation of the long-run, economy-wide effects of labour market policies. This is done on an heterogeneous-agent dynamic general equilib-

rium (GE) model. We depart from the New Deal for the Young People (NDYP), the UK labour market programme targeted at the under-25 that grounds this research. Its stated aim is to improve the labour market attachment of these individuals. The programme provides help and advice for job search, training and education for the least skilled, and access to subsidised employment, among other options (see Bell, Blundell and Van Reenen, 00, Blundell, Costa Dias, Meghir and Van Reenen, 01, Van Reenen, 01, and the introduction to this thesis for a more detailed description of the programme). Because the NDYP is a global programme in the sense that all eligible individuals are expected to participate, GE effects are more likely to occur, meaning that the impact of the programme is probably not restricted to the direct effect it has on participants. Market conditions and prices are likely to be affected too, spreading the effect of such intervention through the whole economy and affecting non-eligible groups. This may happen not only because the programme is costly and financed through taxes on agents that may not be the direct participants, but also because the amount of human capital and the process of selection into work are themselves changed.

To clarify this latter point, let's consider one of the main possibilities the NDYP creates: the chance of having a subsidised job for a period of time. Eventually, the labour supplied by subsidised workers is cheaper, and therefore more desirable. In order to benefit from the subsidy, firms may be willing to take advantage of the benefit when offering jobs, substituting from unsubsidised to subsidised workers. Also, firms may try to replace their workers for new, subsidised ones, the so-called displacement effect. On the other hand, selection and depreciation of skills implies that low-skilled workers should constitute a disproportionately large share of the long-term unemployed, for whom working is not an attractive option given their productivity levels. The new subsidy may change this too, creating the possibility for these workers to earn above their productivity levels for some time while building up new skills that will improve their future prospects. If the NDYP succeeds in improving the skills and employability of the long-term unemployed, the ultimate outcome is a compositional change in the pool of unemployed towards shorter spells and more participative agents, rising the supply of labour and putting a downward pressure on wages.

An evaluation of the kind of effects just described can hardly be done with some kind

of comparison between participants and non-participants outcomes. The main reasons are twofold: First, there is no comparable group out there from which one can draw the appropriate counterfactual since the economy at wide is affected heterogeneously; and Second, the time span of the available data is typically too short to show the overall effects, which are likely to take some time to build up.

The solution adopted here is to construct a global, GE model of labour supply, designed to provide an overview of how the programme works through the economy affecting participants, non-participants and prices. Such a model is then estimated and simulated to provide measures of the long-run effects of a policy intervention and comparisons to what would be identified under the *no spillover* assumption.

To accomplish this task, an overlapping generations model is constructed in the Auerbach and Kotlikoff (87) tradition. The point of departure is the recent developments in empirical dynamic GE models designed to assess the impact of different public policies. Heckman, Lochner and Taber (98a) present a GE model of skills and human capital formation grounded on empirical evidence from the micro literature. The model is applied to evaluate the impact of tuition subsidies on skills and human capital investment decisions (Heckman, Lochner and Taber, 98b and 99). GE effects are found to be very significant, making the usually estimated partial equilibrium (PE) effects potentially misleading. Other applications of the same structural model include the analysis of different tax schemes and their impacts on human capital formation (Heckman, Lochner and Taber, 98c and 99, and Taber, 00).

Another contribution is by Lee (01), who analyses the impact of education policies within a GE model of career choices including labour supply decisions along with skills and human capital formation. Individuals are allowed to move freely between different types of jobs (skills) without any costs other than foregone utility in the rivalrous occupations. As expected, GE effects are considerably smaller than in the Heckman et al. (98a) model.

Tax credit programmes are analysed in Cossa, Heckman and Lochner (99). The main focus of this study is on how wage subsidies affect the accumulation of human capital depending on the initial position of the agent in the scheme. The study compares the predictions by two frequently used models: learning by doing and on-the-job training. Absent from the analysis, however, is an attempt to characterise the overall impact of the

considered policies. A GE approach to the evaluation of social programmes is presented in Mortensen and Pissarides (02) and Richardson (97a and 97b). Wage subsidies' effects on unemployment rates, flows into and from unemployment, and wages are evaluated within a model of job search and matching. However, no account is taken for the role of human capital and exogenously acquired skills define completely separate job markets.

This latter feature is in fact common to most of the literature in the subject. Another example is the evaluation of the effects of the impact of social security reforms by Huggett (96). Again, no account is taken for skills and human capital heterogeneity and formation.

The goal of many of the social policies released, however, is to change skills and human capital of participants and through it improve their labour market perspectives. Moreover, empirical evidence rejects the efficiency units approach and instead supports the comparative advantage principle, showing that skills and human capital are important characteristics in the labour market.¹ We therefore follow Heckman et al. (98a) approach by explicitly modelling skills and human capital formation. As in their model, we incorporate human capital in the model, which can be accumulated on the job, consider the existence of different non-perfectly substitutable skills that can be acquired through formal education, and acknowledge the existence of several dimensions of heterogeneity concerning age and ability, which impact on the type of skills acquired and on the amount of human capital accumulated while working.

This chapter, however, attempts to evaluate a programme of different characteristics of that studied in Heckman et al. (98a). In particular, we are interested in producing an assessment of a wage subsidy policy, meaning that some modifications are required to the original Heckman et al. (98a) model for such a programme to be at all meaningful. For this we consider three main directions. *First*, and following Lee (01), labour supply is included in the model. This is a central feature given the stated goal of evaluating active labour market policies targeted at the unemployed. It is done by considering 1/0 decisions on working, consistent with the finding that participation decisions explain most of the variation in hours worked (Pencavel, 86, Blundell and MaCurdy, 99, and Browning, Hansen and Heckman, 99). We also consider the existence of fixed costs from working in the form of unemployment benefit (as in Low (99)). Such a feature alone is responsible

¹Evidence is presented in Heckman and Sedlacek (85, 90), Topel (86) and Sattinger (93).

for a decision design that is very close to the 1/0 pattern imposed in the present study.

Second, as in Huggett (97), Huggett and Ventura (99) and Lee (01) we allow for idiosyncratic uncertainty, in the present case affecting the working and studying costs/revenues. The shocks have both a transitory and a permanent component, consistent with empirical evidence on the importance of transitory wage shocks to explain enrolment in the programmes (Ashenfelter, 98, Heckman, LaLonde and Smith, 98, Heckman and Smith, 99). Given the independence of shocks across individuals and the dimension of the economy, the model is still one of perfect certainty at the aggregate level and perfect foresight of prices is assumed. The individual level uncertainty is responsible for additional heterogeneity as it influences both savings, working and studying decisions. However, and contrary to what is assumed in Lee (01), we consider preferences under risk aversion consistent with results by Barsky, Juster, Kimball and Shapiro (97) on the proportion of risk averse people and Kimball (90), Carroll (94), Attanasio (99) and Banks, Blundell and Brugnolini (99) on the importance of precautionary savings. Given that studying and working are risky options, such feature is responsible for under-investment in human capital at both the intensive (working) and extensive (studying) margins and an excessively high skill-premium. The additional assumption of decreasing absolute risk aversion makes poorer individuals particularly penalised by their responses to risk and inequality likely to rise even more.

Finally, we also consider the existence of a government which collects taxes from workers' earnings to finance an unemployment insurance and other potential interventions. This aspect is especially important given the great deal of attention that has been recently allocated to the cost/benefit analysis of the labour market policies (Heckman, LaLonde and Smith, 98 and van Reenen, 01). A detailed discussion of more efficient policies and degree of intervention may then be attempted.

The outline of the paper is as follows. The first section presents and discusses the GE model and some of its properties. It draws heavily on the results from the previous chapter in setting up the individual's problem. Section 2 discusses the identification issues, again resorting to previous results on the estimation of the human capital production function. Section 3 discusses the effects of wage subsidies and presents the results from simulations of two policies of different generosity levels. The rich simulated data is a good source of information not only for a discussion of the potential importance of labour

market policies, but also for the identification of the economic mechanisms underlying such reactions and for the analysis of the relative size of GE effects when compared with PE ones. The adequacy of control groups drawn from non-participants to provide the missing counterfactual in evaluations of the direct effects of interventions can, therefore, be assessed as well. We use the simulated behaviour of 15,000 agents in each specified economy to address each of these issues. Section 4 presents some concluding remarks and section 5 includes the appendix to this chapter.

5.1 The Model

This section discusses the GE model of labour supply used to evaluate the impact of labour market policies. The goal of evaluating a policy of the NDYP type will be made clear from the discussion as many of the structural assumptions chosen reflect features of this programme. It is important to notice that this is a global intervention, with wide implementation, targeted at the young, unskilled workers. Participation is compulsory and a range of potential treatments is available, including a tax credit scheme which is the policy being evaluated in this chapter. In what follows the model is formally presented.

Given the stated goal of evaluating labour market policies, we develop a model of labour supply and education. The framework developed in the previous chapter is used and extended to accommodate endogenous education decisions. It is imbedded in a OLG setup, where agents live for a fixed number of periods, A . Such specification allows for policies to be targeted at specific age-groups as is the case with the NDYP. The overall environment is one of GE, considered to provide clues about the whole effect of the programme. Our economy is composed by a continuum of heterogeneous individuals that invest in education, supply human and physical capital for production and live on consumption.

Evidence suggests that participants in the NDYP are drawn mostly from the least educated population (Bell et al., 99). Treatment, however, is not sufficiently long and intense for them to become part of the most educated group unless it motivates future investment in education. To accommodate this fact, we consider three different skills acquired through education and corresponding to the following levels: the basic level, achieved by 16 years of age and equivalent to less than 5 O-levels ($s = 1$), completion of

secondary school or A-levels ($s = 2$) and graduation from college ($s = 3$). Treatment may affect unskilled agents at the margin of investing in education but does not directly make them highly skilled workers. At the individual level, the amount of human capital owned by an agent with educational attainment s is h_s . At the aggregate level, the three types of skills define three different production factors potentially not perfectly substitutable in production. These are denoted by H_1 , H_2 , and H_3 for, respectively, low, medium and high skills.

The discussion that follows is split in three parts, corresponding to the production sector, the individual's problem and the equilibrium conditions.

5.1.1 The production sector

The aggregate production function is assumed to use the three types of human capital along with physical capital in a constant returns to scale technology. The human capital utilised in the production process is decomposed in three types denoted by H_1 , H_2 and H_3 , the indexes standing for the level of skills corresponding, respectively, to low, medium and high levels of formal education. We assume a constant elasticity of substitution (CES) specification for the production function, meaning that the elasticities of substitution between two inputs are constant. This is a simple and parsimonious specification, and one that allows for an easy control of the degree of substitutability between inputs. A nested CES specification is adopted by separating out the unskilled labour. This choice is motivated by the fact that attention is focused on this group, the target of social policies of the NDYP kind. Following results by Heckman et al. (98a), Blundell and Bond (00) and Blundell, Bond and Windmeijer (00), physical capital is assumed to account for a constant share in production.² Thus, the time t aggregate production is specified as follows,

$$Y_t = K^\eta \left\{ p H_{1t}^\alpha + (1-p) [q H_{2t}^\sigma + (1-q) H_{3t}^\sigma]^{\alpha/\sigma} \right\}^{(1-\eta)/\alpha} \quad (5.1)$$

²Our initially preferred specification accounts for the possibility of capital-skill complementarity, as in Krussell, Ohanian, Rios-Rull and Violante (00). It was established by considering high skilled labour to exhibit a special relationship with physical capital and such aggregate to be potentially more complementary to medium skilled labour than to low skilled labour. This was, however, a rather complicated specification, and one difficult to make consistent with the available data.

where the parameters α and σ define the degree of substitutability between inputs and η determines the physical capital share.

5.1.2 The individual's problem

For completeness, this section presents a brief discussion of the agent's problem as it bears strong similarities to the model presented in chapter 4. As before, the agent's problem is to decide optimally about each period level of consumption and activity. Besides working and staying at home, possible activities include investment in education. Again, the three options are taken to be mutually exclusive, the decision being taken at the extensive margin.³ Education is the source of different skills, potentially non-substitutable in production and commanding jobs with different learning contents. Consumption is assumed to be of an homogeneous good and is where the agent derives utility from. For the sake of simplicity, we have ruled out wealth effects as they especially affect the upper tail of the income distribution and are not, therefore, the main focus of a study designed to evaluate programmes targeted at the most disadvantaged workers. A dynamic decision process is considered so that the impact of the programme can be measured throughout the individual's life. Finally, the agent is assumed to be rational, making decisions based on an inter-temporal utility function.

The agents are assumed to be heterogeneous along a number of dimensions. First, of course, there are A generations living together at each moment in time. Second, innate ability varies across individuals, affecting the on-the-job learning process. And third, markets are incomplete through idiosyncratic uninsurable uncertainty. A major source of uncertainty in the agent's life concerns the *returns to knowledge*. It is modelled as a productivity shock, which affects current earnings and future levels of human capital would the agent decide to work. The productivity shock being considered directly affects contemporaneous earnings by adjusting the wage rate to the individual-specific productivity level and indirectly affects future earnings through its effect on the period-specific learn-

³See Blundell and MaCurdy, 99, and Browning, et al. (99) for evidence on the importance of decisions at the extensive margin - whether to participate - as compared to decisions at the intensive margin - whether to supply more time.

ing ability.⁴ This later feature is rationalised by the learning-by-doing framework being considered, where working and learning are non-rivalrous activities. It means that the more effort is dedicated to work, the higher the return in terms of accumulated knowledge or human capital.⁵ Uncertainty also affects the direct cost of education, affecting how much the agent needs to pay if deciding to take such investment. Studying and working are, therefore, potential risky decisions, and labour market choices are made under risk aversion (Attanasio, 99).⁶ At the aggregate level, however, no uncertainty is considered and we assume there is perfect foresight in what respects to the market prices.

The following variables formalise the problem. Consumption, c , working status, d , and studying status, i , describe the agent's decision. The options taken will determine the evolution of some of the state variables, namely assets, k , skills, s , and human capital, h . Skills and human capital are distinguished on the basis that skills, acquired through formal education, determine the *type* but not the *amount* of human capital supplied by the agent. Schooling prepares the individual for different, possibly more demanding tasks. On the other hand, on-the-job experience determines the *amount* of human capital through learning by doing but has no effect on its *type* so that the resulting human capital is perfectly substitutable in production for the previous one. Finally, investment in education is irreversible in the sense that only the highest acquired skill can be supplied. Thus, unskilled workers cannot apply for skilled jobs and skilled workers cannot apply for unskilled jobs.

The rational agent maximises an inter-temporally separable lifetime utility. Each period utility, u , is derived from consumption only as wealth effects have been excluded. u is assumed to be strictly increasing and concave in c . Labour market activity decisions are summarised by the working status (d) and the studying status (i). d is 1 whenever the agent decides to work and i is 1 whenever education is the preferred activity. If staying at

⁴Transitory wage shocks have long been acknowledged as an important component of the treatment selection process: the so-called Ashenfelter dip is a recurrently observed characteristic of the participants' wages data (Heckman and Smith, 99).

⁵Cossa et al. (99) find learning by doing to be more consistent with the data than the rivalrous activities that compose the on-the-job training model.

⁶Empirical evidence on the importance of risk aversion and precautionary savings can be found on Barsky, Juster, Kimball and Shapiro (97), Attanasio (99) and Banks, Blundell and Bruggavini (01).

home both d and i are 0. Decisions are based on the state of the nature the agent faces, completely described by the ability type of the agent (θ), time period (t), age (a), amount of assets (k), level of schooling (s) which determines the type of human capital (skills), the amount of human capital (h), the productivity shock (π) and the shock on the cost of education (ξ) (the latter two variables are the random components of the model):

$$(\theta, t, a, k, s, h, \pi, \xi)$$

An individual of type θ born at t' solves the following life-cycle problem,

$$\max_{\{c, d, i\}_{a=1}^A} E \left\{ \sum_{a=1}^A \beta^a u(c_{ta}) \mid (R_t, W_{1t}, W_{2t}, W_{3t}, B_t, T_t, \tau_t)_{t=t'}^{t'+A-1} \right\}$$

where β is the discount factor and $(R_t, W_{1t}, W_{2t}, W_{3t}, B_t, T_t, \tau_t)_{t=t'}^{t'+A}$ represents the sequence of prices the agent faces over his/her entire life - R_t is the interest rate, W_{st} is the wage rate of skill s , B_t is the unemployment insurance, T_t is the direct cost of education and τ_t is the tax rate imposed on earnings.

The laws of motion for the state variables complete the setup of the problem. Equations (5.2), (5.3) and (5.4) specify the dynamics for the levels of assets, education and human capital, respectively,

$$k_{t+1, a+1} = (1 + R_t)k_{ta} + d_{ta}h_{ta}\pi_{ta}W_{st}(1 - \tau_t) + (1 - d_{ta})[B_t - i_{ta}(T_t + \xi_{ta})] - c_{ta} \quad (5.2)$$

$$s_{t+1, a+1} = s_{ta} + i_{ta} \quad (5.3)$$

$$\begin{aligned} h_{t+1, a+1} &= h_{ta}(1 + \nu(\theta, h_{ta}, s_{ta})\pi_{ta}) && \text{if } d_{ta} = 1 \\ h_{t+1, a+1} &= \underline{h_{s+1}^\theta} && \text{if } i_{ta} = 1 \\ h_{t+1, a+1} &= h_{ta} && \text{if } d_{ta} = i_{ta} = 0 \end{aligned} \quad (5.4)$$

Equation 5.4 dictates that any accumulated amount of human capital of a given skill-type is not transferable to other skills. Though this is an extreme hypothesis, its importance should be assessed within the particular application we are studying, namely the evaluation of interventions targeted at rather disadvantaged workers who have experienced long unemployment spells. Such individuals are less likely to use previously acquired

human capital if significantly improving their skills by getting an important amount of education.

The only differences from what has been described in chapter 4 concerns skills acquisition and its impacts on the other state variables. If deciding to invest in education, the agent accumulates an additional level of skills (equation (5.3)), must pay for the direct cost of the investment which is individual- and time-specific ($i(T + \xi)$ in equation (5.2)) and loses the human capital accumulated under lower skills (equation (5.4)). The reason for this *non-transferability of human capital between skills* assumption is again related with the design of the NDYP. It is targeted at the very unskilled long-term unemployed workers, for whom past working experience is unlikely to be of any value for future jobs if in the meanwhile they decide to invest in education. Equation (5.4) also reflects the fact that no depreciation of human capital is assumed if the agent decides to stay at home. All individuals are equal at birth apart from the characteristic θ which determines their on-the-job learning ability: they all live for the same number of periods, A , being endowed with the same amount of human capital, the same level of schooling, and the same amount of assets. The conditions at birth can be described as,

$$\begin{aligned} k_1 &= 0 \\ s_1 &= 1 \\ h_1 &= \underline{h_1^\theta} \end{aligned}$$

and the budget constraint corresponds to the restriction $k_{A+1} \geq 0$.

Hence, the recursive of the model can now be written as

$$V_{ta}^\theta(k_{ta}, h_{ta}, s_{ta}, \pi_{ta}, \xi_{ta}) = \max_{c_{ta}, d_{ta}, i_{ta}} \left\{ u(c_{ta}) + \beta E_{\pi\xi} V_{t+1, a+1}^\theta(k_{t+1, a+1}, h_{t+1, a+1}, s_{t+1, a+1}, \pi_{t+1, a+1}, \xi_{t+1, a+1}) \right\}$$

where V_{ta}^θ stands for the value function of an agent of type θ aged a at time t and EV_{ta}^θ is its expected value while the information about contemporaneous shocks has not been disclosed.⁷

The results derived in chapter 4, which characterise the individuals' problem, apply straightforwardly to this more general setting. Lemmas 1 and 2 can be re-written as

⁷Notice again that the adopted specification is equivalent to one where consumption and leisure are

follows,

Lemma 5 *Given the particular conditions faced by an individual at a certain age, the working and studying decisions can be described by reservation values for the work efficiency (π) and the direct cost of education (ξ) when the shocks are uncorrelated over time:*

The agent prefers working to staying at home whenever $\pi > \pi^R$.

The agent prefers studying to staying at home whenever $\xi < \xi^R$.

The agent prefers studying to working whenever $\xi < \xi^R(\pi)$.

Lemma 6 *On a bounded and convex state space, if the joint density function of the idiosyncratic shocks on the working efficiency and tuition fees, π and ξ , is continuously differentiable (C^1), and the coefficient of absolute risk aversion is decreasing but not “too much” so, then the expected value function $EV_a(k, h, s)$ is a strictly increasing, C^2 and concave function of k .*

Lemma 3 applies with the same wording. Proofs are omitted as they resemble very closely the ones presented in chapter 4. These results characterise the individual’s decision process, establishing the reservation policies and the Euler equation as the optimal rules.

perfect substitutes. This time, studying is also associated with some disutility in the following way

$$u_{ta}(c, d, i) = f(c + \alpha_d(1 - d) + \alpha_{di}(1 - d)(1 - i))$$

Considering the following transformation of the consumption variable,

$$\bar{c} = c + \alpha_d(1 - d) + \alpha_{di}(1 - d)(1 - i)$$

the consumers problem can again be re-written in terms of the new composite consumption which also includes leisure by performing the following additional transformations,

$$\bar{B}_t = B_t + \alpha_d + \alpha_{di}$$

$$\bar{T}_t = T_t + \alpha_{di}$$

That is, there is a monetary equivalent to the utility cost of studying as well as working that is independent of the shape of the utility function, units or level of utility.

5.1.3 Characterising the equilibrium

At each moment in time, equilibrium is characterised by a set of prices such that aggregate demand equals aggregate supply. We assume that ours is a small open economy in a world with free capital movements, so that the interest rate R is exogenously determined. The wage rates, however, are determined endogenously (Card and Lemieux, 00, present evidence on the responsiveness of wages to changes in the relative supply of different skills in Canada, the US and the UK). There is no aggregate uncertainty, so that equilibrium satisfies the following equilibrium conditions,

$$H_{st}^D = H_{st}^S \quad \text{with } s = 1, 2, 3 \quad (5.5)$$

$$K_t^D : \frac{\partial Y_t}{\partial K_t^D} = R_t \quad (5.6)$$

where the superscripts D and S denote demand and supply, respectively, and capital letters stand for aggregate variables.

Wage rates are determined in a competitive equilibrium, and so each input price simply equals its marginal productivity. Thus, the optimal demand for human capital solves the following equations,

$$H_{st}^D : \frac{\partial Y_t}{\partial H_{st}^D} = W_{st} \quad \text{with } s = 1, 2, 3$$

The level of human capital used in production is the aggregation of individual's labour supply. Conditional on aggregate prices, the latter is well defined almost surely under perfect foresight, making the former also well defined. This is shown in the next lemma,

Lemma 7 *Take a set of aggregate prices at any time t , $(R_t, W_{1t}, W_{2t}, W_{3t}, B_t, T_t, \tau_t)$. For a continuous joint density function of the idiosyncratic shocks on productivity and tuition fees, π and ξ , the measure of the individual's indifference set is zero. If in addition each (a, θ) -group is formed by a continuum of individuals, the aggregate labour supply is uniquely defined and deterministic.*

Proof. See appendix to this chapter (section 5.5.1). ■

Lemma 7 insures that the supply of human capital is uniquely determined for a given set of prices. Lets further postulate that the distribution of θ is constant across cohorts,

the joint distribution of (π, ξ) is independent of time, cohort and ability and the shocks are independent over time and across agents. The aggregate supply of human capital of type s can then be represented as follows,

$$H_{st}^S = \sum_{a=1}^A \int_{\theta \in \Theta} \int_{\pi(a) \in \Pi^a} \int_{\xi(a) \in \Xi^a} h_{tas}^S(\theta, \pi(a), \xi(a)) C(t, a) dF(\pi_1, \xi_1) \dots d(\pi_a, \xi_a) dG(\theta) \quad (5.7)$$

where

$$h_{tas}^S(\theta, \pi(a), \xi(a)) = \mathbf{1}[s_{ta}(\theta, \pi(a), \xi(a)) = s] * d_{ta}(\theta, \pi(a), \xi(a)) * h_{ta}(\theta, \pi(a), \xi(a)) * \pi_{ta} \quad (5.8)$$

In equation (5.7), Θ , Π and Ξ stand for the support of the ability parameter, productivity shock and cost of education shock, respectively. $C(t, a)$ is the size of the cohort aged a at time t . Finally $\pi(a)$ and $\xi(a)$ in equations (5.7) and (5.8) represent the history of shocks experienced by the agent up to age a , so that $\pi(a) = (\pi_1, \dots, \pi_a)$ and $\xi(a) = (\xi_1, \dots, \xi_a)$. It is made explicit that the schooling level, amount of human capital and decision on whether to work and study depend on age, ability type and shocks experienced up to age a . There is also a dependence on the time period as it incorporates information on the prices faced over the agent's entire life. Worth noting, each period prices may depend on the distribution of wealth as measured by assets, k , and so will individual decisions. This is because the agents are assumed to be risk averse, and how much risk they are willing to take for a potential gain may depend on how wealthy they are. This issue will be further discussed later on.

The following results characterise the aggregate labour supply H_{st}^S as defined in equation (5.7).

Lemma 8 *If the support of the ability parameter has finite measure and the support of the productivity shock, π , is bounded, the aggregate supply of human capital of type s , $H_{st}^S(W_1, W_2, W_3 | R, B, \tau)$, is bounded above and below.*

Proof. See appendix to this chapter (section 5.5.2). ■

Lemma 9 *Under the established conditions, the aggregate supply of human capital of type s , $H_{st}^S(W_1, W_2, W_3 | R, B, \tau)$, is a continuous function of the wage rates.*

Proof. See appendix to this chapter (section 5.5.3). ■

Lemma 10 *Under the established conditions, the aggregate supply of human capital of type s , $H_{st}^S(W_1, W_2, W_3 | R, B, \tau)$, is a non-decreasing function of W_s but may not be monotonic with respect to the other wage rates.*

Proof. See appendix to this chapter (section 5.5.4). ■

Finally, the model is closed with the accounting of the public sector. It is assumed that the government collects taxes on workers' earnings and redistribute them through subsidies to the non-workers. The government budget is required to be balanced, which means that the following equality must be verified,

$$\begin{aligned} \sum_{s=1}^3 \sum_{a=1}^A \int_{\theta \in \Theta} \int_{\pi(a) \in \Pi^a} \int_{\xi(a) \in \Xi^a} h_{tas}^S(\cdot) W_{st} \tau C(t, a) dF_1 \dots dF_a dG(\theta) &= \quad (5.9) \\ &= \sum_{a=1}^A \int_{\theta \in \Theta} \int_{\pi(a) \in \Pi^a} \int_{\xi(a) \in \Xi^a} [1 - d_{ta}(\cdot)] B_t C(t, a) dF_1 \dots dF_a dG(\theta) \end{aligned}$$

where h_{tas}^S is defined in equation (5.8) above and F_i stands for the cdf of the shocks at age i , $F(\pi_i, \xi_i)$, $i = 1, \dots, a$. The lhs on the government budget constraint (5.9) represents total collected taxes in period t and the rhs stands for total expenses.

The existence of an equilibrium can now be established.

Lemma 11 *Lets take an economy composed of a number of groups defined by age and innate ability, (a, θ) , each group being formed of a continuum of individuals. Suppose in whole, the measure of the economy is finite. Under the present setting, taking the state space of individual level variables to be a convex bounded set in all the continuous real state variables and a finite set in the discrete state variables, the steady state equilibrium exists for particular choices of the government-established prices, τ and B .*

Proof. See appendix to this chapter (section 5.5.5). ■

5.2 Identification issues

This section presents the fundamental aspects related with the identification of the parameters of the model and discusses the main results. The structural set-up is used whenever

possible to provide conditions for identification. This follows the strategy discussed in chapter 4 and the results on the rates of human capital accumulation are imported from there (see estimates displayed in tables 4.4, 4.5 and 4.3). Thus, the present analysis focus only on the remaining parameters of the individual's problem and the parameters of the production function.

5.2.1 The aggregate technology

The identification of the parameters of the production function (5.1) requires knowledge of the aggregate amount of human capital used in production. By acknowledging that agents are heterogeneous in what concerns to the type and amount of human capital they are able to offer and to how they value participation in the labour market, it becomes clear that the information on the aggregate human capital inputs is not readily available in any data set. In fact, since past working decisions depend on the aggregate and individual conditions faced, the commonly used variables on the total number of working hours or total number of workers cannot provide an accurate measure of the amount of utilised human capital. Even when adjusted for experience, such measures are not adequate because individuals learn at different speeds depending on their own abilities and specific experience on the job.

The method devised to approximate the aggregate human capital series makes use of different information sources as in Heckman et al. (98a). The structural implications of the individual's model are used to establish the wage and participation equations and to compute the wage rates. This is discussed in the previous chapter and formally established by the earnings' models (4.7) and (4.8) among youngsters and (4.11) and (4.8) as the agents get older. In a second step, the following accounting relationship is used to approximate the aggregate levels of utilised human capital (H^s , $s = 1, 2, 3$),⁸

$$H_t^s = \frac{TW_{st}}{W_{st}} \quad (5.10)$$

⁸The true wage rates are identified up to a multiplicative constant $\underline{h}^{\theta=1}$, the homogeneous initial level of human capital among individuals of ability level 1 and any educational status. Such transformation is innocuous, only re-defining the units of measurement by setting $\underline{h}^{\theta=1} = 1$. As a consequence, the aggregate levels of human capital are also identified up to a constant.

where TW_{st} stands for the total wage bill for human capital of type s at time t . Once W_s and H_s are determined, the first order conditions for the firm's problem provide the structure needed to estimate the production function parameters under competitiveness.

To do so, however, one first needs to measure TW_{st} . Its approximation by skills' type relies on a cross-section of households in the UK, the General Household Survey (GHS). The GHS is an annual survey running since 1973 that collects very detailed individual level information. Under the present requirements, a particularly important characteristic of the GHS concerns to the adopted education classification. It matches perfectly the one found in the NCDS, the dataset used to estimate the rates of human capital accumulation and the wage rates in the previous chapter. Thus, using a skills' classification similar to the one applied to the NCDS data, one can construct a measure of the aggregate wage bill by educational level. To do so, we simply aggregate the total weekly earnings for all working agents. Utilised human capital can now be identified from equation (5.10) with data on the wage rates and the total wage bills by skill level.

Estimates on W_{st} come from the analysis in the previous chapter. The earnings selection model estimated among youngsters provide an estimate for the wage rates in 81 and the earnings' growth rate regressions simultaneously provide the wage rates' growth rates.

Since there are only three periods of wage information (81, 91 and 00), we do not attempt to estimate the parameters of the production function but solely to choose a reasonable set of parameters consistent with the data. Table 5.1 shows the approximated parameters.

Table 5.1: Aggregate production function parameters.

| | value | elasticity of substitution |
|---|-------|----------------------------|
| α (low skilled - more skilled labour) | 0.12 | 1.1 |
| σ (medium skilled - high skilled labour) | 0.80 | 5.5 |
| p | 0.40 | - |
| q | 0.46 | - |

The displayed results suggest a large but finite elasticity of substitution between the two levels of skilled labour. Unskilled labour, however, shows up to be less substitutable

to skills than other estimates in the literature (Card and Lemieux, 00, Heckman et al., 98a). Ours, however, are more homogeneous groups in terms of education (skills), and commonly estimated elasticities stand between the two values just identified.

The missing piece of information to completely determine of the aggregate production function is the share of physical capital in production, η . We have chosen to calibrate it to 0.4, following the estimates presented in Blundell and Bond (00). Exogeneity of the interest rate and separability of physical capital in production minimise the potential impact of such choice.

5.2.2 Education decisions and unemployment benefit

Decisions about education depend on its cost, which is determined by the aggregate level of tuition fees and the individual specific shock. Similarly to what is done with regard to the human capital production function in chapter 4, we start by parameterising the distribution of ξ to be a truncated normal distribution $N(0, \sigma_\xi)$ in a domain with symmetric bounds $\xi = -\bar{\xi}$. Identification of T , $\bar{\xi}$ and σ_ξ uses data moments on the aggregate levels of education for the 1958 cohort and selects the combination of parameters that produce similar simulated moments. The same technique is used to approximate the unemployment benefit, B , this time using the aggregate proportion of unemployed individuals.

The education related parameters are displayed in table 5.2 and its performance is analysed in Table 5.3.⁹¹⁰

It is clear from table 5.3 that the model produces over-investment in the medium level of education throughout life and under investment in the high level of education among the youngsters. It does, however, produce a pattern similar to what is empirically observed.

The unemployment benefit has been approximated to 0.37. This figure is below the average (gross) earnings at the beginning of life for an unskilled worker - which amount to 0.47. This is a slightly large figure as compared to true values, where the unemployment

⁹Reported ages correspond to the life periods in the simulated model. In the formal model, the highest level of education can only be achieved by the end of the second period of life, corresponding to 25 years of age. These are the numbers reported under the simulated moments. Data numbers, on the other hand, correspond to 23 years of age.

¹⁰Approximation uses only the aggregate levels of education, not distinguishing between different ability groups.

Table 5.2: Education related parameters.

| | value |
|---------------|-------|
| T | 0.5 |
| σ_ϕ | 1.5 |
| $\bar{\phi}$ | 1.0 |

Table 5.3: Investment in education: Data vs. simulated moments.

Proportion of individuals with different levels of skills.

| | age=3 | age=5 | age=7 |
|-------------------|-------|-------|-------|
| NCDS58 data | | | |
| low skills | 0.557 | 0.527 | 0.459 |
| medium skills | 0.257 | 0.244 | 0.247 |
| high skills | 0.186 | 0.229 | 0.294 |
| Simulated moments | | | |
| low skills | 0.522 | 0.413 | 0.384 |
| medium skills | 0.392 | 0.368 | 0.351 |
| high skills | 0.085 | 0.218 | 0.265 |

benefit accounts to about half the gross earnings of unskilled young (16-20) workers. The main reason relates to the large length of the periods being considered, responsible for the relatively high risk incurred if staying unemployed. As before, Table 5.4 compares the data and simulated moments concerning unemployment rates, showing that, once more, the simulated pattern resembles the empirically determined one.

Table 5.4: Working decisions: Data vs. simulated moments.

| Unemployment rates. | | | |
|---------------------|-------|-------|-------|
| | age=3 | age=5 | age=7 |
| NCDS58 data | 0.237 | 0.203 | 0.147 |
| Simulated moments | 0.243 | 0.175 | 0.128 |

Table 5.5 presents some measure of the precision of the simulated model in predicting

Table 5.5: Labour market transitions: Data vs. simulated moments.

Proportion of individuals moving between employment and unemployment.

| | age 3 to age 5 | | age 5 to age 7 | |
|-------------------|----------------|--------------|----------------|--------------|
| | into | into | into | into |
| | employment | unemployment | employment | unemployment |
| NCDS58 data | 0.525 | 0.308 | 0.347 | 0.148 |
| Simulated moments | 0.562 | 0.213 | 0.697 | 0.156 |

the agents' labour market behaviour, this time in what concerns to transitions in the labour market. Transitions are particularly important in the present context since the frequency of unemployment spells is the major outcome of interest of the NDYP as publicly announced. Apart from transitions into employment between ages 5 and 7, all the other parameters are closely approximated in the simulations produced and discussed below.

5.2.3 Other parameters

Finally, information available in the literature is used to calibrate the remaining parameters. We start by assuming the isoelastic utility function

$$u(c) = \frac{c^{-\gamma}}{-\gamma}$$

and set γ to 0.56 following Attanasio, Banks, Meghir and Weber (99).

The exogenously determined interest rate, R , is set to 3% annually, the average rate for the UK between 1970 and 1986 as reported in Attanasio and Browning (95). Accordingly, the discount rate β is calibrated to 0.02 per year, producing a slightly increasing pattern of consumption over the life-cycle. Obviously, this behaviour is reinforced by risk aversion.

The model is completely determined by imposing a limit to the individual's life span. This is deterministically set to 10 periods of working life, each corresponding to approximately 5 working years, plus an additional two periods of retirement.

5.3 Labour market policies: assessing the overall effects of wage subsidies

This section discusses the impacts of a wage subsidy policy on different dimensions of the agents' lives. Following the design of the NDYP, the scheme is targeted at the young unskilled unemployed workers. It is handed to the agent if he decides to take up a job, though in the considered stylised economy this is equivalent to being handed to the firm when the employer is able to distinguish between subsidised and unsubsidised workers. The possibility of getting a subsidised job is modelled as a fourth labour market activity option available to young agents only. Under the policy, agents may choose to get a subsidised job in the first period of their working life. Eligibility, however, requires some time in unemployment prior to enrolment. This is modelled as a one year unemployment spell, corresponding in duration to one fifth of the periods being considered. Thus, at the start of their working life the agents must choose one of the four following labour market options available. First, enrol into education to acquire skills. Second, get an unsubsidised job to acquire knowledge in unskilled tasks and be paid for the work done. Third, wait for a while in unemployment (one fifth of the period) and subsequently enrol in a subsidised job to acquire knowledge in unskilled tasks and be paid for the work done during the remaining four fifths of the period on the top of the subsidy. And Fourth, remain unemployed for the whole period, collecting the unemployment benefit.

In what follows we discuss the long-run overall effects of such a wage subsidy scheme as measured from the simulations of the described economy. Two levels of generosity for the wage subsidy were tried. The first, under the title of *Policy 1*, amounts to one fourth of the average pre-programme net earnings among young unskilled workers, corresponding to 0.1 monetary units. The second, denoted by *Policy 2*, amounts to one half of the average pre-programme net earnings among young unskilled workers, corresponding to 0.2 monetary units. The “no-policy” economy, called *Baseline Case*, is also presented for comparison purposes.

The effects of the policy are measured along three main lines. The first two concern the incidence of unemployment and the acquisition of skills. They have been used to justify the creation of the NDYP and its design and are, therefore, essential in the evaluation

procedure. The third focus on the individual welfare and wealth impacts of the programme.

Simulation exercises as the one described not only provide information about the potential importance of the policies being experimented but they do so in a structural way, so that the economic mechanism generating the effects are understood. Further information is possible to extract from such studies, as the relative size of GE effects when compared with PE ones and the adequacy of available controls to provide the missing counterfactual if the evaluation is to be performed based on outcome's data. The following analysis addresses each of these points.

Before presenting the simulations, however, we shall discuss the lines along which such policy might work within the present framework. Expected effects are of different sorts, often affecting prices and labour market decisions in opposite directions. Thus, the sign of the overall impact of wage subsidies is unknown *a priori* under this type of setting. This is the subject of the next section, where we start by discussing the direct effects of the programme under perfect certainty. Price effects are then introduced in the discussion and uncertainty is considered at the end.

5.3.1 Directions of the programme effects

The most obvious effect from a tax credit policy of the type described above is to improve the desirability of unskilled work at the beginning of the agents life as there is an extra subsidy to be earned. The symmetric version of this effect is translated in the loss of attractiveness of unemployment and studying options. In particular, the subsidy creates extra indirect costs of education. Treated agents that succeed in accumulating an extra amount of human capital than they otherwise would find working in unskilled jobs more attractive in the future as well. This turns unemployment spells less likely for the whole life of the treated but also rises the costs of education permanently. However, not all participants become more knowledgeable in unskilled tasks as a consequence of treatment. Some among them would be employed in the no-programme world. For these, the time in work during the first period of life is diminished as a consequence of the eligibility rules, reducing their learning attainment. But then, future indirect costs of education are diminished for participants that would otherwise be employed, making education more desirable. For exactly the same reason, however, unemployment will also become more

attractive later in life.

Changes in prices depend on the relative size of the participants' group and its composition. A drop in the wage rate of unskilled labour suggests that the programme has a relatively strong impact in employing youngsters that would otherwise stay out of work, possibly keeping them unskilled in the future. On the contrary, a drop in the wage rate of skilled labour indicates that most of the participants would have decided to work in the first period of their life anyway and participation makes future educational investments more attractive. Quite obviously, changes in the skill premium will affect the composition of the groups by changing the incentives to work, study and enrol into the programme. Non-participants, for instance, may start investing more in education if the skill premium increases, possibly affecting their own future labour market choices towards working. The tax rate is also expected to change as movements in government expenses and revenues must be accommodated. However, whether it decreases or increases depends on whether the fall in the unemployment incidence compensates or not for the extra financial burden introduced by the subsidy.

Lets now consider what happens under uncertainty. To simplify the discussion, we present a brief overview of the effects of uncertainty at the individual and aggregate levels before focusing on its effects when a labour market policy is introduced.

By coexisting with an unemployment benefit and both direct and indirect costs of education, uncertainty is expected to affect individuals' working/education choices. In particular, risk aversion increases the odds of staying at home as working and studying are potentially risky investments. Since wealth effects are excluded from the analysis, the agent chooses to stay at home only if there is a chance of loosing from working or studying today. This happens when the future returns from working or studying, in terms of accumulated human capital or skills, are potentially not enough to cover the present cost of investment, as measured by the loss in income either from taking up a job that pays less than the unemployment benefit or in the form of tuition costs and lost earnings. Individuals, however, may not be affected similarly by a given risk, depending on how attitudes towards risk behave as a function of wealth. Individual's wealth at the start of a period is independent from the risk being evaluated in that period but affects the utility of the investment. The discussion around lemma 2 in the previous chapter made it clear

that under increasing absolute risk aversion, wealthier agents are more willing to run into a given risk and are therefore expected to invest more in education and to take up work more often.

Such behaviour patterns have consequences at the aggregate level that can be described along three main lines. First, this is an economy characterised by under investment in education and human capital through working, the potentially risky options. As the working and studying decisions occur less frequently than what would be desirable, the education (skill) premium and general wages are kept at higher levels than in a first best economy. Second, uncertainty affects the distribution of wealth/income given that investment in skills and human capital is more frequent at the upper tail of the distribution of wealth. Finally, notice that the impact on prices does not come exclusively through the under-investment behaviour. In fact, the distribution of wealth also determines prices by changing individuals' decisions. Therefore, prices are expected to be affected as a consequence of changes in the distribution of wealth in response to uncertainty.

The effects of a wage subsidy policy under risk aversion can now be described. Uncertainty accentuates the direct effects of the programme discussed above by increasing the risk of undertaking education and decreasing the risk of working among participants that succeed to improve human capital levels. Similarly, participants that loose in human capital accumulation also experience stronger impacts of the same kind already discussed. However, uncertainty also generates other effects. First, by changing the wealth accumulated at different stages of the life-cycle the programme affects individuals' willingness to work or study as risk aversion responds to wealth. And Second, changes in the distribution of wealth induce second order effects on prices by further affecting the composition of the skill- and working-groups.

5.3.2 Results from the simulations

Table 5.6 displays the steady state prices under the three experimented scenarios. Columns (1) and (2) contain the figures for the *baseline case*, where no labour market policy is taking place. Columns (3) to (6) present results for the same economy under *policy 1* and *policy 2*, the two wage subsidy policies being considered of different degrees of generosity (*policy 1* is the least generous).

Table 5.6: Steady state prices.

| | Baseline | | Policy 1 | | Policy 2 | |
|--------------------------|----------|-------|----------|-------|----------|-------|
| | gross | net | gross | net | gross | net |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| low skill's wage rate | 0.245 | 0.211 | 0.243 | 0.208 | 0.240 | 0.203 |
| medium skill's wage rate | 0.331 | 0.285 | 0.333 | 0.285 | 0.337 | 0.285 |
| high skill's wage rate | 0.428 | 0.368 | 0.430 | 0.368 | 0.435 | 0.368 |
| tax rate | 0.141 | | 0.144 | | 0.154 | |

The figures suggest that wage subsidies of the type being discussed increase the relative amount of unskilled labour being supplied by inducing young unskilled agents to work and, through skill-specific knowledge accumulation, to remain unskilled after treatment. Notwithstanding, the rise in the skill premium and wage rates for skilled labour registered under both policies may induce some of the participants to let go the experience accumulated on unskilled labour and invest in education. Non-participants find it more attractive to invest in skills once the policies have been introduced and may actually benefit from its existence. Adjustment in the tax rates reveal that the programme does not pay for itself, the rise in government expenses being particularly strong under the more generous *policy 2*. This is reflected on the net wage rates, which change only slightly for skilled labour when the policies are introduced and may even decrease as is the case for the price for high skills under *policy 2*.

Table 5.7: Programme participation rates.

| | Policy 1 | Policy 2 |
|---|----------|----------|
| (1) participation rate | 0.43 | 0.66 |
| (2) Proportion "otherwise unempl." among participants | 0.35 | 0.45 |
| (3) Proportion "otherwise unempl." participants among unempl. | 0.44 | 0.89 |

The high levels of participation and their composition explain the large additional expenses induced by the programmes (see table 5.7). About half of the young population participate in the programme but less than half of the participants are from the initial

target group. In fact, individuals that would be unemployed in the *baseline case* amount to only 35% and 45% of the participants group (row (2) in table 5.7). The others would be working or studying in the *baseline case* and participation may actually deteriorate their labour market perspectives, increasing instead of reducing the odds of future unemployment spells. Notwithstanding, the more generous programme, *policy 2*, drives most of the target population into the programme (row (3) in table 5.7).

The changes in prices reported above constitute a very aggregate measure of the impact of the programme. We shall now focus on the impacts of the programme on a number of individual level variables discriminating between types of agents. An important question rests on the differences between a GE overall analysis and the commonly adopted PE one. Comparisons between the two types of effects are also attempted below, where PE effects are computed by introducing the policy but not allowing the prices to change.

The incidence of unemployment

Figure 5.1 plots the changes in unemployment rates by age for the two policy cases and the type of effect being measured. The dotted lines stand for GE effects and full lines represent the PE effects. PE effects dictate a fall in unemployment rates as the programme induces an important share of the participants to remain out of benefits more frequently in the future. The size of the effect is larger under *policy 2* as it attracts a larger proportion of *otherwise unemployed* individuals. Nonetheless, these are all small numbers that get only close to 1 percentage point shortly after participation. The changes in prices occurred under GE, however, completely reverse the impact identified for PE. As the price for unskilled labour drops, unskilled agents seem to find working less attractive. It is informative to decompose such effects by groups according to the participation status and pre-programme activity. This is done in figure 5.2.

It is obvious from figure 5.2 that different groups are affected in opposite directions, the magnitude of the effects being much stronger than when they are taken together. This figure presents only the case for *policy 2*, which shows more pronounced effects due to its relative generosity.¹¹ The lhs graph in this figure shows a large decrease in the unemployment rates among participants that would otherwise be unemployed, the target

¹¹Qualitatively, the results from both policies are equivalent.

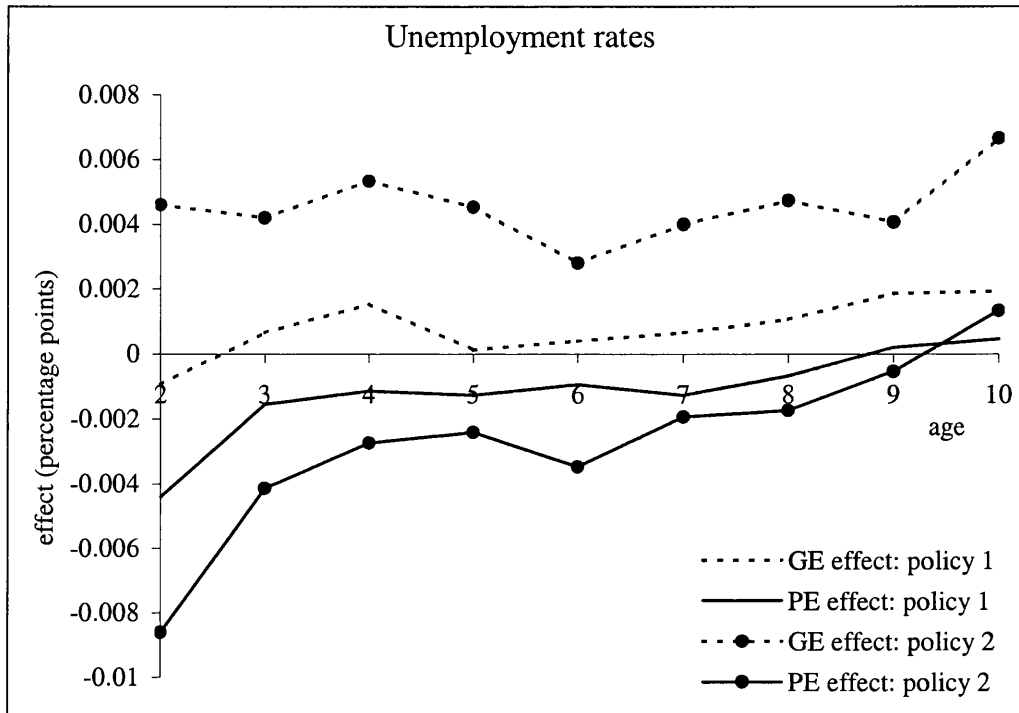


Figure 5.1: Impact of the wage subsidies on the working decisions.

group for the NDYP. Human capital accumulation during the treatment period seems to compensate for the drop in the unskilled wage rate, raising the likelihood of staying out of unemployment in the future by up to 3 percentage points. On the contrary, the labour market perspectives for non-participants in the target group deteriorate considerably, by an order of magnitude similar to that observed among *participants otherwise unemployed*. Thus, attempts to measure the impact of the programme on *participants otherwise unemployed* using *non-participants otherwise unemployed* as control group produces serious bias. The TTE could be over-estimated to the double due to the assumption of no indirect effects.

The rhs in figure 5.2 presents the case for individuals that would be employed in the *baseline case*. Again, both groups are affected by the existence of the programme, independently of the participation status. The drop in the price of unskilled labour strongly impacts on these agents' decisions by making employment less attractive, therefore raising the incidence of unemployment later in life. The *participants otherwise employed*

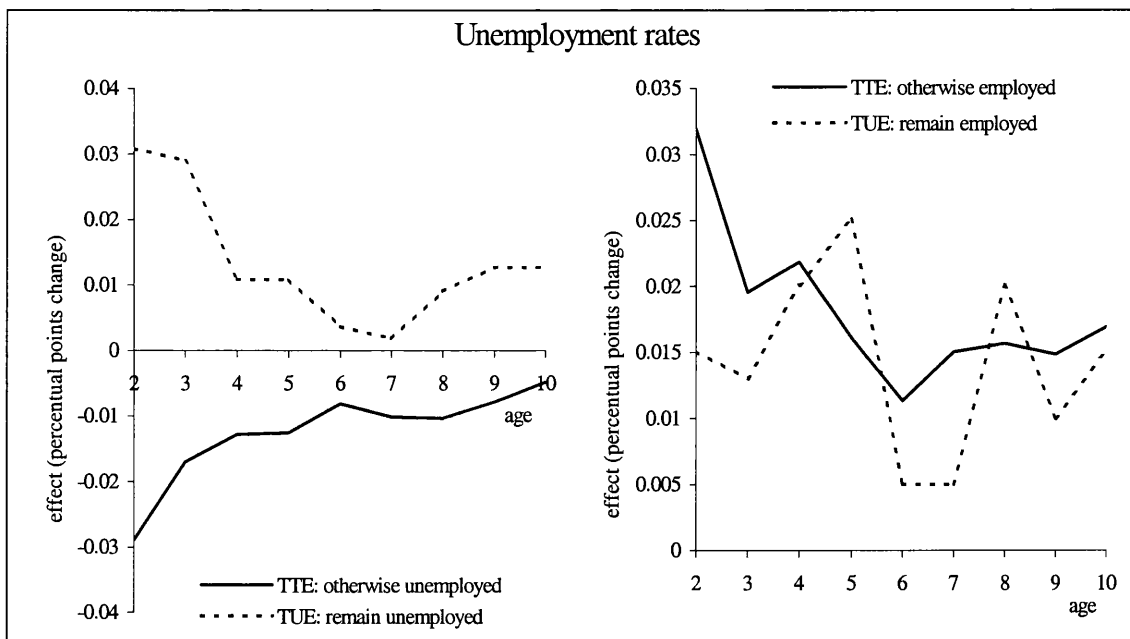


Figure 5.2: GE effects of the wage subsidy (policy 2) on the working decisions by treatment status and working status in the baseline case.

seem to be slightly more affected, at least shortly after treatment. Nevertheless, taking the *non-participants otherwise employed* as controls again introduces strong biases on the estimates of the programme effect, reducing them to near zero.

Investment in education

Figures 5.3 and 5.4 are a reproduction of figures 5.1 and 5.2 when the education decisions are being considered instead of working ones. The effects of the policy on the proportion of unskilled labour by age are plotted. Computed global effects reveal a significant drop in educational investments at the start of the working life (see figure 5.3). As expected, *policy 2* affects these decisions more deeply as it creates higher costs of education, but the pattern of the effect is similar independently of how generous the policy is. Under PE, the initial impact is expected to last for the rest of individuals' life, again showing the importance of additional human capital accumulated through treatment. However, when the wage rates are allowed to adjust, the initial effect of the programme fades away as large skill premiums improve the attractiveness of the education option.

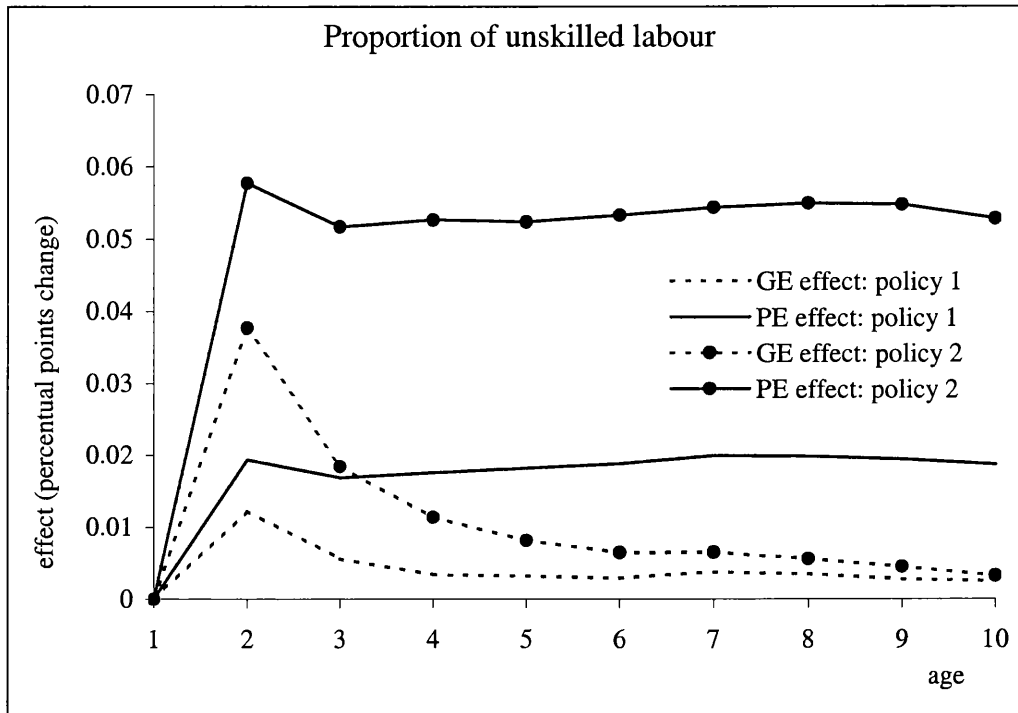


Figure 5.3: Impact of the wage subsidies on the investment in education decisions.

Figure 5.4 shows again that selecting similar agents that have kept out of the programme to construct the control group may be very misleading. The programme being analysed affects both participants and non-participants in this dimension as well, though the impact on non-participants is about half of the magnitude of that on participants when the proportion of agents remaining unskilled is the outcome of interest. Not all groups experience drops in the educational attainment as a result of the introduction of the programme. In fact, such a drop is specific to *participants otherwise unemployed*, the group that benefits the most from participation in terms of accumulated human capital. All the other groups presented respond to the raise in the skill premium by taking up education more frequently.

Distributional effects

This section analysis the impact of the programme on the levels of wealth and welfare in the economy and its distribution. Under risk aversion, the pattern of the wealth distribution

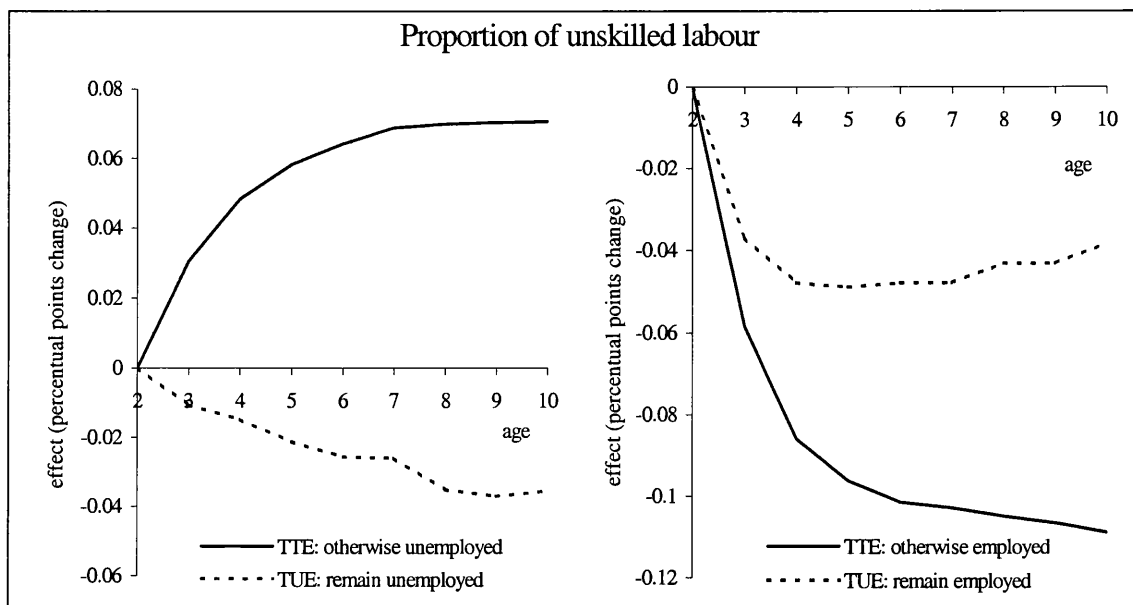


Figure 5.4: GE effects of the wage subsidy (policy 2) on the education decisions by treatment status and working status in the baseline case.

is particularly important as it affects individuals labour market decisions and, through that, prices. Thus, reducing the inequality in the wealth distribution has the potential of diminishing the gap to the first best in an economy of incomplete markets with uncertainty.

Table 5.8 characterises the distribution of net lifetime wealth and welfare under the three considered scenarios (columns (1) to (3) and (4) to (6), respectively). Average levels of wealth are kept almost unchanged by the introduction of the policies. Its distribution, however, becomes slightly less unequal especially due to improved life standards on the bottom of the distribution under *policy 2*. Table 5.9 details the origin of these changes in the wealth distribution by decomposing the group of individuals benefiting from the existence of the programme (columns (1) and (2)). The more generous policy benefits a very large proportion of *participants otherwise unemployed*, the original target group and the one that potentially includes a larger share of the most disadvantaged workers. In fact, agents benefiting from the programme are much more represented in this group than in any other when *policy 2* is being evaluated. The least generous *policy 1* presented in column (1), however, benefits more frequently the non-participants. This is mainly due

to their concentration on skilled jobs, taking advantage of the relative shortage of skills caused by the introduction of the wage subsidy. More generous policies, however, become rather expensive, exerting a pressure on public accounts that is reflected on the tax rates.

Table 5.8: Lifetime net wealth and welfare: descriptives.

| | Net wealth | | | Welfare | | |
|-----------|------------|----------|----------|----------|----------|----------|
| | Baseline | Policy 1 | Policy 2 | Baseline | Policy 1 | Policy 2 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| mean | 3.62 | 3.62 | 3.64 | -16.89 | -16.84 | -16.74 |
| st. error | 1.09 | 1.08 | 1.05 | 2.31 | 2.28 | 2.19 |
| p10 | 2.46 | 2.49 | 2.54 | -20.00 | -19.86 | -19.61 |
| p90 | 5.03 | 5.02 | 4.99 | -13.97 | -13.98 | -13.98 |
| p90/p10 | 2.04 | 2.02 | 1.96 | 0.70 | 0.70 | 0.71 |

Table 5.9: Proportions benefiting from the programme.

| | Net wealth | | Welfare | |
|--------------------------------------|------------|----------|----------|----------|
| | Policy 1 | Policy 2 | Policy 1 | Policy 2 |
| | (1) | (2) | (3) | (4) |
| among treated | 0.483 | 0.637 | 0.546 | 0.703 |
| among treated "otherwise unemployed" | 0.515 | 0.817 | 0.502 | 0.813 |
| among treated "otherwise employed" | 0.466 | 0.427 | 0.569 | 0.462 |
| among untreated | 0.714 | 0.563 | 0.709 | 0.586 |
| Total | 0.615 | 0.612 | 0.639 | 0.664 |

The impact of the policy on individual welfare exhibits similar but generally more pronounced properties to the ones described above. A subsidy at the beginning of the working life is effective in augmenting welfare as it subsidises agents when they are least productive and more risk averse. Moreover, by benefiting the most disadvantaged workers more significantly, *policy 2* significantly reduces inequality in the distribution of lifecycle welfare.

5.4 Conclusions

This chapter develops an estimable GE model of savings, skills and human capital with labour supply and idiosyncratic uncertainty on an overlapping generations setting. The model is designed in the Heckman et al. (98a) fashion, extending that setting by incorporating labour supply, idiosyncratic uncertainty and a government with capacity to introduce labour market policies at the expenses of the tax payer. We use the structure of the model to estimate and calibrate its parameters, being then able to solve it for the steady state. Within a structural setup as the one developed, it is possible to evaluate the overall impact of a wide range of labour market policies, assessing the direct effects on participants, indirect effects on non-participants and GE effects.

Numerical simulations of the long-run impact of wage subsidies were performed. Following the NDYP design, the simulated policies were targeted at the young, unskilled and unemployed workers. Two experiments were ran using the same type of policy under different levels of generosity. The results are qualitatively identical but much less pronounced for the least generous policy. It is suggested that different groups are differently affected by such policies. One important group is formed by the individuals the programme is originally targeted at, namely the unemployed in the *baseline case*. As expected, more generous policies attract more among these, reduces the incidence of unemployment in future periods of their life and benefits most of them both in terms of lifetime wealth and welfare. However, treatment reduces the odds of enrolling into education among *participants otherwise unemployed* by introducing additional indirect costs on the investment. The group that generally benefits the least from the programme is composed by the *participants otherwise employed*. The selection process dictates that not only the original target group will participate as other types of individuals may find it desirable to stay unemployed for a while in order to become eligible. However, *participants otherwise employed* loose in accumulated human capital from participation and the large enrolment rates induces negative changes in the price for unskilled labour that also contribute to deteriorate their future perspectives. Among these, the chances of staying out of work in the future increase significantly and less than half benefit from the existence of the programme. The consequential drop in indirect costs of education, however, induces a sharp increase in the

likelihood of taking up education during later periods in life.

Two important results from the simulations concern the relative importance of the indirect effects and the adequacy of available control groups. Computed GE effects significantly differ from PE effects, and are actually its reverse when measuring the impact of the programme on the incidence of unemployment. Thus, ruling out indirect effects may be a very restrictive assumption as changes in prices are expected to impact on every individual in the economy. In fact, it is showed that non-participants are strongly affected by the sole existence of the programme, some times exhibiting reactions that resemble the magnitude of those of participants. Thus, sampling from some seemingly adequate control group to construct the required counterfactual when trying to measure the direct effect of treatment on the treated may lead to misleading results due to a fundamental violation of the *no indirect effects* assumption.

5.5 Appendix to chapter 5

5.5.1 Proof of lemma 7

Take the group of agents defined by (θ, a, s) and any individual belonging to such group at a time t . The chances that he/she is indifferent between any two labour market activity options when facing a given set of prices is

$$P \left[\begin{array}{l} (V^{d=1} = V^{i=1} \geq V^{d=i=0}) \quad \text{or} \quad (V^{d=1} = V^{d=i=0} \geq V^{i=1}) \\ \text{or} \quad (V^{i=1} = V^{d=i=0} \geq V^{d=1}) \end{array} \right] = \quad (5.11)$$

$$= \int_{\pi^R}^{\bar{\pi}} \int_{\xi^R(\pi)}^{\xi^R(\pi)} dF(\pi, \xi) + \int_{\xi^R}^{\bar{\xi}} \int_{\pi^R}^{\pi^R} dF(\pi, \xi) + \int_{\underline{\pi}}^{\pi^R} \int_{\xi^R}^{\xi^R} dF(\pi, \xi) = 0$$

where any dependencies on the state variables have been omitted for simplicity of notation. Indifference is only important when it happens at the two preferred options. Thus, the three terms in equation (5.11) stand for the odds of being indifferent between working and studying when staying at home is not a better option, working and staying at home when studying is not a better option, and studying and staying at home when working is not a better option, respectively. It makes it clear that the measure of the “indifference set” in this economy is zero, meaning that the individual’s labour market decision is uniquely determined with probability one.

Given the reservation rule characterisation of the labour market activity policy, the existence of a continuum of individuals of the (θ, a) type at each moment in time makes the sub-group defined by (θ, a, s) also composed by a continuum of individuals or otherwise empty. But then, the aggregate supply of human capital of type s by the (θ, a) -group is deterministic and well-defined and so is the aggregate total. ■

5.5.2 Proof of lemma 8

It is obviously bounded below by zero.

The maximum amount of human capital of type s an individual of type θ may ever supply can be bounded as follows,¹²

$$\bar{h}_s \leq \underline{h} (1 + \nu(\theta, s, \underline{h}) \bar{\pi})^{A-s}$$

which, at an aggregate level means that,

$$\bar{H}_s^S < \sum_{a=1}^A \int_{\theta \in \Theta} \bar{h}_s C(a, \theta) dG(\theta) < M_s$$

meaning that the aggregate supply of human capital is bounded from above as well. ■

5.5.3 Proof of lemma 9

The proof follows by induction. Lets take a change in W_s of magnitude δ (say $\delta > 0$, as a similar argument can be produced for the symmetric case).

Individuals aged A .

The maximum change in life-time income at this stage of the life-cycle is

$$\Delta = \delta \bar{h}_s$$

where \bar{h} stands for the maximum attainable level of human capital in life.

Take the conditional value function, $V_A^{d,i}$. It is continuous in k_A (see lemma 6) and the agent is indifferent between a change in income caused by the change in W_s and a change in $k_A/(1+R)$ of similar magnitude (at the most, Δ). Given that δ can be made as

¹²It is being assumed that $\nu(\theta, s, h)$ is decreasing with h , consistently with the results presented in the previous chapter.

small as desired, so can Δ , making $V_A^{d,i}$ continuous in W_s as well. But then, by an implicit function argument, it is shown that the reservation policies, π_A^R and ξ_A^R , are continuous in W_s , meaning that only an infinitesimal fraction of individuals aged A change their working decision in response to a infinitesimal change in W_s .

Individuals aged $a = 1, \dots, A - 1$

Suppose the reservation policies, π_α^R and ξ_α^R for $\alpha = a + 1, \dots, A$, are continuous in W_s . Thus, an infinitesimal change in W_s only affects infinitesimally the probability of future working/studying decisions as well as the expected future income due to changing working/studying paths. Potential future gains resulting from shifts in the working/studying paths can, therefore, be value in present terms as a function of the change in W_s , say $\gamma(\delta)$ such that $\lim_{\delta \rightarrow 0} \gamma(\delta) = 0$. Similarly to what has been done for the last period of life, define Δ_a to be

$$\Delta_a = \delta \sum_{\alpha=a}^A \bar{h}_{s\alpha} (1 + \nu(\theta, s, \underline{h}) \bar{\pi})$$

The agent is willing to accept an increase in k_a of size Δ_a in exchange for the increase δ in W_s . And since the conditional value function $V_a^{d,i}$ is continuous in k_a and Δ_a can be made arbitrarily close to 0 as δ goes to 0, $V_a^{d,i}$ is also continuous in W_s . The same implicit function argument used above proves that the reservation policies at age a , π_a^R and ξ_a^R , are continuous in W_s .

Altogether, these amount to an infinitesimal change in the total supply of human capital in response to an infinitesimal change in W_s , making H_s^S continuous in W_s . ■

5.5.4 Proof of lemma 10

Consider first the aggregate supply of human capital of type s as a function of W_s . Lets take an increase in W_s of size $\Delta > 0$. Any worker that decides to work at W_s will also take the same decision at $W_s + \Delta$. This is obvious if $h\pi(W_s + \Delta) \geq B$ since in this case there is no risk from taking the working option. If $h\pi(W_s + \Delta) < B$ it suffices to notice that the agent is now incurring in a lower cost from working and getting higher future payoffs, therefore taking a lower risk from working.

As a function of other wage rates, W_j , $j \neq s$, the aggregate supply of human capital H_s^S may not be monotonic. Take the example of $H_2^S(W_3)$. An increase in W_3 will make

more people invest in education, acquiring level 2 and 3 skills: more people is going from the second to the third level of education and more people is going from the first to the second level of education. Whether the proportion of people acquiring only a medium level of education increases or decreases depends on the composition of the group of agents deciding to take any education at all and, therefore, on the level of the wages. For a relatively high skill premium, the characteristics of those acquiring education is likely to become more heterogeneous, and the proportion of agents ending up with medium education may increase. A lower skill premium implies that individuals taking education are more homogeneous, making a drop in the proportion of medium-educated people more likely. ■

5.5.5 Proof of lemma 11

The steady state equilibrium is characterised by the set of prices $(W_1, W_2, W_3, R, B, \tau)$ such that the market clearing conditions (5.5), (5.6) and (5.9) are fulfilled.

The proof of the existence of a steady state equilibrium follows in two steps. We start by taking the interest rate and the Government prices, R , B and τ , as given to look for a steady state as a function of the wage rates alone, and then solve for the government budget constraint. R continues to be given as no GE in the financial market is being assumed.

Under the specification (5.1) for the production function, the aggregate demand of labour of type s is a continuous and strictly decreasing function of W_s , converging to $+\infty$ and 0 as W_s goes to 0 and $+\infty$, respectively. But since the supply of human capital of type s is continuous (lemma 9) in W_s and bounded above 0 and below $+\infty$ (lemma 8), the two curves must always cross for given prices for the other skills and for R , B and τ . But then one can always choose $(W_1, W_2, W_3 \mid R, B, \tau)$ that simultaneously clear the three markets for human capital. Conditional on (R, B, τ) there is only one such combination as the supply of human capital of type s is non-decreasing in W_s (lemma 10). Thus, there is always a solution for the steady state problem as a function of the wage rates and this is unique and away from zero.

To solve for the unconditional steady state one must balance the government budget, which depends on the prices B and τ and on the individuals working decisions. An increase

in the tax rate τ affects individual's earnings making them more willing to stay at home, therefore rising Government expenses. Whether Government income decreases depends only on how much more revenue the rise in τ generates and, as usual, total revenues are likely to increase for relatively low values of the tax rate but eventually start to decrease as the tax rate approaches 1. On the other hand, a drop in the unemployment benefit makes staying at home less attractive, therefore reducing the Government debt. Whether there is a solution to the problem depends on whether the Government debt may be pushed to zero, meaning that B and τ need to be chosen so that the problem is possible. ■

Chapter 6

Further developments

There are a number of possible directions still unexplored in what concerns to the evaluation of the New Deal for the Young People (NDYP) and the related subjects addressed in this thesis. The research gaps are revealed not only within each study but also by putting together the reduced-form and the structural approaches, which may help disclosing the relative weaknesses and strengths of each one. Moreover, the quality of the evaluation results depends decisively on the existence of good, detailed data whatever the chosen approach and methodology used. Performing evaluation of social policies exposes the main limitations of the data at use and helps defining the main information requirements for future studies. In what follows, we briefly discuss some of the issues requiring further attention.

First, the main omission in our *ex-post* evaluation work is that we do not consider the longer-term effects of the NDYP. A full evaluation needs to consider whether individuals' employability is enhanced by their full treatment experience, including both the job-search assistance and the options of subsidised work, education, training, etc. A longer period is required for the evaluation process to assess the extent to which individuals become independent of the welfare and to have further insight on the potential importance of indirect effects of the NDYP. Data on earnings together with information on labour market history, participation status and the nature of the received treatment would be most valuable for this type of evaluation. It would inform the researcher about the quality of the job matching after treatment and the returns to the investment, be it on job-search

assistance alone or on other treatments like education and job experience.

Second and similarly to the first point, the long run *ex-ante* evaluation should also be extended to include the several types of treatment the NDYP include. Such analysis would provide important information about the way different treatments available simultaneously affect individual's behaviour and well-being. The interaction of several treatments available at the same time may significantly change the impact identified for one of the treatments when only that one is modelled. For example, let's consider the case of a tuition subsidy operating contemporaneously to the studied wage subsidy. The downward pressure on the unskilled wages and the upward pressure on the skilled wages resulting from the introduction of the wage subsidy is likely to be relaxed as more individuals may find studying a better option. Hence, the roll of who loses and wins from the existence of the programme and from participation is expected to be affected.

The third point concerns the incompatibility of the results on the importance of the indirect effects obtained from the *ex-ante* and *ex-post* approaches. There are a number of reasons for such outcome to be possible, the first being that the *ex-post* analysis is not suited to identify such effects. At the most, such analysis can only provide some clues about their importance, but not in a definitive way. Another potentially relevant factor relates to the nature of the treatments being studied in each case. A wage subsidy may eventually be more successful in moving individuals into work than job-search assistance, affecting the composition of the different labour market groups more effectively and exerting larger pressure on the relative wages. The first and second points addressed above bring together the two approaches in terms of the nature of the policy being evaluated, simplifying the comparison. Notwithstanding, the horizons considered in the two analysis do not match, and further work is also required to perform short-run evaluation within the structural approach.

Fourth, the structural setup developed in chapter 5 is ideal to assess the individual's responses to risk in what concerns to labour market decisions. There is a surprising lack of information about how risk affects individuals education and working choices and how policies may be designed to help insuring the most disadvantaged workers against such risk. An important question is how to disentangle the contribution of the individuals attitudes towards risk on the measured impacts from policies from other aspects of the

same impacts.

Fifth, the estimation of the different components of the general equilibrium model could be significantly improved with richer data. On the individual's problem addressed in chapter 4, the main drawback concerns the frequency of the NCDS58 observations. A ten years' period is overly long, raising concerns on the quality of the estimates. The sensitivity checks performed would be strengthened with further estimates based on more frequent data. On the production side, an integrated dataset including individual workers' and firms' information for a sensible period of time would make it possible to estimate the production function and discuss its specification in much more detail.

Finally, further inspection on the validity of the general equilibrium model is needed. Validation requires that the model predicts empirically observed facts outside the range of what it has been estimated on. That is, data required for validation should be orthogonal to that used for estimation. This is not an easy task as the estimation of a structural model as the one presented here uses a very large amount of data on the different areas touched by the model. Other economic issues are simply not modelled, and therefore not predictable.

Bibliography

- [1] Abadie, Alberto, Joshua Angrist and Guido Imbens (1998), “Instrumental Variables Estimation of Quantile Treatment Effects”, NBER working paper 229, forthcoming in *Econometrica*.
- [2] Altug, Sumru and Robert Miller (1998), “The Effect of Work Experience on Female Wages and Labour Supply”, *Review of Economic Studies*, 65, 45-85.
- [3] Anderton, Bob, Rebecca Riley and Gary Young (1999), “The New Deal for Young People: Early Findings from the Pathfinder Areas”, *Employment Service Research and Development Report*, ESR34, December.
- [4] Ashenfelter, Orley (1978), “Estimating the Effect of Training Programs on Earnings”, *Review of Economics and Statistics*, 60(1), 47-57.
- [5] Ashenfelter, Orley and David Card (1985), “Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs”, *Review of Economics and Statistics*, 67(3), 648-660.
- [6] Athey, Susan and Guido Imbens (2002), “Identification and Inference in Nonlinear Difference-in-Differences Models”, mimeo, NBER.
- [7] Attanasio, Orazio (1999), “Consumption”, in Taylor and Woodford (eds.), *Handbook of Macroeconomics*, volume 1, Amsterdam: Elsevier Science.
- [8] Attanasio, Orazio, James Banks, Costas Meghir and Guglielmo Weber (1999), “Humps and Bumps in Lifetime Consumption”, *Journal of Business and Economic Statistics*, 17(1), 202-35.

- [9] Attanasio, Orazio and Martin Browning (1995), "Consumption over the Life Cycle and over the Business Cycle", *American Economic Review*, 85(5), 1118-37.
- [10] Auerbach, Alan and Laurence Kotlikoff (1987), *Dynamic Fiscal Policy*, Cambridge: Cambridge University Press.
- [11] Banks, James, Richard Blundell and Agar Bruggavini (1999), "Risk Pooling, Precautionary Saving and Consumption Growth", IFS working paper W99/19.
- [12] Barsky, Robert, Thomas Juster, Miles Kimball and Mathew Shapiro (1997), "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey", *Quarterly Journal of Economics*, 112(2): 537-80.
- [13] Bassi, Laurie (1983), "The Effect of CETA on the Post-Program Earnings of Participants", *Journal of Human Resources*, 18(Fall), 539-556.
- [14] Bassi, Laurie (1984), "Estimating the Effects of Training Programs with Nonrandom Selection", *Review of Economics and Statistics*, 66(1), 36-43.
- [15] Bell, Brian, Richard Blundell and John Van Reenen (1999), "Getting the Unemployed Back to Work: An Evaluation of the New Deal Proposals", *International Tax and Public Finance*, 6, 339-360.
- [16] Blundell, Richard and Steve Bond (2000), "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Econometric Reviews*, 19(3), 321-40.
- [17] Blundell, Richard, Steve Bond and Frank Windmeijer (2000), "Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator", *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, 15, 53-91.
- [18] Blundell, Richard and Monica Costa Dias (2000), "Evaluation Methods for Non-Experimental Data", *Fiscal Studies*, 21(4), 427-68.
- [19] Blundell, Richard and Monica Costa Dias (2001), "Alternative Approaches to Evaluation in Empirical Microeconomics", forthcoming in *Portuguese Economic Journal*.

- [20] Blundell, Richard, Monica Costa Dias, Costas Meghir and John Van Reenen (2001), "Evaluating the Employment Impact of a Mandatory Job Search Assistance Program", IFS working paper W01/20.
- [21] Blundell, Richard, Lorraine Dearden, Alissa Goodman and Howard Reed (1997), Higher Education, Employment and Earnings in Britain, IFS Monograph Series.
- [22] Blundell, Richard, Lorraine Dearden and Costas Meghir (1996), The Determinants and Effects of Work-Related Training in Britain, London: Institute for Fiscal Studies.
- [23] Blundell, Richard, Alan Duncan and Costas Meghir (1998), "Estimating Labor Supply Responses Using Tax Reforms", *Econometrica*, 66(4), 827-861.
- [24] Blundell, Richard and Thomas MaCurdy (1999), "Labor Supply: A Review of Alternative Approaches", *in* Ashenfelter and Card (eds), *Handbook of Labor Economics*, volume 3, Amsterdam: Elsevier Science.
- [25] Browning, Martin, Lars Hansen and James Heckman (1999), "Micro Data and General Equilibrium Models", *in* Taylor and Woodford (eds.), *Handbook of Macroeconomics*, volume 1, Amsterdam: Elsevier Science.
- [26] Burtless, Gary (1995), "The Case for Randomised Field Trials in Economic and Policy Research", *Journal of Economic Perspectives*, 9(2), 63-84.
- [27] Card, David and Thomas Lemieux (2000), "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis", NBER working paper w7655, April.
- [28] Card, David and Philip Robins (1998), "Do Financial Incentives Encourage Welfare Recipients To Work?", *Research in Labor Economics*, 17, 1-56.
- [29] Carroll, Christopher (1994), "How Does Future Income Affect Current Consumption?", *Quarterly Journal of Economics*, CIX(1), 111-47.
- [30] Carroll, Christopher and Andrew Samwick (1997), "The Nature of Precautionary Wealth", *Journal of Monetary Economics*, 40(1), 41-71

- [31] Cochran, William and Donald Rubin (1973), "Controlling Bias in Observational Studies", *Sankhya*, 35(4), 417-446.
- [32] Costa Dias, Monica (2002), "Evaluating the Overall Effects of Active Labour Market Policies", mimeo, IFS.
- [33] Cossa, Ricardo, James Heckman and Lance Lochner (1999), "Wage Subsidies and Skill Formation: A Study of the Earned Income Tax Credit", mimeo, University of Chicago.
- [34] Davis, Steven, John Haltiwanger and Scott Schuh (1996), *Job Creation and Destruction*, Cambridge: MIT Press.
- [35] Dearden, Lorraine, Carl Emmerson, Christine Frayne, Alissa Goodman, Hidehiko Ichimura and Costas Meghir (2001), "Evaluating the Education Maintenance Allowance in the UK", mimeo, IFS, invited presentation Royal Economic Society Conference 2001.
- [36] Devine, Theresa and James Heckman (1996), "The Structure and Consequences of Eligibility Rules for a Social Program", *in* Polachek (ed.), *Research in Labor Economics*, volume 15, Greenwich, CT: JAI Press, 111-170.
- [37] Dickens, Richard, Steve Machin and Alan Manning (1999) "The Effects of Minimum Wages on Employment: Theory and Evidence From Britain", *Journal of Labor Economics*, 17(1), 1-22.
- [38] Dolton, Peter, Gerry Makepeace and John Treble (1992), "Public and Private Sector Training of Young People in Britain", *in* Lynch (ed.), *Training and the Private Sector: International Comparisons*, Chicago: University of Chicago Press, 261-82.
- [39] Eckstein, Zvi and Kenneth Wolpin (1989), "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience", *Review of Economic Studies*, 56, 375-390.
- [40] Eissa, Nada and Jeffrey Liebman (1996), "Labor Supply Response to the Earned Income Tax Credit", *Quarterly Journal of Economics*, CXI, 605-637.

- [41] Fan, Jianqing (1992), "Design Adaptive Nonparametric Regression", *Journal of the American Statistical Association*, 87, 998-1004.
- [42] Fisher, Ronald (1951), *The Design of Experiments*, 6th edition, London: Oliver and Boyd.
- [43] Hahn, Jinyong (1998), "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects", *Econometrica*, 66(2), 315-331.
- [44] Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw (2001), "Identification and Estimation of Treatment Effects with Regression Discontinuity Design", *Econometrica*, 69(1), 201-9.
- [45] Hales, Jon, Debbie Collins, Chris Hasluck and Steve Woodland (2000) "New Deals for the Young People and Long Term Unemployed: Survey of Employers" Employment Service Report No. 58.
- [46] Hausman, Jerry, and David Wise (1985), *Social Experimentation*, NBER, Chicago: University of Chicago Press.
- [47] Heckman, James (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", *Econometrica*, 46(6), 931-59.
- [48] Heckman, James (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1), 153-61.
- [49] Heckman, James (1990), "Varieties of Selection Bias", *American Economic Review*, 80(2), 313-318.
- [50] Heckman, James (1992), "Randomization and Social Program", *in* Manski and Garfinkle (eds.), *Evaluating Welfare and Training Programs*, Cambridge, Mass: Harvard University Press.
- [51] Heckman, James (1996), "Randomization as an Instrumental Variable Estimator", *Review of Economics and Statistics*, 78(2), 336-341.

- [52] Heckman, James (1997), "Instrumental Variables: A Study of the Implicit Assumptions Underlying One Widely Used Estimator for Program Evaluations", *Journal of Human Resources*, 32(3), 441-462.
- [53] Heckman, James and V. Joseph Hotz (1989), "Choosing Among Alternatives Non-experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training", *Journal of the American Statistical Association*, 84(408), 862-874.
- [54] Heckman, James, Hidehiko Ichimura and Petra Todd (1997a), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economic Studies*, 64, 605-654.
- [55] Heckman, James, Hidehiko Ichimura and Petra Todd (1997b), "Characterising Selection Bias Using Experimental Data", *Econometrica*, 66(5), 1017-1098.
- [56] Heckman, James, Robert LaLonde and Jeffrey Smith (1998), "The Economics and Econometrics of Active Labor Market Programs", *in* Ashenfelter and Card (eds.), *Handbook of Labor Economics*, Volume III, Amsterdam: Elsevier Science.
- [57] Heckman, James, Lance Lochner and Christopher Taber (1998a), "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics*, 1, 1-58.
- [58] Heckman, James, Lance Lochner and Christopher Taber (1998b), "General Equilibrium Treatment Effects: A Study of Tuition Policy", *American Economic Review*, 88(2), 381-86.
- [59] Heckman, James, Lance Lochner and Christopher Taber (1998c), "Tax Policy and Human Capital Formation", *American Economic Review*, 88(2), 293-7.
- [60] Heckman, James, Lance Lochner and Christopher Taber (1999), "General Equilibrium Cost-Benefit Analysis of Education and Tax Policies", *in* Ranis and Raut (eds), *Trade, Growth and Development*, Amsterdam: Elsevier Science.

- [61] Heckman, James and Richard Robb (1985), "Alternative Methods for Evaluating the Impact of Interventions", *in* Longitudinal Analysis of Labour Market Data, New York: Wiley.
- [62] Heckman, James and Richard Robb (1986), "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes", *in* Howard Wainer (ed.), Drawing Inferences from Self-Selected Samples, Berlin: Springer Verlag.
- [63] Heckman, James and Guilherme Sedlacek (1985), "Heterogeneity, Aggregate and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market", *Journal of Political Economy*, 93(6), 1077-1125.
- [64] Heckman, James and Guilherme Sedlacek (1990), "Self-Selection and the Distribution of Hourly Wages", *Journal of Labor Economics*, 8(1), S329-63.
- [65] Heckman, James and Jeffrey Smith (1994), "Ashenfelter's Dip and the Determinants of Program Participation", mimeo, University of Chicago.
- [66] Heckman, James and Jeffrey Smith (1998), "The Sensitivity of Non-Experimental Evaluation Estimators: A Simulation Study", mimeo, University of Chicago.
- [67] Heckman, James and Jeffrey Smith (1999), "The Pre-Program Earnings Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluation Strategies", *Economic Journal*, 109(457), 313-48.
- [68] Heckman, James, Jeffrey Smith and Nancy Clements, (1997), "Making the Most out of Program Evaluations and Social Experiments: Accounting for Heterogeneity in program Impacts", *Review of Economic Studies*, 64, 487-536.
- [69] Heckman, James and Ed Vytlačil (1999), "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects", *Proceedings of the National Academy of Sciences*, Vol. 96, No. 8, pp 4730-4734.
- [70] Hubbard, Glenn, Jonathan Skinner and Mathew Zeldes (1994), "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving", *Carnegie Rochester Series on Public Policy*, 40(June), 59-125.

- [71] Huggett, Mark (1996), "Wealth Distribution in Life-Cycle Economies", *Journal of Monetary Economics*, 38(3): 469-94.
- [72] Huggett, Mark (1997), "The One-Sector Growth Model with Idiosyncratic Shocks: Steady-State and Dynamics", *Journal of Monetary Economics*, 39(3), 385-403.
- [73] Huggett, Mark and Gustavo Ventura (1999), "On the Distributional Effects of Social Security Reform", *Review of Economic Dynamics*, 2, 498-531.
- [74] Imbens, Guido and Joshua Angrist (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62(2), 467-75.
- [75] Katz, Lawrence (1998), "Wage Subsidies for the Disadvantaged", *in* Freeman and Gottschalk (eds.), *Generating Jobs*, New York, NY: Russell Sage Foundation, pp. 21-53.
- [76] Kemple, James, Fred Doolittle and John Wallace (1993), "The National JTPA Study: Site Characteristics and participation patterns", New York: Manpower Demonstration Research Corporation.
- [77] Kimball, Miles (1990), "Precautionary Savings in the Small and in the Large", *Econometrica*, 58(1): 53-73.
- [78] Knab, Jean, Johannes Bos, Daniel Friedlander and Joanna Weissman (2000) "Do Mandates Work? The Effects to Enter a Welfare to Work Program" Manpower Development Research Corporation, November.
- [79] Krusell, Per, Lee Ohanian, Jose-Victor Rios-Rull and Giovanni Violante, G. (2000), "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis", *Econometrica*, 68, 1029-54.
- [80] Lalonde, Robert (1986), "Evaluating the Econometric Evaluations of Training Programs with Experimental Data", *American Economic Review*, 76(4), 604-620.
- [81] Layard, Richard (2000) "Welfare to Work and the New Deal" *The Business Economist*, 31(3), 28-40.

- [82] Lee, Donghoon (2001) "An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice", mimeo, University of Pennsylvania.
- [83] Low, Hamish (1999), "Self-Insurance and Unemployment Benefit in a Life-Cycle Model of Labour Supply and Savings", IFS working paper 99/24.
- [84] Magnac, Thierry and Jean-Marc Robin (1991), "A Dynamic Model of Choice Between Wage-Work and Self-Employment with Liquidity Constraints", mimeo, INRA.
- [85] Meghir, Costas and Marten Palme (2001) "The Effects of a Social Experiment in Education", IFS working paper 01/11
- [86] Meghir, Costas and Luigi Pistaferri (2001), "Income Variance Dynamics and Heterogeneity", IFS working paper WP01/07.
- [87] Meghir, Costas and Edward Whitehouse (1996), "The Evolution of Wages in the United Kingdom: Evidence from Micro Data", *Journal of Labor Economics*, 14(1), 1-25.
- [88] Meyer, Bruce (1995) "Lessons from US Unemployment Insurance Experiments", *Journal of Economic Literature*, 33(1), 91-131.
- [89] Mortensen, Dale and Christopher Pissarides (2002), "Taxes Subsidies and Equilibrium Labor Market Outcomes", Centre for Economic Performance Discussion Paper, 519, March.
- [90] Orr, Larry, Howard Bloom, Stephen Bell, Winston Lin, George Cave and Fred Doolittle (1994), "The National JTPA Study: Impacts, Benefits and Costs of Title II-A", A Report to the US Department of Labor, 132, Abt Associates: Bethesda, Maryland.
- [91] Pencavel, John (1986), "Labor Supply of Men: A Survey", *in* Ashenfelter and Layard (eds), *Handbook of Labor Economics*, volume 1, Amsterdam: Elsevier Science.
- [92] Richardson, James (1997a), "Can Active Labour Market Policy Work? Some Theoretical Considerations", Centre for Economic Performance Discussion Paper 331.

- [93] Richardson, James (1997b), "Wage Subsidies for the Long-Term Unemployed: A Search Theoretic Analysis", Centre for Economic Performance Discussion Paper 347.
- [94] Rosenbaum, Paul and Donald Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70(1), 41-55.
- [95] Rosenbaum, Paul and Donald Rubin (1984), "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score", *Journal of the American Statistical Association*, 79, 516-524.
- [96] Rosenbaum, Paul and Donald Rubin (1985), "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score", *American Statistician*, 39(1), 33-38.
- [97] Rubin, Donald (1978), "Bayesian Inference for Causal Effects: The Role of Randomization", *Annals of Statistics*, 6(1), 34-58.
- [98] Rubin, Donald (1979), "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies", *Journal of the American Statistical Association*, 74, 318-329.
- [99] Sattinger, Michael (1993), "Assignment Models of the Distribution of Earnings", *Journal of Economic Literature*, 31(2), 831-80.
- [100] Sianesi, Barbara (2001), "An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s ", IFS working paper W02/01
- [101] Sianesi, Barbara (2002), "Swedish Active Labour Market Programmes in the 1990s: Overall Effectiveness and Differential Performance", *Swedish Economic Policy Review*, 8(2).
- [102] Smith, Jeffrey (1997), "Measuring Earnings Dynamics Among the Poor: Evidence from Two Samples of JTPA Eligibles", mimeo, University of Maryland.
- [103] Smith, Jeffrey and Petra Todd (2000), "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?", mimeo, University of Western Ontario.

- [104] Stokey, Nancy and Robert Lucas (1989), *Recursive Methods in Economic Dynamics*, Cambridge: Harvard University Press.
- [105] Taber, Christopher (2000), "Tax Reform and Human Capital Accumulation: Evidence from an Empirical General Equilibrium Model of Skill Formation", mimeo, Northwestern University.
- [106] Topel, Robert (1986), "Job Mobility, Search and Earnings Growth: A Reinterpretation of Human Capital Earnings", *in* Ehrenberg (ed.), *Research in Labor Economics*, volume 8, part A, Greenwich, Conn.: JAI Press.
- [107] Van Reenen, John (2001), "Active Labor Market Policies and the British New Deal for Unemployed Youth in Context" *in* Blundell, Card and Freeman (eds.), *Seeking a Premier League Economy*, Chicago: University of Chicago Press.