The Effect of Youth Training on Labour Market State Transition Processes, Wage Expectations and Search Elasticities: A Matching Approach

JONATHAN ROBERT MOUNSEY

Submitted for the degree of Doctor of Philosophy

Department of Economics & Department of Statistics

University of Newcastle

Newcastle upon Tyne, NE1 7RU

England.

June 13, 2002

Abstract

During the 1980's, unemployed school leavers were encouraged to undertake a period of government sponsored training as a potential stepping stone into work. One such programme was the Youth Training Scheme (YTS), which operated from 1981 to 1994. The effectiveness of such schemes has long been questioned. This work presents an investigation into YTS scheme effectiveness in which we compare estimated labour market elasticities of people with YTS experience against those without. We acknowledge the existence of inherent differences between the two groups which could act to bias our conclusions, then attempt to account for these using matching methods.

In an experiment designed to assess a treatment effect, "independent" individuals who do not receive the treatment are usually used as a control group. Econometricians seldom have such experimental data due to the interference of self selection. One way to allow for the bias introduced by self selection is to use a matching algorithm. The purpose of the matching algorithms we employ is to produce a synthetic control group of individuals who have not experienced training but who are similar to the YTS participants. Having produced various matched datasets, we recalculate the elasticities of an optimal job search model to ascertain whether there is still evidence of a treatment effect when comparing like with like. In contrast to previous research, we also make allowance for the changing nature (heterogeneity) of the YTS scheme over time.

We compare different matching methods and assess their relative performance. Observations lying beyond the region of overlapping support are shown to cause a degradation in nearest neighbour matching performance. Kernel based procedures are employed using the full range of the bandwidth parameter on which they rely.

In the final part of this work we widen our field of vision to include the unemployment to work transition process. We present results which support the hypothesis of different YTS treatment effects in each generation of YTS.

Acknowledgements

I would like to thank my two supervisors, Dr. Richard Boys and Prof. Peter Dolton, for all their help and guidance throughout the course of this work. I would also like to thank the members of staff and postgraduate students from both the Departments of Economics and Statistics who have aided me throughout my time here. Your advice and friendship is greatly appreciated.

I am grateful to the Engineering and Physical Sciences Research Council without whose financial support I could not have undertaken this research.

To all my family, Mum, Dad, Paul, Kate, Nana and Grandad thank you.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Up until the 1970s, UK employers made provision for training the many of their young employees via the apprenticeship scheme. During this decade, the numbers of young people undertaking a traditional apprenticeship began to fall as employment in manufacturing declined. At the same time UK participation rates in full-time further education lagged well behind those of similarly developed nations. Governments acknowledged that an increasingly skilled work force was needed, to compete on international markets. The skill shortage, compounded by rising youth unemployment, could not be met without the introduction of a systematic vocational education and training provision for the majority of school leavers.

During the first term of Mrs Thatcher's Conservative Government plans were brought forward to replace the active labour market policy known as the Youth Opportunities Program (YOP) with a more ambitious scheme. In consequence, 1981 saw the announcement of a new one year Youth Training Scheme (YTS). By 1983 the scheme was fully operational. Since their introduction supporters and opponents of active labour market policies (ALMP's) have argued about the effects of such schemes on those who participate.

This research attempts to address some of the difficulties which lie behind a study of this question. At the same time some of the empirical methods which have been forwarded as possible solutions to them are compared.

1.1 The Youth Training Scheme

Youth unemployment was a significant problem for the UK Government throughout the 1970s and 80s and the problem persists to the present day. The problem has prompted direct active labour market policies designed to provide state funded training in an attempt to alleviate the difficulties experienced by young people facing the transition from school to work. In this thesis we examine the effect of state training schemes in changing the labour market expectations of young people. We suggest that the different forms of the government schemes, which have been implemented over the years, have had quite different effects.

In the late 1970s and early 1980s the UK Government introduced the Youth Opportunities Program (YOP) in an attempt to address what was generally considered to be a temporary excess supply of youth labour. It was suggested that school leavers were receiving high relative wages and as a result were being priced out of the jobs market. It was hoped that the low training allowances, which YOP offered, would act to depress youth wage expectations and thereby increase their chances of finding employment post scheme. The view that high youth unemployment was a temporary phenomenon led to an emphasis on work experience, whilst the training component of YOP was neglected, see Bradley (1995). In the early 80's, with the economy continuing to under perform and the job prospects for the young in decline, the number of persons on the scheme grew rapidly. At the same time the fraction of those who successfully made the transition to a permanent job after the scheme dropped to around a quarter, this led to widespread condemnation of YOP by participants, parents and trade unions. In 1981 the New Training Initiative addressed the perceived shortcomings of YOP with its proposals for the Youth Training Scheme. Under the first version of this scheme (YTSI), all 16 and 17 year olds were given the opportunity to either continue on to further education or proceed to a YTS placement for a period of planned work experience and three months compulsory off the job training at a college of further education. The young were encouraged to learn general transferable skills that employers might find of use.

In the mid 1980s there was recognition of a continued skills shortfall in the labour force, which was viewed as a hindrance to the Government's attempts to produce stable economic growth. This led to YTSI being superseded in April 1986 by a new version of the Youth Training Scheme (YTSII). The funding structure of which encouraged employers to recruit trainees as first year apprentices with employed status and allowances topped up to levels closer to the average for young workers. The scheme's duration was lengthened to two years for 16 year olds. There was increased emphasis on the scheme's vocational training elements (rising to 20 weeks) with the opportunity to gain qualifications. The introduction of YTSII coincided with a fall in the numbers of young people leaving school and an economic boom; these changing conditions would have enhanced the scheme's prospects of being labelled a success.

The late 1980s and early 1990s saw the Government increasingly concerned by the skills gap between the UK work force and that of it's international competitors. A need for yet higher levels of meaningful training amongst the young was identified. In May 1990 Youth Training (YT) replaced YTSII. A main objective of the scheme was to help school leavers attain qualifications of at least NVQ level 2 (a vocational qualification). There was an emphasis on scheme flexibility and with this in mind the length of YT placements and the amount of off-the-job training was tailored to the needs of each placement. YT saw the inception of the Training and Enterprise Councils (TECs) whose aim was to stimulate employer led investment in training[1]. As YT matured the economy experienced a cyclical down turn, which affected the prospects for school leavers during this period, and so the performance of YT may have suffered. Detailed surveys of the historical background of these schemes can be found in a number of studies Chapman and Tooze (1987), Dolton *et al.* (1993), Deakin (1996).

[1]This was not the first time that this had been attempted. ITB's tried to do something similar in 1964.

1.2 Traditional Approaches

Government sponsored training schemes have been introduced in a variety of countries with the intention of improving the labour market prospects of young people. Four separate labour market effects of training schemes have been identified in recent research. These include the *'wage effect'* Ashenfelter (1978), Ashenfelter and Card (1985), and the *'employment effect'* Card and Sullivan (1988) have been measured using data on individual workers. The *'deadweight loss'* and the *'substitution effect'* have been measured using data on firms. Studies in the USA have used both experimental and non-experimental data. UK studies have exclusively used non-experimental data since no alternative data is available. One potential economic effect of these schemes, which has been systematically ignored, is their impact on the wage expectations and search elasticities of the job seekers. As stated, one political view around the time of inception of YTS held that young people were pricing themselves out of jobs by holding unrealistic wage expectations. Hence, successive schemes have had the objective of lowering young people's reservation wages and make them more responsive, in their search behaviour, to changing labour market conditions. The major aim of this thesis is to test this hypothesis and examine the extent to which different forms of Youth Training Schemes may have had different effects.

If we consider that Government training participants receive a treatment effect as a result of their time on a scheme, then the size of this effect is of considerable importance. Most of the empirical work in the literature on evaluating active labour market training programs focuses on average effects and in particular the mean of the direct effect of treatment on the treated. This refers to the average effect of treatment to some outcome measure (such as post scheme labour market state). In contrast to this approach we measure treatment effects as the differences between the outcomes on a number of elasticities of unemployment for those who experience YTS and those who don't.

1.3 Recent Developments

Unfortunately differences are likely to exist between the YTS participants and those who did not experience the scheme. Theory would suggest that those who undertake a period of government training would, on average, have a lower skills base than the rest of the labour market (because the more highly skilled will already be in jobs or higher education). This skill shortage, as well as a number of other factors such as the placement of persons that careers officers considered needed the assistance which YTS might deliver, could act to bias any conclusions which we hope to draw from our analysis of the treatment effect. We address this problem using various methods of matching. Datasets are recast in such a way that members of the control group (non-participants) are deemed to be suitable matches for YTS participants if they have "similar" observed characteristics as measured by a propensity score computed using a probit model for YTS participation.

These methods, for which Rosenbaum and Rubin (1983) developed a theoretical framework of assumptions, have in recent times grown in popularity amongst empirical researchers wishing to correct for selection bias into a labour market state. The topicality of the matching approach has highlighted the need for rigorous assessment of the performance of these methods. Our analysis extends beyond simply applying these methods and encompasses a full appraisal of their relative strengths and weaknesses. We go still further in an attempt to uncover a procedure for identifying whether a given method when employed with a certain dataset will produce a "good quality" matched dataset. The matching methods studied include a nearest neighbour matching estimator similar to that used by Lechner (1999) as well as the kernel based methods described in Heckman *et al.* (1997).

1.4 Thesis Structure

Before describing the research proper, it is necessary to outline some preliminaries. The next Chapter facilitates this analysis by considering much of the background to our study.

It begins by outlining the various datasets, which we will use. Emphasis is placed on the power which panel data gives us when considering such things as duration within a specific labour market state. Section 2.1.1 describes the Youth Cohort Survey (YCS) cohort data from the 1984-94 period, which we use in our estimations. The data contains information on key variables such as unemployment duration, reservation wages and expected wages that enable us to provide estimates of the effect of all YTS variants over the period of survey. Later we examine the National Online Manpower Information System (NOMIS) dataset and discuss its value when linked to the YCS data. Section 2.2 introduces the optimal search model. The structural parameters derived from this model are the focus of our empirical estimations. We demonstrate the ways in which Lancaster and Chesher (1983) derive various elasticities from this model in section 2.3.

Section 2.4 outlines a method for modelling outcomes with and without training scheme experience. This discussion moves on to highlight the problems of self-selection into YTS. A brief demonstration of the power of randomised experiments and the ways in which they allow us to identify counterfactual outcomes is given. It is suggested that methods of matching might allow us to recast our dataset so that it more closely mirrors that which a randomised experiment would have produced. The underlying assumptions of traditional matching methods are introduced. The work of Rosenbaum and Rubin (1983) is shown to be crucial. They were able to show that, when attempting to match treatment and control individuals to one another, a propensity score for treatment (in this case a spell on YTS) may substitute for the group of variables used to model it. We then move on from matching in general to three specific examples of matching algorithms, the nearest neighbour, kernel and local linear matching methods. The penultimate section of Chapter 2 concludes by defining stochastic dominance and examining its use to assess the suitability of a given dataset for nearest neighbour matching. We conclude by considering the statistical procedure of bootstrap sampling.

The analysis within Chapter 3 attempts to assess the overall employment effects of various YTS incarnations. In turn we consider: models of the potential effect on the elasticities of unemployment following any spell on any form of YTS scheme; the effect when we allow for YTS scheme heterogeneity; the effect of any YTS scheme having nearest neighbour matched individuals with YTS experience to non-YTS persons in an attempt to create a synthetic control group of non-participants to account for the inherent differences between those people who participate in YTS training and those who don't; the effect of each form of YTS scheme having again used nearest neighbour matching to account for the inherent differences. Section 3.2 compares the estimates of the structural parameters obtained using the non-parametric approach developed by Lancaster and Chesher (1983) for both "all men" and "all women" [2] to those obtained in previous studies. Section 3.3 presents estimates of the search elasticities for men and women and by whether they experienced any form of YTS training. Subsection 3.3.1 makes allowance for the evolution of YTS over the sample period and estimates the structural parameters for each incarnation of the scheme. Section 3.5 contains an appraisal of the economic climate in which each incarnation of the YTS scheme operated. Knowledge of which allows us to interpret our findings. In Section 3.6 we consider our findings on the effectiveness of YTS in assisting young people to make the transition from school to work.

Chapter 4 begins with reference to the assumptions of matching methods, as defined in Chapter 2. We note that these methods may only produce "high quality" matches if these assumptions remain valid. We then propose to investigate the relative performance of the nearest neighbour method against the, kernel and local linear methods. We suggest that versions of the nearest neighbour algorithm with exclusions when matches are found to be poor are flawed. Elasticities as calculated using datasets generated with the kernel and local linear methods are presented. We identify the need to develop a method for selecting the bandwidth parameter, which these algorithms require. Our solution is to use a bootstrap iterative procedure to generate a value for the bandwidth. Results are then interpreted using synthetic control group elasticities and wage rates, generated using our

[2] Without regard to the differences between the three YTS types, those who participate in YTS and those who don't.

estimated bandwidths. We suggest, without the introduction of new bias, that a series of random exclusions from the treatment group could allow us to adjust for common support.

We move on to consider the YTS treatment effect for labour market state transition rates. Chapter 5 introduces the 3-state labour market model. We estimate a series of Cox proportional hazard models for the unemployment to work transition using unmatched and matched datasets of various types. The specification includes indicators for whether individuals experienced a spell on YTSI, YTSII or YT. Results complement those of previous chapters, revealing the implications of the various movements in reservation wages and the job search elasticities for those with YTS experience. Chapter 6 concludes the work by considering our attempts to answer the twin questions of matching algorithm performance and YTS performance. Conclusions are drawn and suggestions for future work are considered.

Chapter 2

Preliminaries

Before we present the results our research has uncovered, it is necessary to review some of the theory behind the analysis of active labour market policies, which we shall make use of throughout this work. Those wishing to draw a series of conclusions from what will follow must first have an understanding of the data. As such we begin this chapter by outlining the initial design of the Youth Cohort Survey and its evolution over the period we chose to investigate. Later we will introduce the search theory framework for exit from unemployment, which Lancaster and Chesher (1983) used to derive a series of elasticities. The values of these elasticities cast light on the job search process and which we rely on during the early part of our analyses.

Next we introduce the concept of matching and its use as a solution to government training scheme evaluation problems. Motivation for the procedure is given by way of a model for labour market state outcomes. Having demonstrated how the use of randomised training scheme exclusions can facilitate the evaluation of such schemes, we then outline the theory of matching methods. We then explain how they can allow us to uncover the effects of training schemes from panel data in the absence of a randomised control group.

We then introduce three matching methods of increasing complexity and propose to employ them in an analysis of the YCS datasets. Starting with nearest neighbour matching, we describe a complex variant of the group of such methods. Much as Lechner

(1999) did, we employ a variable width calliper whilst matching actual non-participants to persons with any form of government training scheme experience. Next we illustrate the kernel density and local linear regression methods of matching, which make use of all non-participants to create a single synthetic non-participant for each scheme participant. We move on to present a brief outline of the probit model specification, which led to our propensity scores for YTS participation. The ways in which the various matching methods make use of these score predictions when creating matched datasets are crucial to our interpretation of the results from any analyses of them.

The penultimate section of this Chapter deals with the concepts of first and second order stochastic dominance and their use when attempting to gauge the suitability of any dataset for the application of a matching method. We conclude the Chapter with a brief outline of the method and statistical basis behind bootstrap sampling.

2.1 Data

The main body of data used in this study consists of sweeps one to three of cohorts one to six of the Youth Cohort Study (YCS). We were also able to make use of extra exgonenous local labour market information from the NOMIS (National Online Manpower Information Service) database. The subsections, which follow, outline these two datasets and include some discussion of their sampling framework and final survey samples.

2.1.1 The YCS Dataset

The Youth Cohort Survey (YCS) is a Department for Education and Skills (DfES)[1] funded longitudinal programme of surveys and was originally designed to facilitate an exploration into the behaviour, decisions and routes of transition from school into a labour market state of a representative sample of young people aged 16 to 17 at the time of initial postal contact in England and Wales. The survey was intended to be the government's main

[1]Formally the Department for Education and Employment (DfEE).

source of information on the school to work transition process. It is a national survey of the experiences of 16 to 19 year olds on YTS, in and out of work, at school and further education, and in other forms of training. Cohorts 1-6 were re-surveyed on each of the first three years after individuals reached their year of minimum school leaving age.

The structure of the YCS is presented in Table 2.2. Only a random sample of those eligible for survey was selected for each new cohort. Not all persons who where selected to be surveyed were contacted or returned their questionnaires. The sample size was therefore affected by survey non-response. When interpreting the results of Table 2.2 numbers in parenthesis relate to the percentage of respondents to a given sweep as a proportion of those who responded to the previous sweep. The subsequent sweeps are hence a subset of the respondents to sweep 1. Final response rates are recorded in the last column. This was the sample design for cohorts 1-4. For cohorts 5 and 6 an attempt was made to contact the whole target sample of sweep 1 at both subsequent sweeps. For the purpose of this investigation information used during the course of our study from cohorts 5 and 6 has been made comparable to that of cohorts 1-4 since information obtained at sweeps subsequent to an individual's non-response are removed from the dataset. An example of this structure can be seen using cohort 3. With reference to the YCS design of Table 2.2, Cohort 3 (YCS III) began with a postal survey dated March 1987 of young people who completed their compulsory education during the school year 1985-6. Respondents were sampled again in March 1988 and March 1989.

Usually, the levels of response for postal surveys of this kind are low. However this is not the case for the YCS dataset. Those who oversaw the surveying procedure undertook a number of actions in an attempt to minimise the levels of non-response. As well as the questionnaire, persons also received four separate reminders as certain dates elapsed. The middle two reminders included replacement questionnaires and succeeded in capturing most late respondents, whilst the final one involved a telephone call (where a person's phone number was available) which only managed to increase the final response rate by 1% – 2%.

Since YCS cohorts 1-6 cover such a large passage of time it is not unsurprising to discover that the questionnaires used underwent a number of revisions. The resulting differences between cohort datasets mean that only a limited number of variables can be used when appending all six cohort datasets together. Differences in the sampling techniques between cohorts may also introduce bias into our analyses. Although the main sampling features on which the original cohort was founded remained in place over the course of the subsequent cohort samples, some variations were inevitable. Sample size is one area, which was allowed to develop within later cohort samples. Cohorts 1 and 2 sampled 10% of the individuals from a selected set of representative schools whereas later cohorts took a sample of 20%. The sample size variations should not in themselves reduce the representative nature of all cohorts to their respective populations; yet the larger samples of cohorts 3 to 6 may lead to smaller variations in the variables of interest which may in themselves then lead to the discovery of significant results within the analyses of such samples or a domination by these larger populations within YCS 1-6 when taken as a whole.

To summarise the YCS underwent two major changes over the period covered by cohorts 1-6. We saw that the sample size as a percentage of the population of interest increased from 10% to 20% after cohorts 1 and 2. Hence we would expect that those individuals within cohorts 1 and 2 would resemble each other most, as would those of cohorts 3, 4, 5 and 6. It will be interesting to see whether this prior belief about the structure of the data manifests itself in our analyses.

A publication from the Government Statistical Service (GSS) (1996) covers the data within YCS cohorts 1 to 6. Several of the findings in that work are relevant when interpreting the results, which we present. The GSS work suggests that YCS reflects a number of trends highlighted in official Government figures for areas such as educational attainment. Estimates during the period covered by cohorts 1 to 6 show a growth in the proportion of young people participating in education, training and achievement of qualifications. Asian youths were particularly likely to remain in full-time education for

Regressors	Description	Un-Mtchd Males Mean	Un-Mtchd Females Mean	Matched Males Mean	Matched Females Mean
Reservation Wage	Responses to the question: '*What is the lowest weekly take home pay you would consider for a full time job?*'	96.16908 (0.49596)	88.6961 (0.42353)	98.75319 (0.65870)	88.61614 (0.52228)
Expected Wage	Responses to the question: '*How much weekly take home pay do you expect to earn in your next job?*	126.1087 (0.66737)	115.8429 (0.54831)	126.8414 (0.90866)	114.2719 (0.69372)
Ethnic Origin	White: Respondent is white, Other: Respondent belongs to another ethnic group.	0.081128 (0.27306)	0.073842 (0.26154)	0.066863 (0.24985)	0.069749 (0.25477)
Education Score at 16	These are education scores calculated from a persons GCSE (or equivalent) results. The scores are computed as A=5 points, B=4 points, C=3 points, D=2 points and E=1 point.	8.084804 (9.72350)	8.861285 (9.66383)	4.265487 (5.51636)	5.339342 (6.40732)
Maths GCSE	1 Respondent held a GCSE (or equivalent) in Maths at grade C or above, 0 Otherwise.	0.215931 (0.41152)	0.188986 (0.39154)	0.086529 (0.28121)	0.107367 (0.30964)
English GCSE	1 Respondent held a GCSE (or equivalent) in English at grade C or above, 0 Otherwise.	0.256618 (0.43682)	0.325824 (0.46873)	0.139626 (0.34668)	0.206505 (0.40488)
No. of Siblings	The total number of siblings (not a dummy).	1.807843 (1.55990)	1.858156 (1.62853)	2.073255 (1.69402)	2.023119 (1.72236)
Career Service Interview	1 Respondent attended a career service interview, 0 Otherwise.	0.141667 (0.34875)	0.111181 (0.31439)	0.174041 (0.37924)	0.142633 (0.34977)
Live With Parents	Respondent lived with parent(s).	0.977696 (0.14769)	0.952232 (0.21330)	0.97296 (0.16224)	0.939655 (0.23817)
YTSI	Respondent had a spell on the first incarnation of YTS.	0.083333 (0.27642)	0.075511 (0.26424)	0.167158 (0.37321)	0.14185 (0.34896)
YTSII	Respondent had a spell on the second incarnation of YTS.	0.116667 (0.32106)	0.131206 (0.33766)	0.234022 (0.42349)	0.246473 (0.43104)
YT	Respondent had a spell on YT	0.053922 (0.22589)	0.06237 (0.24185)	0.09882 (0.29849)	0.111677 (0.31503)
Pred. Training allowance	An estimate of the respondents training allowance.	- (-)	- (-)	3.998719 (0.16567)	3.952304 (0.14224)
Regions	North, Yorkshire & Humberside, East Midlands, East Anglia, Greater London, South East, South West, West Midlands, North West and Wales. North of England is used as a reference group.	- (-)	- (-)	- (-)	- (-)
Cohorts	Dummies for Cohorts 1 to 6. Cohort 1 is used as a reference group.	- (-)	- (-)	- (-)	- (-)
Exogenous Variables					
Youth Unemployment	These variables were collected by region and come from Regional Trends. Wages are adult figures and are dis-aggregated by gender, so females in the database are assigned mean female wages for their home region in the relevant time period. Regional training place figures are normalised by using the population of 16-19 year olds in the region for each year. Population figures provided by the Office for National Statistics (ONS).	0.064805 (0.03561)	0.064578 (0.03562)	0.076662 (0.04027)	0.075117 (0.03982)
Regional YT Places		0.083314 (0.02429)	0.083677 (0.02434)	0.090544 (0.02228)	0.090273 (0.02213)
Average Wages		233.1733 (32.5365)	159.0188 (24.2911)	220.9955 (24.4659)	150.5093 (19.6396)
LEA Unemployment	Using the National On-line Manpower Information System (NOMIS), unemployment figures were collected for each month for each Local Education Authority (LEA). Rates were then calculated by dividing through by Census population figures of the number of individuals in the labour market for each LEA. The 1981 Census is used for the period to 1991 and the 1991 Census is used thereafter.	10.01333 (4.13270)	9.922007 (4.14132)	11.11412 (4.36587)	10.96832 (4.31522)
Sample Size		4115	4770	2034	2552

Table 2.1: Summary Statistics of the Four YCS Subsets Used in this Study

all ages of study and young people from all ethnic minority groups were more likely to be in full-time education than those who classified themselves as white. In 1994, as the final sweep of cohort 6 drew to a close, 87% of Asian 16 year olds and 66% of Asian 18 year olds were in full-time education. Compare these to figures of 72% of all 16 year olds and 40% of all 18 year olds. At the same time there were increases in the proportion in full-time education among 16 year olds from groups with traditionally low levels of educational continuity. These groups included those with lower GCSE attainment, parents in unskilled manual occupations and those who stated that they had truanted during year 11. The key activity at age 18 continued to be full-time education. In 1994, 40% of 18 year olds responded that this was their main labour market state. Young men were more likely to be in a government sponsored training scheme at all sweeps compared to women.

			Year in Which Sweep Was Conducted										
Cohort	Of School Leaving Age In:	Sample Size	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	Final Response Rate
1	1983-84	8064	1 (69%)	2 (75%)	3 (84%)								43%
2	1984-85	14430		1 (74%)	2 (80%)	3 (83%)							49%
3	1985-86	16208			1 (77%)	2 (76%)	3 (76%)						44%
4	1987-88	14116					1 (71%)	2 (74%)	3 (78%)				41%
5	1989-90	14511							1 (72%)	2 (75%)	3 (77%)		42%
6	1990-91	29922								1 (69%)	2 (74%)	3 (75%)	40%

Table 2.2: The Cohort Design Structure of the YCS Survey

As the non-response literature would suggest, see Taylor (2000), those with higher achievement, both educationally and with respect to their labour market outcomes are more likely to have responded to the survey. Hence these groups are over represented, whilst those in other activities are under-represented in sweeps 2 and 3. In an effort to counter this effect, those who administered the survey applied weightings. However the attrition present in sweeps 2 and 3 re-introduces the bias. As a result, data from sweep 1 is more reliable than that from sweep 2, which is in turn more reliable than that of sweep 3. Variables used in the weighting process were population variables for sex, school type, region, year 11 attainment and participation in full-time education at 16. The weighting methods differ for each cohort to reflect changes in sample selection.

Having weighted the initial sample, some bias still remained once the data was collected. This was the result of differing response rates among different groups of young people. At sweep 1, the data were weighted to education census figures for participation in full-time education. As an example of the need to weight the initial sample to bring it into line with population estimates, as well as the bias which remained post survey, Table 2.3 contains estimates of how the main labour market activity taken from YCS compares with administrative estimates of those in full-time education and on government supported training for cohorts 5 and 6. As this table demonstrates, YCS consistently over represents those in full-time education, whilst under predicting the proportion in Government supported training. It should be noted that YCS included a sub-sample from independent colleges of education, which the DfEE excludes from its figures.

The YCS is a clustered dataset, because it's school-based. This work does not consider the implications for the design effect which clustering can have. Throughout this work we do not estimate robust formulae for standard errors, nor do we attempt to identify any school effects.

Some questions, which we considered of high importance to our study, were only put to a subset of those sampled. For instance, questions regarding reservation/expected wages were only put to those who were unemployed at the time of survey, whilst training allowances were only available for those persons participating in a training scheme at the time of the survey. We follow other work in this area (such as the research presented in papers by Lynch (1983), Lancaster and Chesher (1983), Main and Shelley (1988), Jones (1988), Gorter and Gorter (1993), and Dolton and O'Neill (1995)) and use as a measure of the reservation wage denoted ξ, the individual's response to the question *'What is the lowest weekly take home pay you would consider for a full time job?'*. To measure the expected wage denoted x, we use responses to the question *'How much weekly take home pay do you expect to earn in your next job?*

Of course all the usual caveats apply to the use of this type of question to construct information of the kind we seek. The details of the derivation of these variables are provided in Table 2.1. The distributions of these variables are graphed in Figure C.3. It is reassuring that these graphs look reasonable and that the joint distribution of the two variables provided in Table 2.4 throws up only 5% of young people in the sample who are irrational, in the sense of having either $\xi > x$, or $\xi < b$ or $x < b$,. where b equals the unemployment benefit level. This proportion is very low for studies of this kind[2].

16 Year Olds	1992		1994	
	DfEE	YCS	DfEE	YCS
FT education (not on Government supported training)	65.1%	66.0%	71.4%	72.0%
Government supported training	15.7%	14.0%	13.6%	12.0%
16 Year Olds	**1992**		**1993**	
	DfEE	**YCS**	**DfEE**	**YCS**
FT education (not on Government supported training)	47.9%	52.0%	53.1%	57.0%
Government supported training	18.6%	16.0%	16.5%	15.0%
16 Year Olds	**1993**		**1994**	
	DfEE	**YCS**	**DfEE**	**YCS**
FT education (not on Government supported training)	32.8%	39.0%	36.6%	40.0%
Government supported training	7.4%	7.0%	7.5%	7.0%
		Source: YCS Sweep 1 of Cohorts 5 to 7		

Table 2.3: Comparison of YCS Results with DfEE Estimates

[2]This proportion is also consistent with people who did not understand the question or who made a mistake with their response.

2.1.2 Duration and Dated Variables

Much of the analyses, which we present towards the end of this work, are concerned with labour market state durations and transition processes. YCS provides two separate sources of information about dated events, which we use to define our duration and transition variables. In the first, respondents are asked to complete a diary for the year prior to each sweep (6 months for sweep 1) in which they recall their labour market status in each month. Although respondents could include a large number of labour markets states, we have chosen to aggregate the classification of labour market states into a single consistent classification over time: (i) unemployed; (ii) on YTS (or other government training scheme); (iii) in full time employment; (iv) in full time education; (v) other. YCS gives each individual's state for around 30 months over the period when they were between 16 and 19 years old. Taken together the first six cohorts in the YCS survey cover the 114 months from September 1984 to February 1994. Every person within our dataset has a fully completed labour market state diary. This data is fully described in Dolton *et al.* (1999). Short Fortran programs were written to derive variables such as total time spent unemployed during the survey period or individual labour market state spell length from this invaluable and rich dataset. The code used to extract unemployment durations for those who then exited to a full time job is reproduced in Appendix A.1. The second source of dated information comes from the responses to particular questions on the survey questionnaire at each sweep. Both sources of information were used to construct the variables in this study.

In order that we might maintain the sample size when estimating models for labour market state transition processes it was necessary to predict the reservation/expected wages and training allowances of those who did not report them before we could include them in the model. Descriptions of the specifications for these estimations are provided during the presentation of our empirical findings.

2.1.3 The NOMIS Dataset

Variables such as the unemployment rate by the Local Education Authority (LEA) within which individuals undertook their secondary education, youth unemployment rate by region[3] of individuals residence and available youth training places as a proportion of the population by region of residence were constructed from the National Online Manpower Information System (NOMIS) data resource. Summary statistics for all the NOMIS derived variables are included in Table 2.1.

LEA and youth unemployment measures were selected due to their relevance to our dataset, also youth labour force unemployment rates were not available at the national level. These measures were believed to improve on national unemployment rate figures in that the disaggregation of each by LEA and region would reflect the real state of employment prospects for individuals due to the wide variations in labour market conditions between such regions.

It was hoped that a measure of Youth Training Places might allow us to capture a greater part of the extent to which 16 and 17 year old school leavers were unemployed. This was of concern to us following the law change of September 1988 when such persons were no longer eligible for unemployment benefit (Job Seekers Allowance) and so disappeared from the unemployment register as defined as those actively seeking work. Youth Training Places might also capture the affluence of a region. Wealthier regions with more healthy employment prospects for the young can at the same time afford a higher level and quality of training scheme provision which in turn might increase the long-term employment prospects of scheme recipients.

[3]The Survey took as its population all persons of school leaving age in the relevant years within England and Wales. This was divided into regional areas as follows: North, Yorkshire and Humberside, East Midlands, East Anglia, Greater London, South East, South West, West Midlands, North West and Wales.

	Reservation Wage									
Expected Wage	Under 27.5	100	125	150	175	200	225	250	Over 250	Total
Under 27.5	3	10	2	1	0	0	0	0	0	16
100	21	3305	84	35	11	3	4	1	2	3466
125	2	2539	797	16	9	0	1	2	0	3366
150	0	583	832	210	2	4	1	0	1	1633
175	0	133	390	216	85	2	2	0	1	829
200	0	75	135	101	49	29	0	0	0	389
225	0	15	34	49	27	18	11	0	0	154
250	0	18	26	20	40	22	9	10	0	145
Over 250	0	7	16	9	13	12	8	17	18	100
Total	26	6685	2316	657	236	90	36	30	22	10098

Table 2.4: The Joint Distribution of Reservation and Expected Wages within the YCS Dataset

2.2 The Search Theory Framework for Exit from Unemployment

This section briefly outlines the standard model of optimal job search. Job offers arrive at random intervals. In particular it is assumed that risk neutral agents receive job offers according to a Poisson process with arrival rate λ. The length of an interval between any two successive job offers is therefore random. The implication of this is that the probability of obtaining an offer in any short interval of time is directly proportional to the length of the interval. These risk neutral agents seek to maximise the expected present value of their income, discounted to the present over an infinite horizon at rate ρ. Income maximisation is not crucial to the general implications of the model. However its use leads to simpler derivations than if we were to assume the more realistic situation of utility maximisation. The probability of receiving at least one job offer within a short interval of length t is $\lambda t + o(t)$, where $o(t)$ is the probability of receiving more than one offer in the interval and $o(t)/t \to 0$ as $t \to 0$.

On receiving a wage offer w, an individual has to decide whether to accept that offer and earn w forever, or reject the offer and search for a better offer next period. Once rejected an offer cannot be recalled. Under the assumption we have made a recall option would never be exercised, as with a constant reservation wage, an offer, which is unacceptable now will be unacceptable at any time in the future. What is important is that the value of occupying a job forever is an increasing function of the wage rate, w. Successive job offers which arrive during a period of unemployment are independent realisations from a known wage offer distribution with finite mean and variance, cumulative distribution function $F(w)$ and density $f(w)$. Each individual faces a wage offer distribution, which reflects his or her status in the labour market. Knowledge of this distribution allows them to follow an optimal search strategy.

While searching the individual may receive non-labour income in the form of a benefit level, b. Let V_u denote the value of searching during the next period, i.e the present value of future net income given that the optimal strategy will be pursued in the future and V_w the present value of stopping, accepting the offer and working forever at that wage.[4]

Since a person's net income flow during a spell of unemployment is constant, job offers are independent and identically distributed and assuming stationarity for both $F(w)$ and λ, it follows that V_u will be constant over the entire spell. We can characterise the problem.

$$V_u = \frac{1}{1+\rho t} bt + \frac{\lambda t}{1+\rho t} E_w[\max\{V_u, V_w\}] + (1-\lambda t)\frac{1}{1+\rho t} V_u + o(t) \tag{2.1}$$

[4] It is straightforward to allow for job loss, with probability say, s, in a given period in this model. However the inclusion of job loss in this fashion does not alter the basic results presented below. A fixed date retirement may also be included, but as the length of a working life is long and the persons within our dataset young, the practical implications of this change are negligible. (see Lynch (1983)).

The first term on the right side of the equation translates as the discounted present value of net income whilst unemployed over the short time interval t. The second represents the probability of receiving an offer in the interval t multiplied by the discounted expected value of following the optimal policy if a job offer of w is received, where V_w denotes the present value of accepting the offer. The third term is the probability of no offer in the time interval t multiplied by the discounted value of optimal job search ever after. The final term signifies the returns to job search when more than one wage offer is received. $o(t)$ is the probability of receiving more than one offer and K is the value of the optimal job search policy given that more than one offer is obtained. Assuming the Poisson arrival rate, then $o(t)/t \rightarrow 0$ as $t \rightarrow 0$. V_w is therefore the present value of expected lifetime income at wage rate w. This is characterised by the equation

$$V_w = \frac{w}{\rho}. \tag{2.2}$$

Given that V_w is continuous and strictly increasing with w and V_u does not depend on w, it follows that the optimal strategy is a stationary reservation wage policy. The individual accepts the offer if $w \geq \xi$, where the reservation wage ξ is a minimum acceptable wage offer as defined by equating the expected present value of employment and the expected present value of continued optimal search. Therefore

$$V_\xi = \frac{\xi}{\rho} = V_u. \tag{2.3}$$

If we substitute equations (2.2) and (2.3) for V_w and V_u in equation (2.1) we get

$$\frac{\xi}{\rho} = \frac{1}{1+\rho t}bt + \frac{\lambda t}{1+\rho t}E_w\left[\max\left\{\frac{w}{\rho}, \frac{\xi}{\rho}\right\}\right] + \frac{(1-\lambda t)}{1+\rho t}\frac{\xi}{\rho} + u(t). \tag{2.4}$$

For an individual faced with the above problem the optimal strategy is to choose a reservation wage ξ, such that the individual will accept the first wage offer greater than or equal to ξ. When rearranging terms and taking limits, equation (2.4) translates to

$$\xi = b + \frac{\lambda}{\rho}\int_{\xi}^{\infty}(w-\xi)dF(w). \tag{2.5}$$

The rate at which individuals accept job offers and exit is equal to the hazard

$$h = \lambda[1 - F(\xi)]. \tag{2.6}$$

From equation (2.6) we see that the probability of exiting unemployment depends on the rate at which job offers arrive and the probability that an individual will accept an offer, which in turn is a function of the reservation wage. It can be shown that under certain regularity conditions: a reduction in b, a reduction in ξ or an increase in λ will all lead to an increase in the probability of exiting from unemployment; see Van Den Berg (1994).

2.3 Derivation of Job Search Elasticities

Lancaster and Chesher (1983) have shown that by differentiating both the reservation wage given by (2.5) and the hazard function given in (2.6) with respect to b and λ one can obtain estimates of the responsiveness of the reservation wages and re-employment probability to changes in the offer arrival rate and the level of unemployment benefits. The advantage of this approach is that by combining the restrictions implied by the optimal job search model with information on reservation wages and expected wages we can deduce many of the structural parameters of the model without specifying the nature of the unknown wage offer distribution. The procedure also facilitates a comparison of the magnitudes of the estimates obtained in this fashion, with those obtained using more standard parametric procedures.

Using the restrictions embodied in reservation wage equation (2.5), Lancaster and Chesher (1983) show that the elasticities of the reservation wage with respect to the benefit level and the arrival rate of offers can be written as

$$\frac{d \log \xi}{d \log b} = \frac{b\,(x - \xi)}{\xi\,(x - b)} \tag{2.7}$$

and

$$\frac{d\log\xi}{d\log\lambda} = \frac{(\xi - b)}{\xi}\frac{(x - \xi)}{(x - b)} \tag{2.8}$$

where b, λ and ξ are as defined above and x is the expected wage given that this wage is accepted, that is

$$x = E[w|w \geq \xi].$$

Using these equations and information on b, x and ξ we can determine the extent to which the reservation wage is affected by changes in the benefit level and the offer arrival rate. In order to determine the impact of these variables on the hazard rate h (the probability of leaving unemployment given that a job has not yet been found) we need to specify a form for the wage offer distribution. Lancaster and Chesher adopt the Pareto Distribution, with density

$$f(w) = \begin{cases} \frac{\alpha w_0^\alpha}{w^{\alpha+1}}, & w \geq w_0 \\ 0, & \text{otherwise} \end{cases} \tag{2.9}$$

where w_0 and α are the lower bound and scale parameter of the distribution, respectively. The Pareto distribution is characterised by a long tail. If $\alpha > 1$, the expected value of a wage offer, w is $\alpha w_0/(\alpha - 1)$ if $\alpha > 2$ the variance is $\alpha w_0^2/\{(\alpha-1)^2(\alpha-2)\}$. Although this distribution is often used in analyses of this kind, economic theory does not present a compelling argument for its use. As such the choice of this wage offer distribution is a potential source of lack of fit or misinterpretation of results. Having made this assumption the elasticities can be written as

$$\frac{d\log h}{d\log b} = -\frac{b}{\xi}\frac{x}{(x - b)} \tag{2.10}$$

and

$$\frac{d\log h}{d\log\lambda} = \frac{d\log\xi}{d\log b} \tag{2.11}$$

To examine the robustness of our estimates to the Pareto assumption we also estimate these last two elasticities assuming that the wage offer distribution has an exponential rather than Pareto distribution. Under the exponential assumption it can be shown that:

$$\frac{d\log h}{d\log b} = -\frac{b}{(x-b)} \tag{2.12}$$

and

$$\frac{d\log h}{d\log \lambda} = \frac{(x-\xi)}{(x-b)}. \tag{2.13}$$

Given information on the individual's level of b, ξ and x we can calculate the above elasticities for each individual using that individual's value of b, ξ and x. The results we report represent the mean values of these elasticities across individuals.[5]

2.4 Matching: A Solution to the Government Training Scheme Evaluation Problem

Having outlined the theory behind the derivation of the job search elasticities we hope aim to produce estimates of these indicators for the datasets used throughout this work. Early work focuses on the YTS treatment effect for scheme participants. We suggest that those who experienced YTS were subject to a process of self-selection and as such our measurements of the effect of treatment on such individuals and their job search methods will be biased. Having acknowledged the existence of an effect, we will make use of several matching procedures to correct for such bias. This section outlines the nature of experimental datasets and details how matching can allow us to employ non-experimental datasets when asking experimental questions.

[5] We have also computed these elasticities by taking the mean values of b, ξ and x, for the sample first and then applying the formula. This is the method used by Lancaster and Chesher (1983). The qualitative results are not sensitive to this method of calculation.

The evaluation of Government Training Schemes is a problem of missing data. The attempts we make to isolate the mean treatment effect of YTS participation are complicated by the simple fact that persons can exist in either one of two "scheme states" (Participating or Non-participating) but not both at the same time. This would not be such a problem if those who had not experienced YTS were a suitable control for YTS participants. However one must consider that self-selection into government training schemes will exist and as a result non-participants will form a biased control group. The following argument helps to explain the nature of our predicament.

When we compare the outcome of people with YTS experience against the outcome for those without our conclusions will be biased by the inherent differences between those who need to spend time on such schemes and those who don't. It seems reasonable to expect that on average, persons in the non-treatment group would possess different personal and labour market characteristics compared to YTS participants. These enhanced characteristics may have an effect on the outcome of treatment. If this were the case then it would not be sensible to compare people with YTS experience to such a group as we may erroneously be attributing the estimated differences to a treatment effect when in fact it is merely due to the different types of people who undertake YTS and not.

In a well designed experiment we would expect the non-treatment group (or control group) differed only from the treatment group in the sense that they lacked treatment. One way to approximate such an experiment is to use a matching algorithm. The purpose of the matching algorithms described below are to produce a synthetic control group of individuals who have not experienced training but who are similar to the YTS participants. Ideally we would like to generate a control group who possess identical values for a given set of their labour market characteristics to be used as dependent variables when measuring the YTS treatment effect as experienced by those of the YTS treatment group. As we will demonstrate, the assumptions, which underpin the methods of matching which we describe, allow us to analyse our non-experimental data set in such a way that it identifies the same parameters as an experimental analysis.

2.4.1 Modelling Outcomes

A formalisation of the problem helps us to understand its nature and at the same time the way in which matching allows us to overcome it. We begin by introducing the notation of Heckman *et al.* (1997). Let

$$D = \begin{cases} 1 \text{ if a person participated in YTS,} \\ 0 \text{ otherwise,} \end{cases}$$

and

$$Y = \begin{cases} Y_1 \text{ denotes the outcome for participants,} \\ Y_0 \text{ denotes the outcome for non-participants.} \end{cases}$$

Also, X denotes the other characteristics which we use as conditioning variables, and $P(X) = Pr(D = 1|X)$. Using this notation we can define the outcome for any given individual as the sum of the observed outcomes, Y_1 and Y_0 such that

$$Y = DY_1 + (1 - D)Y_0. \tag{2.14}$$

Outcomes are then a function of observables, X and unobservables, (U_1, U_0):

$$\begin{aligned} Y_1 &= g_1(X) + U_1 \\ Y_0 &= g_0(X) + U_0 \end{aligned} \tag{2.15}$$

where $E(U_1) = E(U_0) = 0$. We assume that g_1 and g_0 are non-stochastic functions. The evaluation of the treatment effect can concentrate on the construction of many features of the missing data. In other words, Y_1 and Y_0 can take many forms. However the main parameter, which researchers often seek to identify from non-experimental data, is "The Mean Impact of Treatment on the Treated".

This gain from participating in the program is defined conditional on the set of personal and labour market characteristics X as

$$\begin{aligned} E(Y_1 - Y_0|X, D = 1) & \\ &= E(\Delta|X, D = 1) \\ &= g_1(X) - g_0(X) + E(U_1 - U_0|X, D = 1) \end{aligned} \tag{2.16}$$

where $\Delta = Y_1 - Y_0$. Using the matching methods described below we can estimate an averaged version of (2.16) over a subset of the support of X, $S_X, using$:

$$\Delta_{D=1} = \frac{\int_{S_X} E(\Delta|X, D = 1) f_x(X|D = 1) dX}{\int_{S_X} f_x(X|D = 1) dX} \tag{2.17}$$

where $f_x(X|D = 1)$ is the density of X.

If we could simultaneously observe Y_1 and Y_0 for the same person, there would be no evaluation problem and the need for matching would be removed, since one could construct $\Delta_{D=1}$ for everyone. This parameter may be interpreted as the gross gain from treatment experienced by participants.

2.4.2 Randomisation as a Solution to the Evaluation Problem

A randomised experiment with random attrition from the YTS program would allow us to recover the conditional distributions of Y_0 and Y_1, $F_0(y_0|D = 1, X)$ and $F_1(y_1|D = 1, X)$ respectively, by randomising at the point of enrolment so that some of those who would normally have been accepted for the scheme would be refused a placement. Our evaluation problem is the result of the non-experimental nature of our YCS data set. Such a randomised experiment would append our data with the necessary information to construct the missing counterfactual, $E(Y|D = 1, X)$ and thereby calculate (2.17).

If, having applied and been accepted for a position on the YTS scheme, participants had then been randomly refused entry, then randomisation would exist on the sub-population for whom $D = 1$. Let us define a new variable R such that

$$R = \begin{cases} 1 \text{ if a person from the } D = 1 \text{ population is randomised into YTS,} \\ 0 \text{ otherwise.} \end{cases}$$

It is not assumed that $E(U_1|X) = 0$ or $E(U_0|X) = 0$, therefore X is not required to be exogenous. In the presence of a randomisation and using the notation of previous sections observed outcomes, Y as defined in equation (2.14) is now

$$Y = D[RY_1 + (1-R)Y_0] + (1-D)Y_0 \tag{2.18}$$

and as a result

$$E(Y|D=1, R=1, X) = E(Y_1|D=1, X) = g_1(X) + E(U_1|D=1, X) \tag{2.19}$$

$$E(Y|D=1, R=0, X) = E(Y_0|D=1, X) = g_0(X) + E(U_0|D=1, X). \tag{2.20}$$

It is from equation (2.19) that we are able to estimate the outcomes for YTS participants. Equation (2.20) allows us to obtain the counterfactuals for these outcomes. If we then subtract equation (2.19) from (2.20) we arrive at

$$\begin{aligned} E(Y|D=1, R=1, X) &- E(Y|D=1, R=0X) \\ &= g_1(X) - g_0(X) + E(U_1 - U_0|D=1, X) \\ &= E(\Delta|D=1, X). \end{aligned} \tag{2.21}$$

If the YCS data set was the result of such a randomised experiment then we would be able to estimate the treatment effect via the above. In general $E(U_1, U_0|D = 1, X) \neq 0$. As we will show the methods of matching which we have used assume that given X

$$E(Y_0|D=1, X) = E(Y_0|D=0, X). \tag{2.22}$$

In essence we are assuming that there exists no group of variables Z which D depends on and which are not independent of $Y|X$. A randomised experiment would remove the need for this assumption when attempting to calculate the treatment effect for participants. However the absence of randomisation means that the distributions of the X variables for participants and non-participants may well be different. Whenever the support of the distribution of X conditional on YTS participation is different to that for X conditional on non-participation equation (2.22) will not hold and so we cannot retrieve information regarding the outcome Y. Since estimation of $\Delta_{D=1}$ (gross treatment effect on the treated) requires that the support of $X|D = 1$ is equal to that of $X|D = 0$ it is often necessary to examine the effect of treatment on a single variable of interest.

2.4.3 The Underlying Assumptions of Traditional Methods of Matching

As we have already touched upon, there exist several key assumptions which underpin the methods of matching we have employed. If these assumptions hold then it is possible to use matching to generate counterfactual estimates from which we may derive $\Delta_{D=1}$ even if the data is generated in a non-random setting.

Firstly we assume that having conditioned on X, the outcomes (Y_1, Y_0) are orthogonal to D, that is

$$(Y_1, Y_0) \perp D|X. \tag{2.23}$$

As long as this assumption holds we may infer that

$$F(y_0|X, D = 1) = F(y_0|X, D = 0), \tag{2.24}$$

which may be interpreted to mean that, conditional on X, persons who do not participate in the scheme $(D = 0)$ attain outcomes (y_0) with the same distribution as the outcomes which those who participated $(D = 1)$ would have received if they themselves were non-participants. Since we are already able to measure the outcome, y_1 for participants it follows from equation (2.14) that

$$E(Y_0|X, D = 1) = E(Y_0|X, D = 0). \tag{2.25}$$

Secondly we assume that

$$0 < Pr(D = 1|X) < 1 \quad \forall \ X \tag{2.26}$$

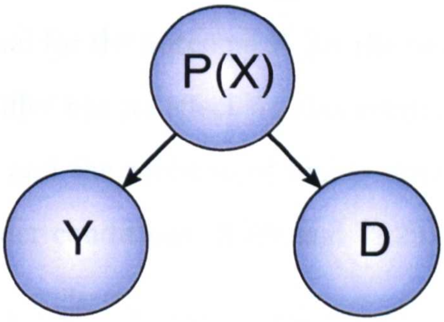

Figure 2.1: Relationship Between $P(X)$, the choice of D and the Outcome Y.

so that (2.16) can be defined for all X.

If these assumptions hold, non-experimental data sets such as our YCS data can identify the same parameters as an experimental data set with randomised scheme exclusions. In fact, for estimation of (2.16) it is enough to assume

$$Y \perp D|X. \tag{2.27}$$

Assumptions (2.23) and (2.26) which Rosenbaum and Rubin (1983) refer to as "strong ignorability" are all that is required to allow us to match on X. However, in reality X may consist of many variables which would lead to problems of dimensionality in large data sets. Fortunately Rosenbaum and Rubin (1983) demonstrate that assumptions (2.23) and (2.26) together imply

$$(Y_1, Y_0) \perp D|P(X) \tag{2.28}$$

and hence

$$Y_0 \perp D|P(X) \tag{2.29}$$

where $P(X) = Pr(D = 1|X)$ and is a propensity score[6] for scheme participation. Figure (2.1) illustrates the nature of the relationship between $P(X)$, a person's choice of D and the outcomes Y. Simply put, if we have two persons with the same $P(X)$, one with

[6]See section 2.6 for a description of propensity scores.

treatment and one without, then the diagram suggests that the outcome Y_0 for the non-participant is the counterfactual for the outcome Y_1 for the treatment person and that the fact that their choices of D differ has no effect on this relationship. Thus matching may be performed on $P(X)$ alone and the problem of highly dimensional X is replaced by a one-dimensional solution. Under conditions (2.28) and (2.29), we can use the estimate

$$\Delta_{D=1} = n_1^{-1} \sum_{i=1:D_i=1}^{n_1} Y_{1i}(X_i) - E(Y_{0i}|P(X_i), D_i = 0) \tag{2.30}$$

where n_i are the number of participants with X values that satisfy assumption (2.26).

2.5 Estimating Elasticities using Matched Samples

Having examined the underlying theory and the reasons behind the use of matching to correct for the non-experimental nature of many datasets we now present three algorithms, which can be used to perform such procedures.

2.5.1 The Nearest Neighbour Matching Protocol

We begin by introducing the simplest and most intuitive method, which is often referred to as nearest neighbour matching. The method of matching, known as Mahalanobis matching with propensity score callipers, is broadly similar to that described in Lechner (1999), and an advance on simple propensity score nearest neighbour matching, an early example of which is used by Cochran and Rubin (1973). An individual from that section of the population who did not participate in YTS, which we denote as the control group, is deemed to be similar to a given treatment person if their covariate structure is near to that of the treated individual.

Although our version of the nearest neighbour algorithm mirrors that of Lechner (1999) we choose to present it using the notation of Heckman *et al.* (1997). This ensures a consistency of argument when taken together with the theory from proceeding sections. Hence, we denote labour market variable responses for treatment or non-treatment (controls) as

Y_{1i} and Y_{0i} respectively. Variables which remain unaffected by treatment we label X_i and D_i is defined to be an indicator of YTS participation, taking value $D_i = 1$ if individual i receives treatment and $D_i = 0$ otherwise. The difficulty in estimating the treatment effect is highlighted by this notation. If an individual, i, receives treatment then we can observe Y_{1i}, however it is then impossible to observe Y_{0i}. If we make certain assumptions then it may become possible to determine the average causal effect of treatment, $\Delta_{D=1}$, as

$$\begin{aligned}\Delta_{D=1} &= E(Y_1 - Y_0|D = 1) \\ &= E(Y_1|D = 1) - E(Y_0|D = 1),\end{aligned} \tag{2.31}$$

where $E(\cdot|D = 1)$ is equal to the population mean for all those who participated in YTS. To determine the causal effects of treatment Lechner (1999) introduces the Stable Unit Treatment Value Assumption (SUTVA). SUTVA must be satisfied for the whole population. The main implication of which is that values of Y_{1i} and Y_{0i} do not depend on the treatment of individuals other than i.

We are unable to identify $\Delta_{D=1}$ due to the unobservability of the second term in (2.31), $E(Y_0|D = 1)$. If the assignment into YTS participation was random, then outcomes would be independent of assignment, in which case $E(Y_1|D = 1)$ would equal $E(Y_0|D = 0)$ and non-participating individuals could form a control group. However analysis of the data suggests that it is unreasonable to assume that assignment is independent of final outcome. As we showed in subsection 2.4.3, the work of Rubin (1977) and Rosenbaum and Rubin (1983) suggests a weaker condition known as the Conditional Independence Assumption (CIA), which they call "random assignment conditional on a covariate". They assume that assignment is independent of potential training outcomes conditional on a covariate, that is

$$Y_0 \perp D|X = x. \tag{2.32}$$

Equation (2.32) is assumed to be valid in all the support of x. Under CIA, $E(Y_0|D = 1, X = x) = E(Y_0|D = 0, X = x)$. The probability of participation in YTS conditional on x $[P(D=1|X=x)]$ is called the propensity score which we denote $P(x)$. Given that $0 < P(x) < 1$ holds, then $E(Y_1|D=1) = E[E(Y_0|D=0, X=x)|D=1]$ can be estimated for large samples. If equation (2.32) holds, then Y_0 is also independent of D conditional on $P(X) = P(x)$ and so

$$E[Y_0|D=1, P(X)=P(x)] = E[(Y_0|D=0, P(X)=P(x)]. \qquad (2.33)$$

Hence estimation is possible using

$$E(Y_0|D=1) = E\{E[Y_0|D=0, P(X)=P(x)]|D=1\}. \qquad (2.34)$$

Thus the dimension of the estimation has been greatly reduced, however the price to pay for this simplification is the need to estimate the assignment probability for each individual using a probit model.

To begin with we select variables, which are to be used to gauge similarity. These variables, defined as V, are the independent variables in a probit model estimation for YTS participation. So that we might examine the differing effects of YTS on both men and women the data set is split before matching to prevent persons being matched to individuals of the opposite sex. This leads to a need to run the matching procedure twice. We then compute $v\hat{\beta}$ and its conditional variance $var(V\hat{\beta}|V=v)$ for each observation in the dataset. Persons are split into a treatment vector (people with YTS experience) and a control vector. The order of individuals within these two vectors is then randomised.

The first person in the randomised T vector is selected. A calliper is constructed (see Appendix C.1, Matching Algorithm) so that it forms a 90% confidence interval around the propensity score, $v_{n_t}\hat{\beta}$ of this individual. Members of the C vector who lie within this caliper are denoted j. If $j \geq 1$ then we select the observation j which minimises the distance

$$\gamma(j) = (v_j\hat{\beta} - v_{n_t}\hat{\beta})^2. \qquad (2.35)$$

If $j = 0$ then it is necessary to calculate the distance for all members of the C vector and match to the one with the minimum $\gamma(j)$. The matched individual is then removed from the C vector and the matching process is repeated for the next individual in the T vector. This is repeated until all individuals from the treatment vector T, are matched with a control from the C vector. The algorithm can be adjusted so that matched controls are not excluded from the dataset and as a result may appear more than once in the matched dataset. This may lead to an over dependence on a few controls. Hence we choose to employ a method with matched control exclusions.

2.5.2 The Kernel Regression Matching Estimator

Unlike the "nearest neighbour" matching protocol described in subsection 2.5.1, kernel regression matching uses the whole comparison sample. A kernel regression estimator assigns each member of the control group a weight so that the control observations closer in terms of distance between propensity score, $|P(X_{1i}) - P(X_{0j})|$, to a treatment person receive greater weight. These weighted people are then fused together to form a synthetic control person.

Let $W(i,j)$ be the weight placed on observation j in forming a mean synthetic comparison individual for treatment observation i with

$$\sum_{j=1}^{N_0} W(i,j) = 1 \quad \text{and} \quad 0 \leq W(i,j) \leq 1 \tag{2.36}$$

where N_0 is the total number of control group individuals. Then the weighted mean synthetic comparison individual for treatment person i is

$$\overline{Y}_{0i} = \sum_{j=1}^{N_0} W(i,j) Y_{0j}. \tag{2.37}$$

Using this equation we can obtain "synthetic" control responses for each person in the treatment group using the propensity scores. This can be done for all variables required for analysis. During this work we refer to collections of synthesised variables generated to match those of a treatment person as a "Synthetic Control Person". Repetition of this process for the whole of the treatment group produces a "Synthetic Control Group". The weights are calculated using a kernel method as

$$W(i,j) = \frac{K\left(\dfrac{P(X_j) - P(X_i)}{h_n}\right)}{\displaystyle\sum_{j=1}^{N_0} K\left(\dfrac{P(X_j) - P(X_i)}{h_n}\right)}, \tag{2.38}$$

where K is a kernel and h_n is the bandwidth (smoothing parameter). A commonly used kernel function is the biweight kernel of the form presented in equation (2.39)

$$K(x) = \begin{cases} \dfrac{15}{16}(x^2 - 1)^2 & \text{if } |x| < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.39}$$

This function weights and combines all persons in the control group differently for each person in the treatment group with a unique $P(X_i)$. The impact of treatment is then estimated by way of the mean difference across all treated individuals as

$$m = \frac{1}{N_1}\sum_{i=1}^{N_1}\left(Y_{1i} - \overline{Y}_{0t}\right) = \frac{1}{N_1}\sum_{i=1}^{N_1}\left(Y_{1i} - \sum_{j=1}^{N_0} W(i,j)Y_{0j}\right). \tag{2.40}$$

2.5.3 Local Linear Regression Matching Estimator

Here we present a method of matching which is similar in nature to that of kernel matching. The weights, which the local linear method assigns to non-participants, are again a reflection of their similarity to the participant of interest and take the form

$$W_j(P(X_i)) = \frac{K_{ij}\sum_{k=1}^{n_0} K_{ik}(P_k - P_i)^2 - [K_{ij}(P_j - P_i)][\sum_{k=1}^{n_0} K_{ik}(P_k - P_i)]}{\sum_{j=1}^{n_0} K_{ij}\sum_{k=1}^{n_0} K_{ik}(P_k - P_i)^2 - [\sum_{j=1}^{n_0} K_{ij}(P_j - P_i)]^2} \tag{2.41}$$

where

$$K_{ij} = K\left(\frac{P(X_i) - P(X_j)}{h_n}\right)$$

and $K(\cdot)$ is the biweight kernel function as defined in equation (2.39). Heckman *et al.* (1997) recommend the weights that the above function delivers over those produced by the kernel regression method due to their high rate of convergence at boundary points. In fact, the order of the bias of local linear regression weights is the same at the boundary points as elsewhere. The ability to adapt to differing design densities is also a strong feature of this estimator since its bias does not depend on the density of $P(X)$. As with the kernel regression matching estimator, the weights from equation (2.41) are used to calculate the treatment effect via equation (2.40).

2.6 Calculation of the Propensity Score

An individual's propensity score is simply the probability of them having a certain characteristic or doing a certain activity, given their personal set of characteristics and that they have been exposed to a particular series of external factors. Since they are in fact predicted probabilities, all propensity scores lie between 0 and 1.

All three of the matching methods, which we have outlined during this chapter, require us to have already calculated a series of propensity scores for YTS participation. When using these propensity scores we will rely on the result of Rosenbaum and Rubin (1983), section 2.4.3, where we showed how, through a series of assumptions, they arrived at result 2.29, which implied that a propensity score constructed from a model containing a series of relevant variables could be matched upon and that problems of highly dimensional X were then replaced by a one dimensional solution.

We can estimate propensity scores for a group of individuals using one of a number of models. Heckman *et al.* (1997) choose the conditioning variables in X such that the logit model they employed to estimate each individuals propensity to participate in YTS produces the highest number of correct participating state predictions. We elected to make use of probit models for the construction of our YTS participation propensity scores. A probit model is defined to be

$$P(D_i \neq 0) = F(XB) \tag{2.42}$$

where $F(\cdot)$ is the cumulative distribution function, and XB is the probit score. Those wishing to interpret the coefficients from the model should bear in mind that since XB has a Normal distribution, each unit increase in a dependent variable with a coefficient of, say 0.8, implies an increase in the probit score of 0.8 standard deviations. However, interpretation of the results of our probit model is not the main focus of this work. The scores, which we obtain, are needed for use in the matching procedures we have described. With the scores defined to be $P(y_i \neq 0)$ we are generating probabilities for persons having the outcome $y_i = 1$, which in this case is the probability that they took part in YTS.

Probit models are employed when the dependent variable that you wish to run a regression for takes binary discrete values. Our dependent variable took the value 1 if a person indicated that they had undertaken a period of YTS training and 0 otherwise. The results of our propensity score estimations by gender using these models are presented in Table C.4. A wide variety of labour market characteristics, personal (education level etc.) and locational (regional and cohort dummies) were included in the model. See Table 2.1 for a description and summary statistics of the variables used. Notice that the female sample consists of around 700 more individuals than the male sample. This is consistent with the higher response rate amongst females for the whole of the YCS dataset. See Table 2.2 for an outline of the YCS cohort design structure as well as response rates per sweep for each cohort.

2.7 Stochastic Dominance

Imagine that we have generated a set of propensity scores for YTS participation. We can separate the persons for whom these scores were predicted into two groups, treatment and control. All those in the treatment group experienced YTS; all those in the control group did not. When attempting to match on these scores we will be assuming that there are sufficient high propensity control persons to match to those who received treatment. All this despite prior knowledge suggesting that those from the treatment group would, on average, have higher propensities. Hence we shall need a control group sufficiently large in number to offset this lack of high propensity controls. This poses the problem, how do we know if our control group is large enough to supply us with the persons we need?

We begin by introducing the notion of stochastic dominance and explain what we mean when we say that one distribution dominates another. Next we investigate whether, by only observing the properties of the cumulative distribution functions of the treatment and control propensity score distributions, we can say if $F(P(X_1))$ dominates $G(P(X_0))$ or not? The concepts of first and second order stochastic dominance can help us to form a conclusion as to whether our dataset will produce good matches. We conclude this section with an explanation of how to achieve this goal.

2.7.1 First Order Stochastic Dominance

If for any propensity score there is more chance of there being a person with that score from population F than population G, then we can say that, on the whole, F dominates G.

If a propensity score distribution has a lot of probability mass skewed towards higher propensities then it is more likely to contain high propensity persons than a distribution which has a lot of probability mass skewed towards low propensities. If this is the case, then we have what is called "first order stochastic dominance". This is defined as follows.

Definition 2.1 First Order Stochastic Dominance: F is said to dominate G on $[a, b]$

according to first order stochastic dominance if and only if $F(p) \leq G(p)$ for all $p \in [a, b]$. This is denoted $F \geq_{D^1} G$.

Figure 2.2 demonstrates this more clearly. When comparing the cumulative distribution function H with either F or G. Specifically, note that for any $p \in [a, b]$, $F(p) \leq H(p)$ and $G(p) \leq H(p)$ as H lies uniformly above F or G. Considering any $p \in [a, b]$, then we see immediately that there is more "area" under the curve between a and p than there is under the F or G curve between a and p. Thus, there is a greater probability mass under $H(p)$ than $F(p)$ or $G(p)$ for any $p \in [a, b]$, i.e. the probability that any t is less than p under H is greater than the probability that any t is less than p under F or G. Thus both F and G dominate H.

However, note that this criterion fails when comparing the cumulative distribution functions F and G with each other. Obviously, neither F nor G are uniformly above each other thus it neither F dominates G nor G dominates F. In particular, note that below

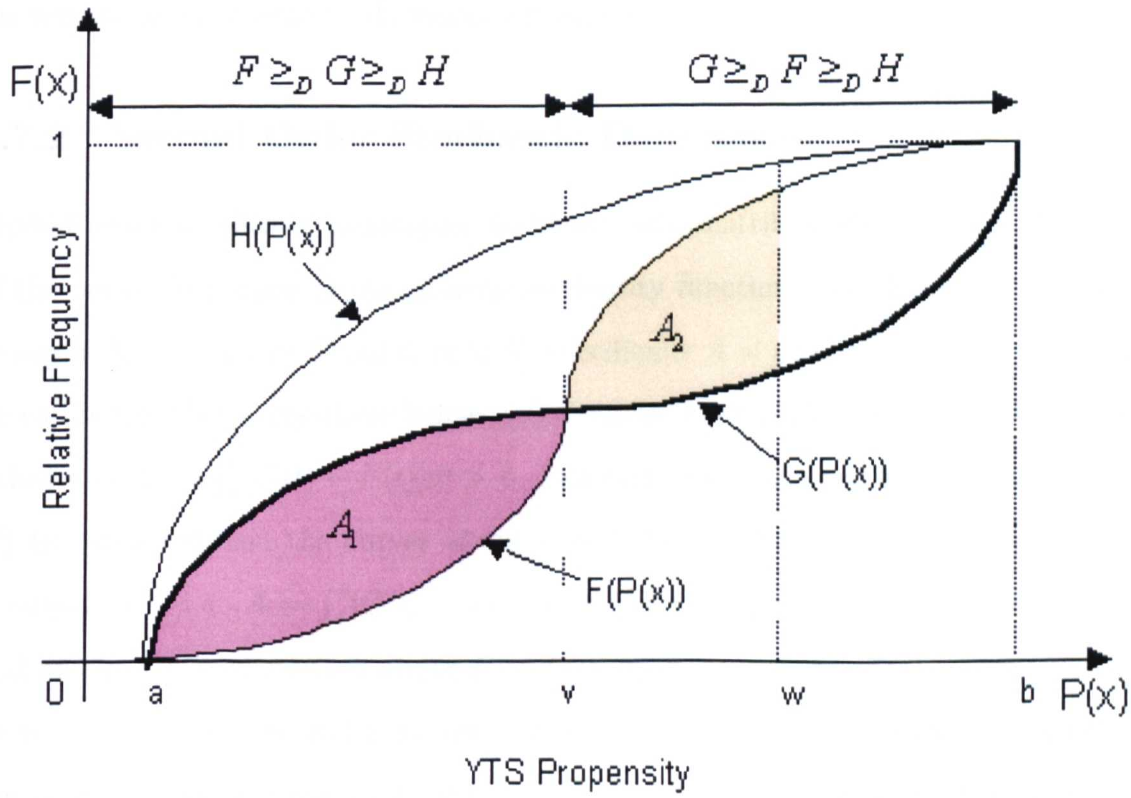

Figure 2.2: First and Second Order Stochastic Dominance

point v, $F(p) \leq G(p)$, whereas above point v, $F(p) \geq G(p)$, or $[1-G(p)] \leq [1-F(p)]$. As we cannot compare F and G according to the first order stochastic dominance criterion, then $\geq_{D^1}$ is only a partial ordering on the space of cumulative distribution functions.

The purpose of the first order stochastic dominance is to enable us to order (if only partially) distributions according to the degree to which the distribution is skewed towards the right and high propensities: in other words, $F \geq_{D^1} H$ implies that F is unambiguously likely to contain more high propensity persons as a percentage of its total, than H because it deposits the bulk of its probability mass amongst higher propensities, and thus will have a higher expected mean propensity. But, as we see in Figure 2.2, we cannot compare F and G by this criterion. If v is the mean, then F and G have the same expected mean propensity. However, in Figure 2.2, it also seems as if the probability mass of F is "less dispersed" than G. An alternative criterion, then, would be to argue that a particular distribution contains more high propensity persons than another as a percentage of its total if it has an unambiguously higher expected mean propensity. This is the purpose of the second order stochastic dominance criterion.

2.7.2 Second Order Stochastic Dominance

Second order stochastic dominance helps us rank distributions in terms of the spread of the probability mass of the cumulative density functions. Let us define A as the area between the two curves F and G in $[a, b]$, specifically $A = \int_a^b [G(t) - F(t)]dt$. In Figure 2.2, we can notice that everywhere below v, G is above F, thus the area in between the curves below v is $A_1 = \int_a^v [G(t) - F(t)]dt > 0$. However, above v, G is below F, thus (negative of) the area between the curves above v is $\int_v^b [G(t) - F(t)]dt < 0$. But notice that A is cumulative, i.e. $A = \int_a^v [G(t) - F(t)]dt + \int_v^b [G(t) - F(t)]dt$ for all $p \in [a, b]$, thus we add up the areas where G is above F and subtract from it the areas where G is below F (note: we can allow G and F to rise above and dip below each other several times over the range). Thus, in Figure 2.2, $A(w) = \int_a^w [G(t) - F(t)]dt$ is the sum of the areas A_1 and A_2 where, $A_1 > 0$ and $A_2 < 0$.

If $A \geq 0$ for all a, b p is undefined here., then the cumulative area where G is above F is greater than the area where F is above G, i.e. the probability mass of G is more spread out than the probability mass of F, which with reference to Figure 2.2, implies that G is dominated by F. Conversely, if $A \leq 0$ for all $p \in [a, b]$, then the mass of F is more spread out than G. We want to "rank" distributions F and G according to whether A is positive or negative over the entire range. Thus, we now define the following:

Definition 2.2 Second Order Stochastic Dominance: F dominates G according to second order stochastic dominance, or $F \geq_{D^2} G$, if $A = \int_a^b [G(t) - F(t)]dt \geq 0$ for all $p \in [a, b]$.

Notice in Figure 2.2 that it is quite probable that $F \geq_{D^2} G$ by this criterion, thus F dominates G according to second order stochastic dominance. Notice also that if A compares H with F or G, then it is indeed true that $F \geq_{D^2} H$ and $G \geq_{D^2} H$. Thus, first order stochastic dominance implies second order stochastic dominance, but not vice-versa.

2.7.3 Judging whether a Dataset will produce Good Matches

Ideally we want to develop a set of rules for good matching practice. A dataset whose properties are deemed to adhere to these rules would then be judged to be suitable for matching to occur. The method we propose is more *ad hoc* than constituting a *universal solution.* Knowledge of the nature of the dataset which we are interested in is the basis from which we develop a test. Hence, although our method may be adapted to test for suitability amongst other datasets an understanding of the difference in propensity score distributions between treatment and controls must be obtained before considering the use of the concept of stochastic dominance.

Prior understanding of the nature of YTS participation leads us to expect that those who undertake a period of YTS training will produce a propensity score distribution which is skewed to the right whilst those from the non YTS grouping will have lower average YTS propensities. Hence, it is at the right side of the distribution that we are likely to find too few controls. We are searching for fragments along the propensity score range

($0 \leq P(x) \leq 1$) for which, given that both treatment and control groups contain the same number of individuals, the control group would have insufficient numbers to match to those from the treatment group. Formally we test for second order stochastic dominance of the control group propensity scores verses the treatment group scores. If the control group does not stochastically dominate the treatment group then we may conclude that the treatment group has more high propensity persons as a percentage of it's total, leading to the need for us to sample a larger number of controls to offset the shortfall.

A plot of both the treatment and control group CDFs allows us to visually identify areas of concern. Any section of the treatment curve with a gradient in excess of that corresponding section of the control group will, in samples of equal size lack support for high quality matching. As already observed, it is the high propensity end of the dataset, with which we are likely to uncover a lack of support.

When attempting to identify whether data will produce a well matched dataset there are three criteria which must be met. If we can prove that a pair of treatment and control propensity score distributions meet all three, then we can say that there are enough control group people whose propensity to participate in YTS is sufficiently similar to those from the treatment group over the full range of propensity scores ($0 \leq P(x) \leq 1$) to produce a well matched dataset.

1. **Maximum and Minimum Values of the Propensity Score Distributions.** The control group propensity score distribution must sit to the left and right of the treatment group propensity score distribution. It is possible that the control group propensity score distribution can stochastically dominate the treatment group propensity score distribution whilst still producing a series of poor (perhaps very poor) matches. Figure 2.3 presents an example of a pair of propensity distributions which would fail to produce good matches. If we imagine that $F(P(x))$ represents the treatment group propensity score distribution and $G(P(x))$, the control group propensity score distribution, then we see that the control group distribution stochastically dominates the treatment distribution. Hence, the control group has

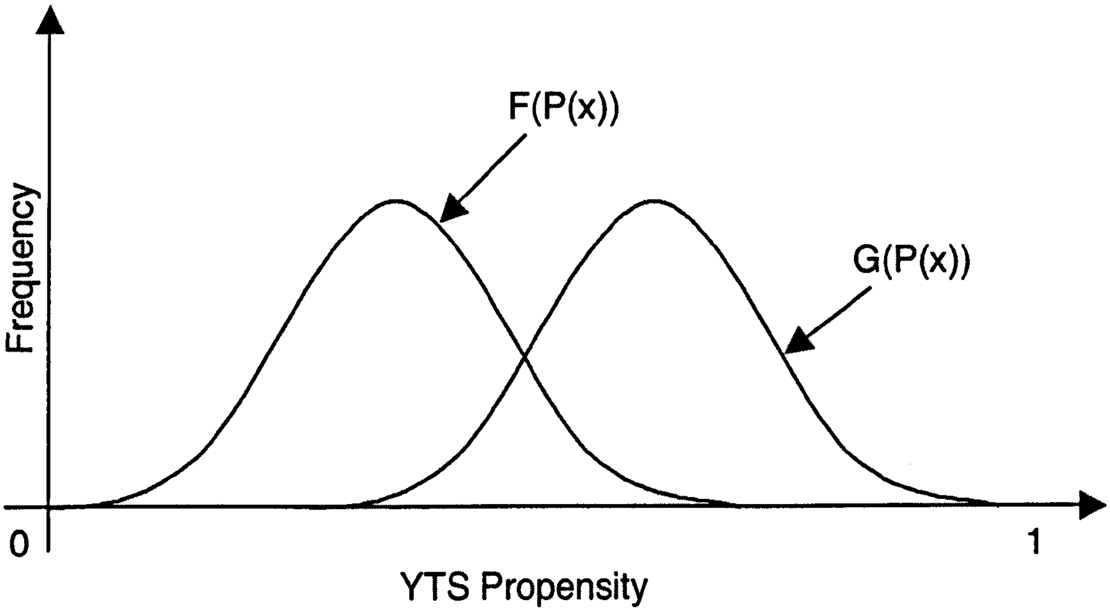

Figure 2.3: Maximum and Minimum Values of the Propensity Score Distributions

a large number of high propensity individuals. However, in this case there are insufficient low propensity persons to match with the treatment group. To counteract this we begin by ensuring that the propensity score distributions obey the following rule

$$\min(P_{1i}(x), P_{0i}(x)) = \min(P_{0i}(x)) \tag{2.43}$$

$$\max(P_{1i}(x), P_{0i}(x)) = \max(P_{0i}(x)) \tag{2.44}$$

where P_{1i} is a sample point from the treatment group and P_{0i} is a sample point from the control group.

2. **Establish the Nature of Dominance.** The above ensures that the control group distribution spans the range of the treatment group distribution. We still need to uncover the relationship between the treatment and control group propensity score distributions. If the treatment group propensity distribution is found to 1st order stochastically dominate that of the control group (our prior belief), then we can say that if $n_0 = n_1$ the treatment group contains more high propensity individuals as a

percentage of its total than the control group.

If after plotting the cumulative distributions of the treatment and control groups we discover that, over the full range of $P(x)$, neither distribution 1st order stochastically dominates the other, then we must consider second order stochastic dominance. When investigating second order stochastic dominance relations the question of whether one distribution dominates another depends on what range of the variable you are considering. Returning to the pair of distributions, G and F, from Figure 2.2 then we can say that if v represents the halfway point in the range of $P(x)$ then F 1st order stochastically dominates G in the range a to v and that G 1st order stochastically dominates F in the range v to b. If we were to substitute the control group propensity score distribution for distribution F and the corresponding treatment group distribution for G, then we could say that for distributions of equal size the control group would not contain sufficient numbers of individuals in the tails of its propensity score distribution to match to those from the treatment group. Since we are interested in discovering any fragments of the propensity score range for which we lack sufficient control persons for matching it makes sense to consider the distributions in fragments of infinitely small size. Hence we note the areas for which the gradient of the control group CDF is less than that of the treatment group.

A full investigation of the nature of the dominance relations between the treatment and control group distributions will facilitate an understanding of the ranges of propensity scores for which the control group will lack sufficient numbers to perform high quality matches when using groups of equal size. The YCS dataset contains far more control than treatment individuals.

3. **Account for the Relative Size of the Treatment and Control Groups.** The discovery of a range of propensity scores over which the treatment group stochastically dominates the control group does not necessarily indicate that there will be insufficient control persons to perform a full series of high quality matches. Instead

it indicates that for treatment and control groups of **equal size** there will be insufficient high quality matches. Hence, we must also examine the ratio of treatment and control group sample sizes. Since if $N_0 > N_1$, this may be sufficient to offset the shortfall, in the control group.

To summarise, the definitions of first and second order stochastic dominance, combined with the relative sizes of the treatment and control groups, enable us to predict whether a dataset will produce a series of good matches when using a nearest neighbour matching algorithm. If these concepts reveal a lack of high propensity controls, then we could randomly exclude some persons from our treatment group. The number of exclusions needed being a function of the degree to which the treatment group propensity score distribution stochastically dominates that of the control group and the relative size of the two sub-samples. We will return to this issue in Chapter 4. Cochran and Rubin (1973) examine the degree of concordance of key variables between pre-match and post-match datasets. This research does not contain a detailed examination of concordance. We focus on YTS treatment effects and the sensitivity of matched datasets to the matching procedure being used.

Chapter 3

Investigating the YTS Treatment Effect

3.1 Introduction

In this chapter we shall present some early results from our investigations into the relationship between YTS and the job search process for YCS school leavers. We calculate the job search elasticities as derived by Lancaster and Chesher (1983) (see section 2.3) for the men and women of the YCS dataset in an attempt to uncover the ways in which a spell on YTS affected their job search strategies. The introduction of a series of important considerations leads us to refine the model further and present alternate estimations. The need to allow for self selection into YTS necessitates use of the nearest neighbour matching algorithm for the first time.

3.2 Comparisons with Previous Studies

We begin by calculating elasticities of the structural search model as set out in Lancaster and Chesher (1983) and described in section 2.2 for males and females in the YCS dataset. For purposes of comparison various population elasticities estimated from previous studies that used this approach are included in one table. Our sample size of 4,115 young men, compares favourably to those used in many previous studies (the sample sizes in Gorter and Gorter (1993), Jones (1988), Lancaster and Chesher (1983), Main and Shelley (1988) and Lynch (1983) were 213, 845, 639, 838 and 52 respectively). The relatively large sample size should allow us to get quite precise estimates of the structural parameters. Comparing our results for all young men and women in the final two columns of Table 3.1[1] with those of the other studies we find that while our results are within the range of estimates from the previous studies, there are some differences.

One reason for the relative difference between our estimate and that of the other studies is that our data consists solely of young people. In contrast the sample in Dolton and O'Neill (1995) consisted of workers who were unemployed at least 6 months, the Lancaster and Chesher (1983) estimates refer to a sample from the stock of *all* unemployed workers. As a result the reservation wage reported by the individuals in the other studies may already be at the minimum level necessary to insure a positive payoff to work. To the extent that the reservation wage may become more rigid in a downward direction the longer the unemployment spell continues and the older the sample is, we might expect changes in benefits to have more of an effect on reservation wages in our sample of young people than in other studies.

Our estimate of the elasticity of the reservation wage with respect to the offer arrival rate is 0.1776 for men and 0.1754 for women which again is higher than the estimates of 0.11, 0.15 and 0.14 reported by Lancaster and Chesher (1983), Lynch (1983) and Main and Shelley (1988) respectively. This may be due to the responsiveness of young people to an exogenous change in labour market circumstances. It is possible that older job seekers and the longer term unemployed are less responsive to a changing offer arrival rate.

The third row shows the elasticity of re-employment with respect to unemployment benefits under the Pareto assumption, while the fourth row reports the estimate under the Exponential assumption. Our estimate of the elasticity for men is close to −0.41

[1]Throughout this work figures in parenthesis are standard errors

	Other Studies					This Study	
Elasticities	Lancaster & Chesher	Lynch	Main & Shelly	Gorter & Gorter	Dolton & O'Neill	All Men	All Women
$\frac{\partial \log \xi}{\partial \log b}$	0.14 (-)	0.11 (-)	0.16 (0.007)	0.28 (-)	0.11 (0.002)	0.1033 (0.0016)	0.1151 (0.0017)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.11 (-)	0.15 (-)	0.14 (0.004)	0.09 (-)	0.10 (0.002)	0.1865 (0.0018)	0.1825 (0.0015)
$\frac{\partial \log h}{\partial \log b}$ +	-1.0 (-)	-0.48 (-)	-0.91 (0.03)	-2.89 (-)	-0.99 (0.03)	-0.4437 (0.0064)	-0.4891 (0.0043)
$\frac{\partial \log h}{\partial \log b}$ ++	-0.96 (-)	-0.56 (-)	- (-)	- (-)	-0.87 (0.03)	-0.3405 (0.0061)	-0.3740 (0.0037)
$\frac{\partial \log h}{\partial \log \lambda}$ ++	0.24 (-)	0.25 (-)	- (-)	- (-)	0.21 (0.004)	0.2897 (0.0030)	0.2976 (0.0028)
Sample Size	639	52	338	112	3520	4115	4770
+ Pareto Assumption, ++ Exponential Assumption							

Table 3.1: Elasticities for Previous Studies

under the Pareto assumption and -0.33 under the exponential.[2] The corresponding two elasticities for women are -0.48 and -0.38 indicating that, with regard to the hazard of leaving the state of unemployment, women may well be more sensitive to a change in the benefit level than men.

The fifth row of Table 3.1 reports the elasticity of re-employment probability with respect to changes in the offer arrival rate, under the Exponential assumption.[3] These

[2]Our figure is substantially lower than the estimate of -2.89 obtained in Gorter and Gorter (1993). However using parametric procedures Gorter and Gorter (1993) are unable to find a significant effect of either benefits or reservation wages on duration.

[3]In the Table we present the estimates reported by Lancaster and Chesher (1983) and by Lynch (1983) in their papers. However, these authors seemed to have overlooked the fact that under the Pareto assumption these elasticities should be equal to the elasticity of the reservation wage with respect to the

results suggest an estimate around 0.27 for men and 0.28 for women. An alternative interpretation of these figures can be obtained if one assumes the hazard function for a given individual is constant over time. In this instance an individual's completed unemployment duration has an exponential distribution, with mean $1/h$. The elasticity of mean unemployment duration with respect to the offer arrival rate is then the negative of the estimates given in rows 1 and 5. Our results imply that a 10% fall in the mean time between offers is associated with around a 1% to 3% reduction in the average length of unemployment duration.

It is particularly interesting to consider the extent to which our results mirror those of Lynch (1983). Her dataset was collected from a survey of young people living in Greater London during March of 1979 who planned to leave school that summer at the minimum legal age of sixteen. Individuals who remained in the labour force were swept every six months after departure from school. Although the original sample consisted of 1922 individuals, the dropout rate was substantial. The results which we reproduce here were calculated using the sweep of November 1980, by which time only about 70% of the original sample remained in the labour force. Beyond the obvious similarities between the YCS and Lynch's datasets, she also used responses to questions such as "What is the lowest weekly wage you would accept before tax and other deductions?" to represent reservation wages, "How much do you expect to earn before tax and other deductions?"for expected wages and "How do you manage for money while you are out of work and how much does that amount to each week?" as a measure of benefits received; All of which mirror the questions which were put to YCS respondents. It is a source of some encouragement that her results are so similar to ours, in both sign and magnitude for all five elasticities of interest. Perhaps the differences between our results and those of Lynch (1983) when compared to those of Lancaster and Chesher (1983), Main and Shelley (1988) and Gorter and Gorter (1993) are in fact a reflection of the nature of the datasets employed.

All of these studies, whose authors based their calculations on the measurement of

benefit level.

reservation and expected wages as responses to questions at interview, present evidence which broadly supports the existence of a number of relationships which the job search model of section 2.2 predicts. Results suggest that higher benefit levels lead to higher reservation wages, whilst higher job offer arrival rates lead to an upward pressure on reservation wages. At the same time the hazard of exiting from unemployment into work falls when benefit levels rise.

3.3 Search Model Elasticities

	Samples					
Means	**All Men**	**Men YTS=1**	**Men YTS=0**	**All Women**	**Women YTS=1**	**Women YTS=0**
Expected Wage	126.1087 (0.66737)	**124.5477** (1.07512)	126.7068 (0.82591)	115.8429 (0.54831)	**111.327** (0.89479)	117.631 (0.67612)
Reservation Wage	96.16908 (0.49596)	**98.03103** (0.83039)	95.45559 (0.60731)	88.6961 (0.42353)	**86.06494** (0.68807)	89.73794 (0.52371)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	0.10328 (0.00164)	**0.08422** (0.00222)	0.11058 (0.00209)	0.11506 (0.00170)	0.11208 (0.00306)	0.11624 (0.00203)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.18646 (0.00176)	**0.17300** (0.00333)	0.19162 (0.00207)	0.18255 (0.00153)	**0.17650** (0.00282)	0.18494 (0.00182)
$\frac{\partial \log h}{\partial \log b}$+	-0.44374 (0.00638)	**-0.40713** (0.00444)	-0.45776 (0.00865)	-0.48906 (0.00426)	-0.49138 (0.00587)	-0.48814 (0.00547)
$\frac{\partial \log h}{\partial \log b}$++	-0.34046 (0.00608)	**-0.32291** (0.00382)	-0.34718 (0.00828)	-0.37400 (0.00371)	-0.37930 (0.00476)	-0.37190 (0.00482)
$\frac{\partial \log h}{\partial \log \lambda}$++	0.28974 (0.00303)	**0.25722** (0.00525)	0.30220 (0.00365)	0.29761 (0.00283)	**0.28858** (0.00526)	0.30119 (0.00335)
Sample Size	4115	1140	2975	4770	1353	3417

+ Pareto Assumption, ++ Exponential Assumption

Table 3.2: Job Search Elasticities and Wages for Men and Women

Examination of the different government training schemes which existed over the 1984-94 period reveals important differences with respect to their duration, training content and how they were perceived by employers. In addition the various incarnations of the scheme operated against different background unemployment rates and training allowances. These background factors are graphed in Figures C.5 and C.6. Without losing sight of the historical economic environment, it is interesting to examine whether different YTS schemes induce different job search behaviour and expectations on behalf of their recipients. In Table 3.2 we estimate the elasticities for and by whether or not men and women have ever been on YTS. This allows us to compare the parameters across subgroups. The first column of Table 3.2 reports the estimates for all men, while the next two columns compare elasticities of those young men who have had a spell on youth training with those who have not. The last three columns report the corresponding estimates for young women.

If we assume that both reservation and expected wages are normally distributed then the computation of elasticities for the different subgroups facilitates a direct statistical test of the differences in the means between the comparison groups. A two sample t-test with null hypothesis of $H_0 : \mu_1 = \mu_2$ verses a two tailed alternative was performed, where μ_1 and μ_2 were the mean values of a given elasticity or expected/reservation wage for those with YTS experience and those without. Test results are presented in Table 3.2. Throughout this work all emboldened elasticities are significantly different to the corresponding elasticity of their alternate specification at the 10% level; those which are shaded are significant at the 5% level.

The same statistical test is used to investigate differences in the means of the reservation and expected wages between subsamples. The difference between the reservation wages of men who had some experience of YTS and those who didn't was significantly different from zero at the 5% level whilst the expected wage was significant at the 10% level. Tests on the exit status mean values failed to reject the null hypothesises[4].

Tests on elasticities for the "YTS participation" subgroups led to some interesting

[4]Comparisons between the elasticities of those who exit from unemployment before the end of the survey and those who don't prove inconclusive and are not reproduced here. None of the elasticity couples between these two subgroups proved significantly different from each other.

conclusions. Young men who spend time on YTS have reservation wage elasticities with respect to both the benefit level and the arrival rate of job offers which are significantly lower than for those with no YTS experience. Under both the Pareto and Exponential wage offer distributional assumptions we see that young men with YTS experience have an unemployment leaving hazard elasticity with respect to the benefit level which is significantly lower than for individuals without YTS training. The hazard elasticity with respect to the arrival rate of job offers under the Exponential assumption is also significantly less for male YTS participants.

These findings taken together suggest that those young people who are more employable, have better qualifications and hence higher reservation wages, and those who have been on YTS, have lower wage expectations. It would seem that the propensity to change one's reservation wage following a benefit change or a change of the offer arrival rate is affected by whether one has been on YTS.

Tests on the mean reservation and expected wages of the female participants in the YCS Cohort survey indicated that differences for both wage types were significantly different from zero at the 5% level if the women had a spell on YTS[5]. Elasticity estimation produced rather mixed results. The elasticity of the reservation wage with respect to the job offer arrival rate was significantly different between women with a spell on YTS and those without. The unemployment leaving hazard elasticity with respect to the arrival rate of wage offers under the exponential wage offer distributional assumption was significantly different with YTS participation.

When comparing the results for males and females in Table 3.2 one could conclude that there is some evidence to suggest that for females, the experience of YTS lowers reservation and expected wages but that experience of YTS has a more marked effect on the sensitivity of the elasticities for men than for women.

[5]As with the male subsamples, mean values for the exit status subsamples were not significantly different from one another.

3.3.1 Heterogeneity of YTS Treatment Types

Up until now our examination of the YTS treatment effect has allowed for the difference between the male and female experience of YTS. However there is another potential bias in the data which casts a shadow over the validity of our conclusions thus far. In this section we attempt to account for our second area of concern, that of heterogeneity of YTS treatment types.

In Table 3.2 the grouping together of all YTS types obscures the differences between these schemes. The literature points to the considerable differences existing between these schemes, prompting us to consider the estimation of separate elasticities for different training regimes. Table 3.3 contains elasticities for men and women by YTS scheme type. Taking men and women separately we investigate the treatment effects participants received from the three YTS schemes. We compare the shaded cells contain elasticities and mean reservation/expected wages which are significantly different to the corresponding elasticities and wages of Table 3.2, (YTS=0), at the 5% level.

The expected wages of men on YTS I and women on types YTS I and YTS II were significantly different from their YTS=0 mean values at the 5% level. The mean for men with YT experience was significant at the 10% level. Reservation wages were significantly different from their YTS=0 equivalents at the 5% level for YTS II and YT men and significant for YTS I and YTS II women at the 5% and 10% levels respectively.

The elasticity of reservation wages with respect to a change in the benefit level was significantly different to that of men with no YTS experience for all versions of YTS and for women who had spells on YT. The elasticity of the reservation wage with respect to the arrival rate of wage offers was significantly different to the elasticities of YTS=0 for YTS I and YTS II for both sexes.

The elasticities of the unemployment exit hazard with respect to changes in the benefit level were significantly different to those of YTS=0 for both sexes for all but men with YTS I experience under the exponential assumption and women with YT experience under both assumptions. The elasticities of the hazard with respect to the wage offer arrival

Means	Men YTSI=1	Men YTSII=1	Men YT=1	Women YTSI=1	Women YTSII=1	Women YT =1
	Samples					
Expected Wage	**120.0951** (1.67594)	126.1986 (1.50209)	**131.1896** (3.16567)	**104.0202** (1.27979)	**113.5821** (1.30736)	120.6493 (2.51888)
Reservation Wage	93.78476 (1.25718)	**101.0787** (1.21100)	**100.3479** (2.36464)	**81.44474** (0.98226)	**88.15212** (1.01707)	90.23697 (1.95561)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	**0.090126** (0.003586)	**0.077312** (0.003212)	**0.088285** (0.005795)	0.111862 (0.004846)	0.106149 (0.003994)	**0.127983** (0.009581)
$\frac{\partial \log \xi}{\partial \log \lambda}$	**0.176359** (0.005218)	**0.164938** (0.004874)	0.186588 (0.009023)	**0.168513** (0.004511)	**0.17723** (0.004212)	0.191212 (0.006943)
$\frac{\partial \log h}{\partial \log b}$+	**-0.42637** (0.006582)	**-0.39268** (0.006962)	**-0.39837** (0.01122)	**-0.5171** (0.008436)	**-0.47624** (0.00845)	-0.47736 (0.017721)
$\frac{\partial \log h}{\partial \log b}$++	-0.33625 (0.005784)	**-0.31537** (0.005924)	**-0.31009** (0.009677)	**-0.40524** (0.00692)	**-0.37009** (0.007253)	-0.34938 (0.012653)
$\frac{\partial \log h}{\partial \log \lambda}$++	**0.266484** (0.008392)	**0.24225** (0.007639)	**0.274873** (0.013838)	**0.280375** (0.008586)	**0.283379** (0.007522)	0.319195 (0.013652)
Sample Size	458	499	183	495	620	238

+ Pareto Assumption, ++ Exponential Assumption

Table 3.3: Job Search Elasticities and Wages for Non-Matched Men and Women by Training Type

rate was significantly different for men and women of all YTS experience types except women with a spell on YT.

3.4 An Application of the Nearest Neighbour Algorithm

It is at this point that we acknowledge that there may be a degree of self selection into YTS, which could cause a bias in the results thus far presented. In an attempt to adjust for this effect we apply the Nearest Neighbour matching algorithm to the dataset. For a motivation and description of matching in general and Nearest Neighbour methods in particular see section 2.4.

3.4.1 Heterogeneity of YTS Treatment Types in Matched Samples

Having obtained two full synthetic control groups for the male and female samples; each containing a single matched individual for every member of the YTS treatment group we re-estimated the elasticities for all men, all women, those with YTS experience and those without. Table 3.4 contains the computed elasticities for the new matched sample of males for both the YTS treatment and synthetic control groups. Base numbers differ to those of Table 3.2 where persons could not be matched with sufficient quality to te smaller treatment groups. In contrast to our non-matched male sample (see Table 3.2) there now appears to be no significant difference between the expected and reservation wages of those men in the YTS=0 control group and those with YTS experience (YTS=1).

The effects of a YTS spell on the search elasticities of men seem to follow much of the pattern we saw for the non-matched male sample. There are significant reductions in the elasticities of ξ, reservation wages, and h, the hazard, with respect to a change in b, the benefit level, under both the Pareto and Exponential wage offer distributional assumptions for men with YTS experience at the 5% level. The elasticity of h with respect to λ, the arrival rate of job offers under the Exponential assumption, is also significantly lower for males with YTS experience, although only at the 10% level.

The last three columns of Table 3.4 present the estimated elasticities for the female

Means	All Men	Men YTS=1	Men YTS=0	All Women	Women YTS=1	Women YTS=0
	Samples					
Expected Wage	126.8414 (0.908658)	126.3593 (1.19755)	127.3236 (1.367361)	114.2719 (0.693717)	**112.8773** (0.950637)	115.6665 (1.009437)
Reservation Wage	98.75319 (0.658696)	98.9926 (0.914834)	98.51378 (0.948336)	88.61614 (0.52228)	**86.93963** (0.707553)	90.29264 (0.765829)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	0.088191 (0.001838)	**0.084814** (0.002343)	0.091569 (0.00283)	0.107151 (0.002179)	**0.112029** (0.003233)	0.102274 (0.002916)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.177601 (0.002513)	0.175579 (0.003579)	0.179623 (0.003528)	0.175349 (0.002115)	**0.178741** (0.002944)	0.171957 (0.003035)
$\frac{\partial \log h}{\partial \log b}$ +	-0.4131 (0.0043)	**-0.40175** (0.004545)	-0.42445 (0.007285)	-0.4818 (0.005878)	-0.48407 (0.006098)	-0.47953 (0.010054)
$\frac{\partial \log h}{\partial \log b}$ ++	-0.32491 (0.003739)	**-0.31694** (0.003826)	-0.33288 (0.006416)	-0.37465 (0.005262)	-0.37204 (0.004751)	-0.37725 (0.009393)
$\frac{\partial \log h}{\partial \log \lambda}$ ++	0.265792 (0.004022)	**0.260393** (0.005615)	0.271192 (0.005759)	0.282501 (0.003792)	**0.29077** (0.00539)	0.274231 (0.005328)
Sample Size	2034	1017	1017	2552	1276	1276

+ Pareto Assumption, ++ Exponential Assumption

Table 3.4: Job Search Elasticities and Wages for Matched Men and Women

matched sample of the YTS treatment and control groups. As with the mean reported expected and reservation wages of the last three columns of Table 3.2 there is a significant fall (at the 5% level) in values for those women with YTS experience.

The elasticities of ξ and h (under the Exponential assumption) with respect to a change in λ are significantly higher for the YTS experienced women than for women without YTS experience. This is also true for the elasticity of the reservation wage with respect to a rise in the benefit level.

As with our analysis of Table 3.2 the effects of YTS participation which we may infer are coloured by the inherent difference between those with YTS experience and those without. As before it is likely that this difference will act to inflate the perceived effects of a spell on YTS. Therefore Tables 3.5 and 3.6 present these elasticities recalculated using the male and female matched datasets of Table 3.4 respectively.

Means	Men YTSI=1	Control Men YTSI=0	Men YTSII=1	Control Men YTSII=0	Men YT=1	Control Men YT=0
	Samples					
Expected Wage	122.8377 (2.098533)	124.9937 (1.984718)	127.1116 (1.540591)	129.9099 (2.189117)	130.5346 (3.269163)	125.14 (3.108129)
Reservation Wage	**95.92751** (1.523901)	98.6012 (1.462629)	101.4469 (1.261891)	99.45786 (1.444596)	98.36523 (2.396757)	96.13016 (2.28167)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	0.087658 (0.003931)	0.083478 (0.004099)	**0.079458** (0.003357)	0.094703 (0.004431)	0.092684 (0.005723)	0.097834 (0.00681)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.177005 (0.006259)	0.173917 (0.00569)	**0.168128** (0.004934)	0.182235 (0.005448)	0.190813 (0.008835)	0.183089 (0.007734)
$\frac{\partial \log h}{\partial \log b}$ +	-0.41283 (0.00727)	-0.41836 (0.013605)	**-0.38908** (0.006586)	-0.41945 (0.009832)	**-0.41301** (0.011487)	-0.44661 (0.016925)
$\frac{\partial \log h}{\partial \log b}$ ++	-0.32517 (0.00633)	-0.33488 (0.012831)	**-0.30962** (0.005186)	-0.32475 (0.007815)	**-0.32033** (0.010438)	-0.34877 (0.015512)
$\frac{\partial \log h}{\partial \log \lambda}$ ++	0.264663 (0.009778)	0.257395 (0.00898)	**0.247587** (0.007825)	0.276938 (0.008901)	0.283497 (0.013679)	0.280923 (0.013157)
Sample Size	340	340	476	476	201	201
+ Pareto Assumption, ++ Exponential Assumption						

Table 3.5: Job Search Elasticities and Wages for Matched Men by Training Type

The elasticities for the synthetic control groups of each variant of YTS are reported in the columns to the right of each YTS treatment group. Treatment effects are deemed to be significant if the YTS treatment group elasticities differ from the corresponding synthetic control group elasticities.

Unlike the non-matched male half of Table 3.3, where there was a marked treatment effect for young males who had participated in YTSI, Table 3.5 shows no significant differences between those with YTSI experience and those without. As with Table 3.3 results for men with a spell on YTSII indicate that there is a significant treatment effect which acts to lower all five of the elasticities under consideration here. The final two columns of Table 3.5 indicate a moderate YT treatment effect. Only the elasticities of the

Means	Women YTSI=1	Ctrl. Women YTSI=0	Women YTSII=1	Ctrl. Women YTSII=0	Women YT=1	Ctrl. Women YT=0
	Samples					
Expected Wage	**104.8057** (1.495201)	113.4125 (1.897633)	114.9681 (1.341568)	115.2291 (1.413484)	118.5152 (2.309596)	119.4951 (2.200196)
Reservation Wage	**81.90739** (1.071004)	89.60041 (1.538963)	89.61378 (1.003444)	89.07742 (1.051336)	**87.42957** (1.768356)	93.8539 (1.584971)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	0.108035 (0.005163)	0.104001 (0.005839)	0.101906 (0.003819)	0.10721 (0.004276)	**0.139443** (0.009613)	0.089185 (0.005088)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.171995 (0.005269)	0.166186 (0.005528)	0.174638 (0.004213)	0.177481 (0.004302)	**0.196365** (0.006427)	0.167097 (0.006715)
$\frac{\partial \log h}{\partial \log b}$ +	-0.50617 (0.008998)	-0.51265 (0.029877)	-0.4655 (0.008105)	-0.47879 (0.009737)	**-0.49699** (0.017047)	-0.4391 (0.010832)
$\frac{\partial \log h}{\partial \log b}$ ++	-0.39814 (0.007434)	-0.40865 (0.029167)	-0.36359 (0.006908)	-0.37158 (0.00796)	-0.35755 (0.011279)	-0.34991 (0.009131)
$\frac{\partial \log h}{\partial \log \lambda}$ ++	0.28003 (0.009556)	0.270187 (0.010054)	0.276544 (0.007378)	0.284691 (0.007657)	**0.335809** (0.012712)	0.256282 (0.010923)
Sample Size	362	362	629	629	285	285

+ Pareto Assumption, ++ Exponential Assumption

Table 3.6: Job Search Elasticities and Wages for Matched Women by Training Type

hazard with respect to the benefit level under both wage offer distributional assumptions are significantly different to the corresponding mean values for the synthetic control group.

Table 3.6 presents the female matched dataset YTS treatment effects by scheme type and takes the same form as Table 3.5. Results suggest that, other things being equal, women with experience of YTSI will on average have lower expected and reservation wages than those without. This result mirrors that found in Table 3.3. However unlike Table 3.3 these results do not support the hypothesis that participation in YTSI has an effect on any of the five elasticities. Table 3.6 indicates that, on average YTSII training has no effect on the wages or elasticities of women who enter into it. Whereas Table 3.3 contained only one significant treatment effect for women on the third scheme, YT, Table 3.6 contains results which point to a treatment effect for many of the indicators under consideration. The mean value for the reservation wage of YT participants is significantly lower than that of the synthetic control group. The elasticities of ξ and h (under the Pareto wage offer distributional assumption) with respect to a change in the benefit level are significantly higher for those with YT experience. The elasticities of ξ and h (under the Exponential wage offer distributional assumption) with respect to λ are also higher for YT participants.

3.5 Context and Implications of Results

In this chapter we have presented evidence indicating the presence of some YTS treatment effects on wage expectations and search elasticities. Section 3.3 Table 3.2 suggested the existence of a marked change in the mean values for the structural parameters and expected/reservation wages of individuals who had undertaken a period of YTS training. However human capital theory would suggest that there may well be some fundamental differences between those who sought YTS training and those who had no experience of such schemes.

Our results support the heterogeneity of YTS treatment types. We allowed for the schemes' evolution by subdividing our male and female samples into the three groups of YTS (YTSI, YTSII and YT). We then re-estimated the structural parameters and expected/reservation wages for the non-matched dataset. Table 3.3 suggested that YTS participants of both sexes could expect to experience large treatment effects for all three YTS variants; the early scheme(YTSI) having a largely detrimental effect. However the estimates by scheme type for the matched datasets presented in Tables 3.5 and 3.6 cast a shadow over the results of Table 3.3. They offer little evidence of a YTSI effect on the structural parameters and some evidence for a reduction in expected/reservation wages following a spell on YTSI. YTSII appeared to have a marked effect on male participants but there was no evidence to support the hypothesis of a female treatment effect. Evidence for the third variation, YT, suggested that female participants experienced a marked treatment effect for both their structural parameters and reservation wages. Males who undertook a period of YT training exhibited an effect only through a drop in the magnitude of the elasticity of the hazard of exiting unemployment, h (under both the Pareto and Exponential wage offer distributional assumptions), with respect to a change in the benefit level, b. The explanation of these results must lie in labour market factors. An examination of the economic climate over the period of YTS evolution provides some clues for the explanation.

To understand the implications of these results we need to examine the labour market conditions experienced by different generations of government training recipients and compare our unconditional estimates with the conditional ones which may be estimated controlling for regressors. Table C.5 contains regression model estimations for reservation and expected wages by gender for both control and treatment individuals. Notice the fact that all three YTS types have a positive effect on both wage types for men whilst only YT has any effect on the wages for women. This result would indicate that experience of YT/YTS actually increases wage expectations. Presumably because YTS recipients would feel that their human capital had been enhanced. Training allowances and the average wage rates for under 18's also have a significant role in determining reservation and expected wages. Both youth and Local Education Authority (LEA) unemployment[6] are negatively significant for both sexes, indicating that adverse local labour market conditions can lower expected and reservation wages. Women who live with their parents have much lower wage levels. Perhaps due to the fact that living costs are lower whilst they remain at home and so jobs offering lower levels of pay become acceptable to these women.

Appendix Figures C.5 and C.6 contain plots of various economic indicators over the time span of interest, some personal to the persons being studied, some regional and some national. In 1986 when cohort 1 sweep 2 and cohort 2 sweep 1 people were questioned and YTS took the form of YTSI the unemployment rate (which has been scaled to

[6]The unemployment rate for the area in which an individual finished his/her schooling.

overlay the wage data) was high (14%). After the introduction of YTSII unemployment fell steadily until 1990 when it bottomed out (8%). As the third YTS variant YT was introduced unemployment was again on the rise and had climbed back up to around 11% by 1994. This trend coupled with the result from Table C.5 that unemployment rises led to falls in reservation and expected wages[7] which might account for the YTS treatment effects uncovered in Tables 3.5 and 3.6 if the unemployment rate affected those with YTS experience differently to our control group.

Male expected wages followed actual average male wages closely for the whole period of study. It is interesting to note that the male school leavers in our sample returned expected wage levels which were above the average rate of pay for males under 18 in the years 1986 and 1994. Female under 18's expected wages were much lower than average female wages for most of the period of interest. Young women appear to have expectations for wages below the female average. Minimum training allowances were all but frozen for all six cohorts and as a result the gap between reservation wages and training allowances had widened considerably by 1994.

An important respect in which young men and young women are different is that over the 1980s and 1990s girls in school were out performing boys in terms of educational scores[8]. Our results may be consistent with the view that although human capital has been rising among young women their labour market expectations have not risen accordingly.

Another explanation for the differences which we observed between the treatment effects for men and women is the existence of gender discrimination by employers. This long recognised phenomenon could be causing female job seekers to revise down their wage expectations in the light of poor wage offer distributions. After the introduction of YT and as the unemployment rate rose it appears that women with YT experience revised downward their expectations and their search elasticities. This indicates that for

[7]Remember that reservation and expected wages are scaled by changes in the Retail Price Index (RPI) relative to 1998 prices.

[8]See Dolton *et al.* (1999) who examine the average GCSE exam score difference between boys and girls.

female school leavers the later YTS schemes had the effect which the government sought. The fact that there was not a similar effect for their young male counterparts (except for the elasticity of h with respect to b) might well be due to the high proportion of women (relative to men) who enter the service and leisure sectors, which were growth areas in the mid 1990s.

3.6 Conclusions

There are two principal empirical findings of this Chapter. Firstly, that the different types of youth training have distinct effects on expected and reservation wages and search elasticities. Secondly, that the effect of government training on males and females seems to be radically different. The effect on the former is to enhance optimistic expectations for men, but modify those of women. The implications of these findings are that in empirical analysis of the labour market outcomes of young people it is necessary to consider men separately from women. In addition it would also appear to be wrong to categorise all government training as the same[9].

On a wider methodological level our estimation strategy also suggested that erroneous conclusions could be reached if non-comparable groups, those receiving treatment and those not receiving treatment, had been compared. Our approach of using a matched sample suggests that one may reach less dramatic comparative conclusions, but ones we may have confidence in. We suggest that the matching methodology has wide applications.

However, there are a number of performance issues relating to the nearest neighbour algorithm. In the next chapter we will investigate the performance of the Nearest Neighbour algorithm as used above. This investigation will lead us to consider alternate matching solutions, methods whose performance will also be discussed.

[9]The much of the literature has treated YTS as a single unchanging entity. See Mealli *et al.* (1996)

Chapter 4

Matching Estimator Performance and its Bearing on Our Conclusions

As we were able to demonstrate in section 2.4.3 the performance of any given matching algorithm is dependent on the validity of the assumptions, which underpin it. In this chapter we investigate the performance of the differing matching algorithms, which we have employed and attempt to assess the validity of the conclusions we might draw from subsequent statistical tests on our matched datasets. The concept of Stochastic dominance is used to assess the region of overlapping support. Several plots are used to highlight the weaknesses which are inherent within traditional matching methods.

4.1 Nearest Neighbour Matching Protocol Performance

The "Nearest Neighbour" matching protocol described in subsection 2.5.1 matches an individual from the control group to a person from the treatment group by selecting the control whose propensity score is closest to that of the said treatment person. The control is then removed from the population of unmatched controls and is unavailable for matching to a subsequent treatment person. This method has the advantage that it matches real people to each other facilitating any further investigation we may wish to perform[1].

Such an approach can be expected to produce close matches for as long as the propensity score distributions take a form similar to that of Figure 4.1 Case 1. Here we depict a pair of stylised propensity score distributions. The yellow distribution represents the propensity score distribution for the treatment population, whilst blue represents that of the control population. Notice how the blue distribution completely envelops the yellow distribution. This is a highly desirable property for data, which we propose to perform a "Nearest Neighbour" matching procedure on as it follows from this that for any treatment individual there exists a control whose propensity score is similar to theirs. However, if the propensity score distributions of the treatment and control populations take the form of Figure 4.1 Case 2, we can expect the algorithm to produce some poor matches with respect to propensity score.

Area A represents the portion of the treatment propensity score distribution, which is not enveloped by the blue of the control group. At some point during a run of the "Nearest Neighbour" algorithm all the controls with a propensity score, which lies within Area B, will have been matched and removed from the unmatched control group. However there will still exist treatment individuals who possess a propensity score within this range. These treatment persons must then be matched to individuals from the remaining unmatched control group who reside to the left of Area B and therefore possess propensity scores, which are poor matches. That is to say that when we match on data which has the form of Figure 4.1 Case 2, we can expect that as successive treatment persons are matched to controls the remaining unmatched control population will look less and less like the remaining unmatched treatment population.

Having described the "Nearest Neighbour" matching algorithm in subsection 2.5.1, we then performed it in section 3.4 for males and females of the YCS dataset, cohorts 1-6. The extent to which the conclusions that we drew are valid depends on the degree to which

[1]Since our matched controls are actual people and not synthesised, they retain their full set of covariates post match.

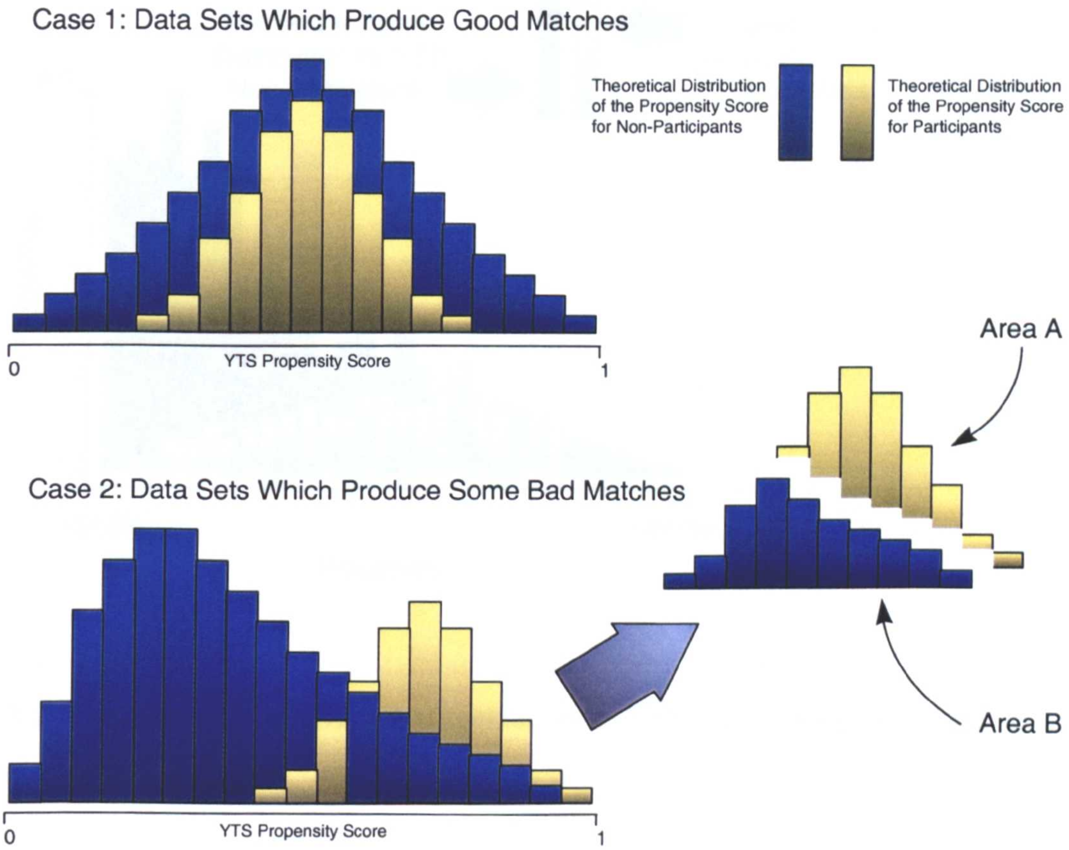

Figure 4.1: Theoretical "Nearest Neighbour" Matching Algorithm Performance

the propensity score distributions of the YTS participants and non-participants overlap. Figure 4.2 contains the actual propensity score distributions for the male sub-populations. Notice how the YTS non-participants propensity distribution (blue) envelops all of the left side of the YTS participants' propensity distribution (yellow). However, the right tail contains a number of YTS participants for whom there will be no close matches. This is true whenever the blue columns are smaller than the corresponding yellow columns.

In Figure 4.3 each pair of points represents the average distance between the propensity scores of a matched couple for every twenty matches. An examination of Figure 4.3 reveals how the effect of this lack of high propensity control group persons causes the mean distance between matched propensity scores to widen as the algorithm proceeds. Notice how the algorithm produces consistently good matches until those of the 43rd

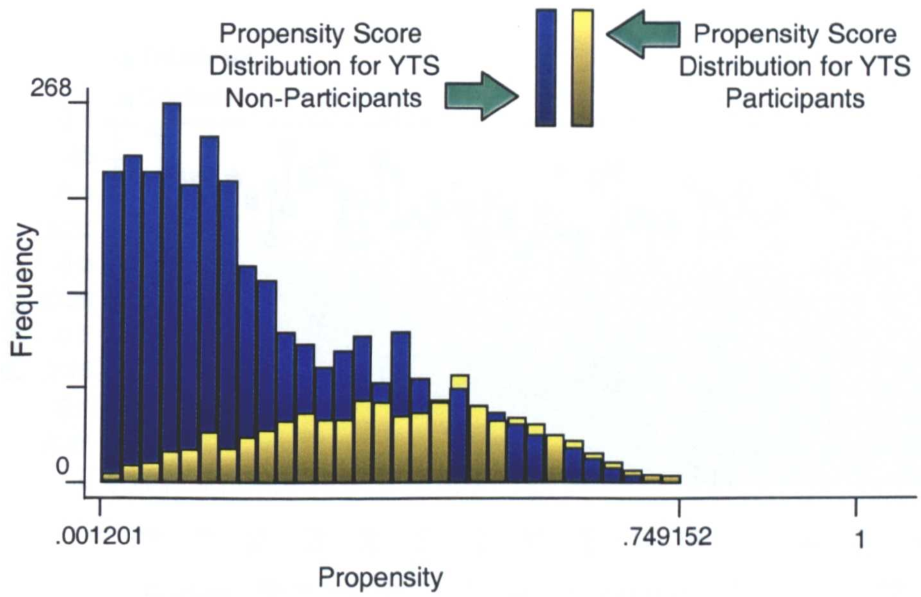

Figure 4.2: Propensity Score Distributions for both YTS Participants & Non-Participants (Males)

average distance point, after which treatment persons are matched to controls of lesser propensity.

Any conclusions drawn from analysis of this matched dataset will be tainted by the existence of these poorer matches within it.

One solution, which the literature suggests for the above problem, is to drop any treatment individual from the dataset for whom the algorithm fails to find a match of sufficient quality. Cochran and Rubin (1973) suggest a version of "Nearest Neighbour" caliper matching which makes a match to person i only if there exists a person j, from the control group such that,

$$\|P(X_i) - P(X_j)\| < \epsilon \tag{4.1}$$

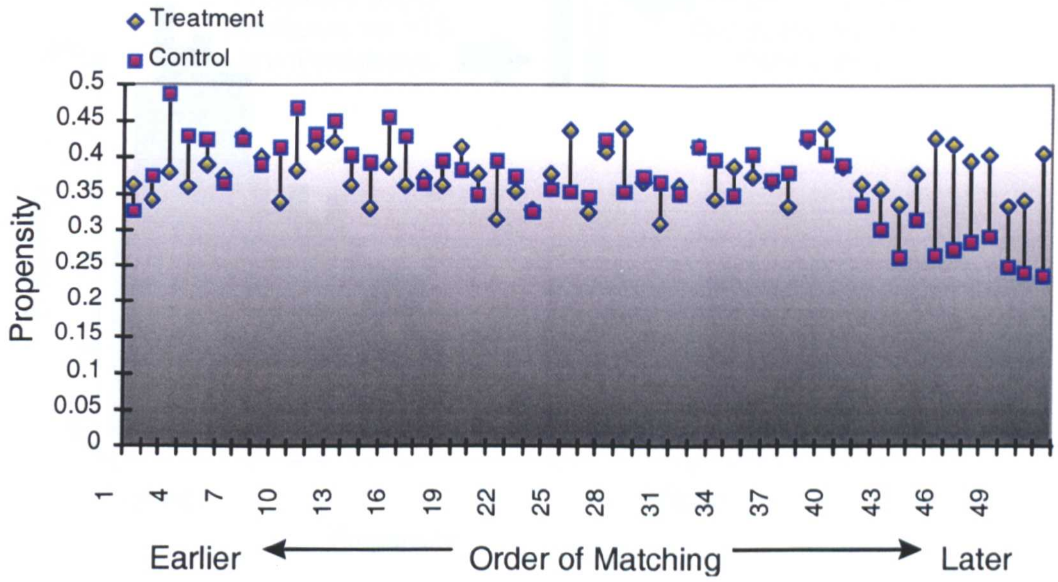

Figure 4.3: Matching Performance from First Match to Final Match (Males)

where ϵ represents the caliper boundary. In fact our version of caliper matching could be easily adapted to perform a more advanced variation of this method if we only made a match to person i when there was some control person j, whose propensity score lay within the caliper of person i, I.e:

$$P(X_j) = v_j\hat{\beta} \in \left[v_{n_t}\hat{\beta} \pm C\sqrt{var(v_{n_t}\hat{\beta})}\right] \tag{4.2}$$

See Appendix C.1 for a description of the full matching protocol we used[2]. There is however a major flaw in the caliper rejection approach in that it removes persons from the treatment group non-randomly. Under such a scheme we would lose people from the right tail of the treatment group propensity score distribution, causing us to generate a biased matched sample containing too few high propensity treatment-control matches. In our attempt to remove one potential bias we would have introduced another. This point is of crucial importance when an investigation is concerned with something such as

[2]Based on Lechner (1999)

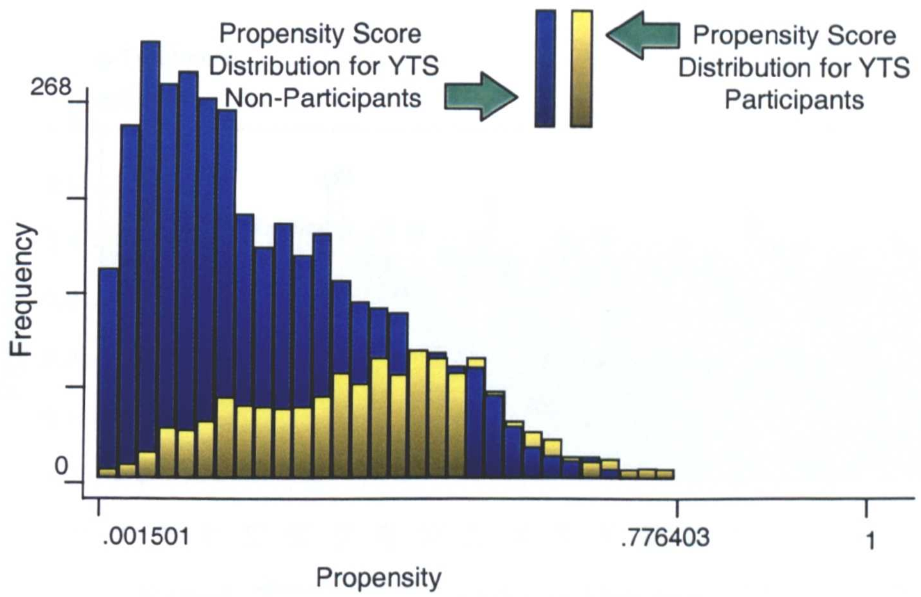

Figure 4.4: Propensity Score Distributions for both YTS Participants & Non-Participants (Females)

measurement of the YTS treatment effect. People with high YTS propensities are likely to be part of the self-selection phenomenon. Indeed, it is these people who the scheme was designed to help. As such they must remain in the dataset for us to accurately evaluate the YTS treatment effect. Previous studies often overlook this fact when choosing to drop poorly matched persons. Although, the recently released[3] STATA 7 .do program "psmatch", automatically excludes those persons whose propensity scores lie beyond the region of support when performing it's version of matching it is possible to override this feature.

Figures 4.4 and 4.5 depict the female versions of the plots presented in figures 4.2 and 4.3. As with their male counterparts we again see a shortage of high propensity control persons with which to match. This leads to a reduction in match quality after the forty ninth-point average. Too many poor quality matches for high propensity treatment

[3]Barbara Sianesi, University College London and Institute for Fiscal Studies, June 2001

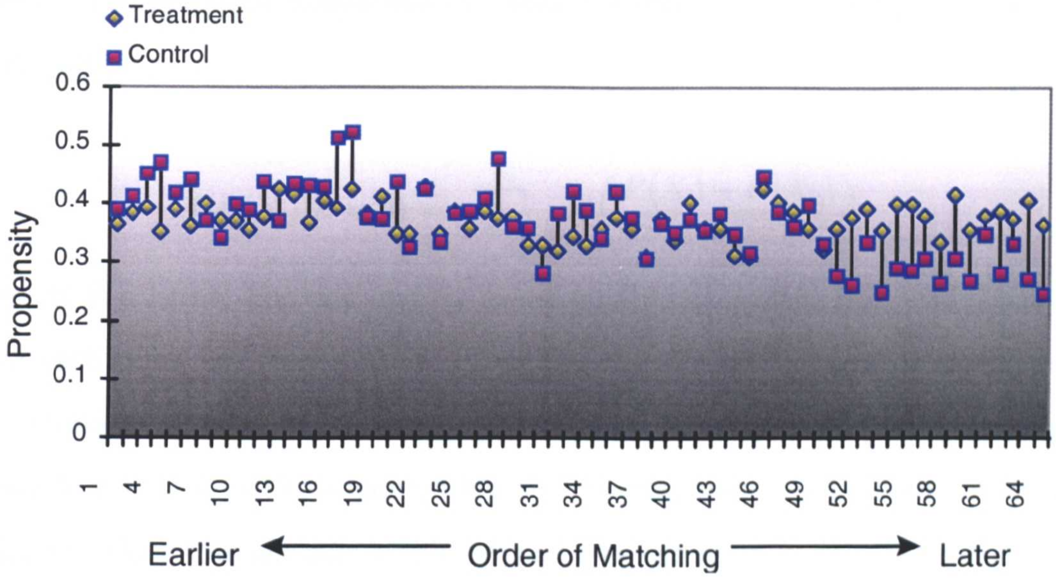

Figure 4.5: Matching Performance from First Match to Final Match (Females)

individuals may therefore bias conclusions we draw from work on the female matched sample.

4.2 Minimising the Number of Poor Matches

4.2.1 Overlap of the Propensity Score Distributions

Ideally we would wish to possess a method of calculating some statistic, which we could use to judge whether a given treatment and control dataset would produce good matches when employing the method of nearest neighbour matching. Todd (1999) presents a method of discriminating between data, which will produce potential good and poor matches. She terms the support of $P(X)$ for which both $f_x(P(X)|D = 1) > 0$ and $f_x(P(X)|D = 0) > 0$ as the "region of overlapping support". She also describes a method for identifying the observations that lie in the region in which the density $f(P(X_i)|D = 0)$ (using $D = 0$

control group data) at each of the $P(X_i)$ values observed for $D_i = 1$ observations is calculated. The standard nonparametric density estimator used is reproduced here in equation (4.3).

$$f(P(X_i)|D=0) = \sum_{k=1,D_k=0}^{n_0} K\left(\frac{P(X_i) - P(X_k)}{h_n}\right) \tag{4.3}$$

where K is a kernel function and h_n is the bandwidth parameter. The density estimates at each point are then ranked. From this the 1% or 2% quantile of the positive density estimates is obtained. A value of $P(X_i)$ with an estimated density greater than this quantile is considered to reside within the "overlapping region of support". Values of $P(X_i)$ less than this quantile reside beyond the region of support and Todd (1999) recommends they be excluded in estimation.

The weakness of this method is that we are again rejecting persons from our sample in a non-random manner. As we have already discussed it seems likely that for our dataset the region of non-overlapping support will lie to the right of the propensity score distributions and as such a method such as we have just recounted would reject mainly high propensity YTS persons from our sample and hence bias any conclusions which we attempt to draw from analyses of the matched dataset. Furthermore it is not unreasonable to assume that these high propensity YTS individuals are the group who require a scheme such as YTS to help them develop their employment prospects and as such it is the effect of the scheme on these persons, which we should be most interested in. Clearly rejecting persons of this type is not desirable given our aim is to uncover the YTS treatment effect on the treated.

Todd (1999) does recognise that the majority of the treatment group could be outside of the region of overlap, if the model for participation predicted unusually well. The solution proposed is to re-estimate the propensity scores using an alternative set of X variables. This solution seems to rely on using a poorer model for YTS participation to increase the region of overlap. As a result this solution is flawed.

Having stated our reservations as to the validity of any method which proposes to remove persons from the treatment group for whom there are insufficient controls pos-

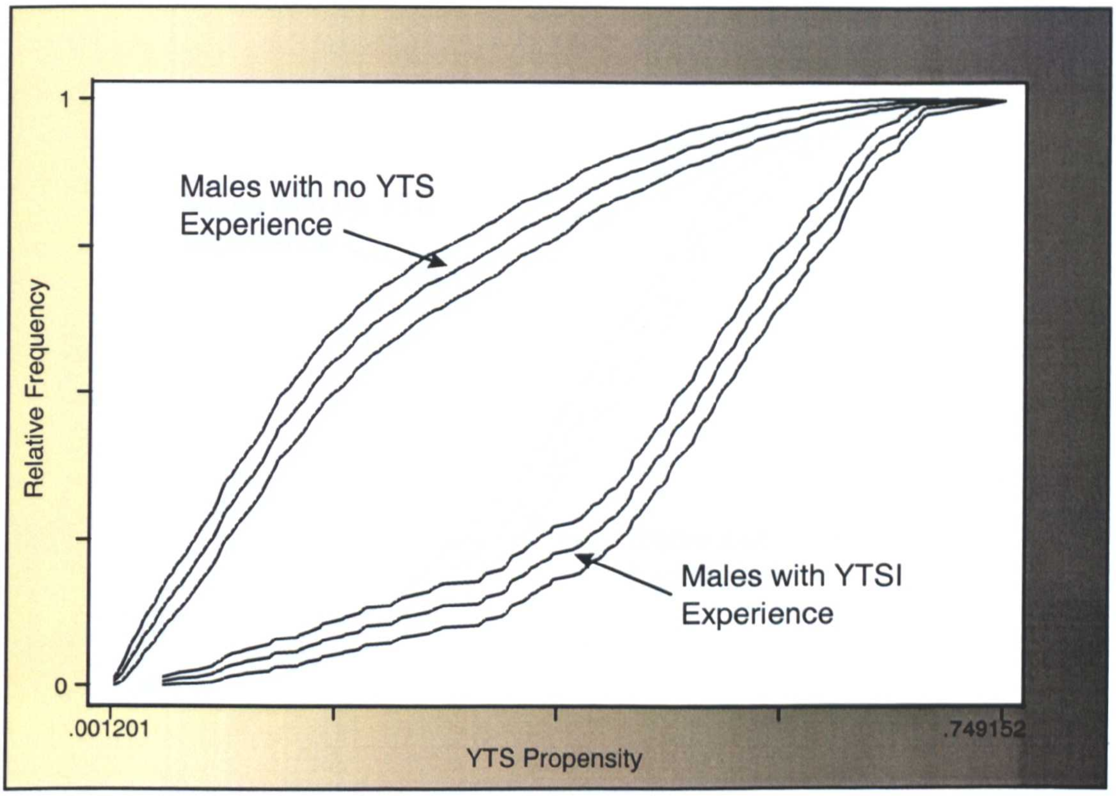

Figure 4.6: Cumulative Distributions of Males with YTSI Experience and those without any YTS scheme experience

sessing a similar YTS scheme propensity before proceeding to match on the remaining dataset, we now suggest that it is as valid to retain all treatment persons and adjust our conclusions from analyses carried out after matching on this complete dataset using the knowledge that there were some poor matches rather than reject treatment persons non-randomly and make no allowance post match for the bias this has induced.

One way in which we may hope to gain an understanding of the degree of overlap in the YTS propensity score distributions is to compare the cumulative distributions of the propensity scores for the control group and the various YTS scheme participants. It is recommended that the reader become acquainted with the notion of stochastic dominance before proceeding any further. See section 2.7 for an overview of the concept and it's relevance to the problem we now face.

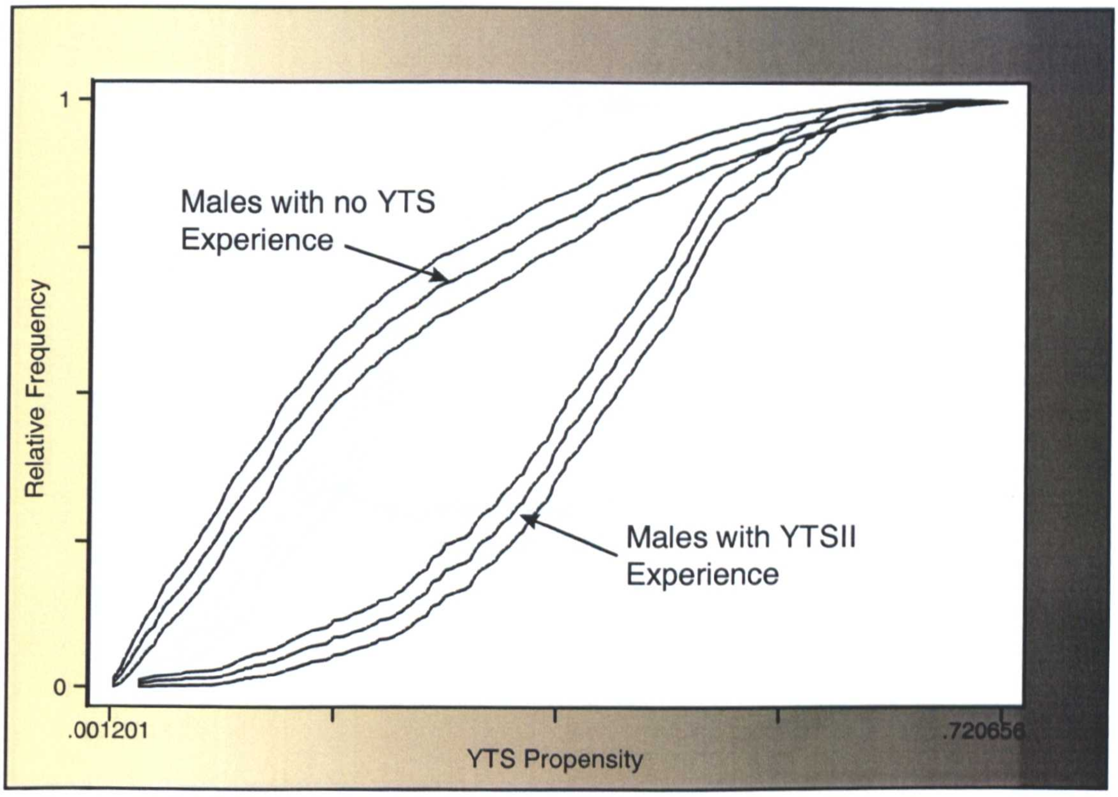

Figure 4.7: Cumulative Distributions of Males with YTSII Experience and those without any YTS scheme experience

If the CDF of the propensity score for YTS participants, $F(P(X_1))$, dominates that of non-participants, $G(P(X_0))$, according to second order stochastic dominance and the control group is not sufficiently large enough to compensate, then we would expect that nearest neighbour matching, performed using these scores, would produce some poor matches for those treatment persons with high YTS propensity scores. These high propensity individuals are the persons whose outcomes we are most eager to examine. The scale of the problem would be directly related to the "degree" of stochastic dominance and the sample size shortfall.

Figure 4.6 shows the relative CDFs and their respective 5% confidence regions for the male control group and males with YTSI experience. Notice how the two CDFs never cross. This indicates that the propensity score distribution for the control group stochas-

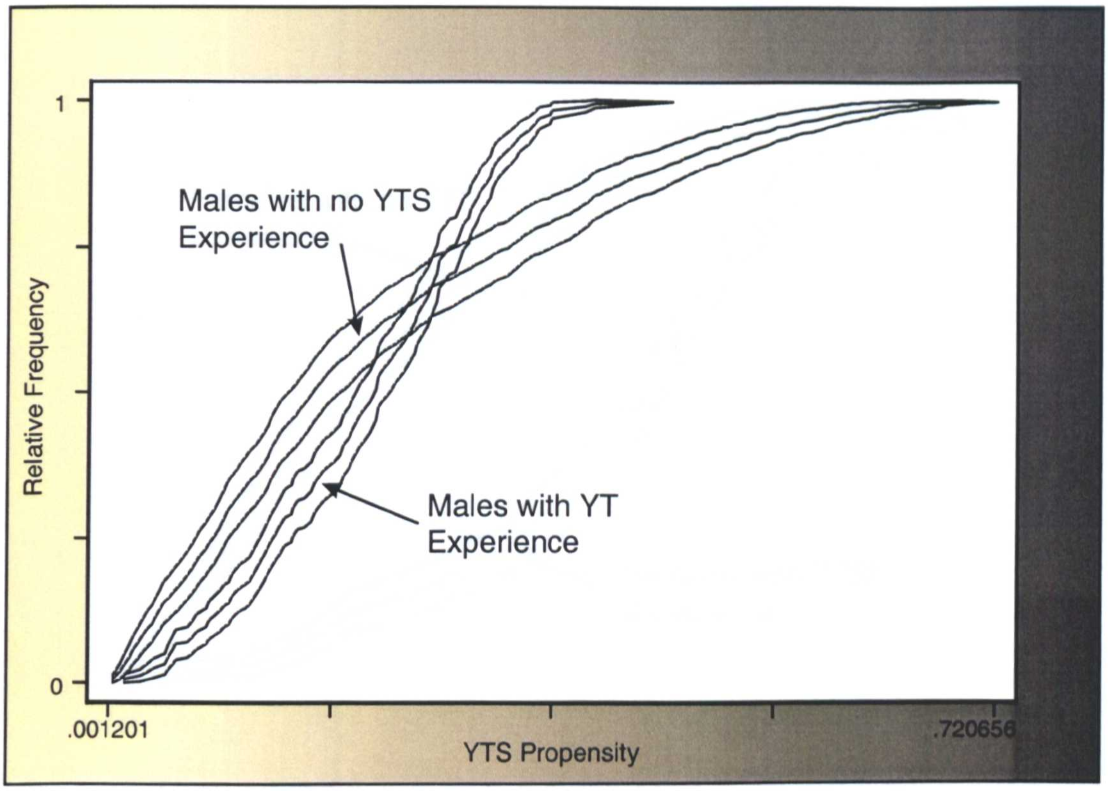

Figure 4.8: Cumulative Distributions of Males with YT Experience and those without any YTS scheme experience

tically dominates that of the male YTSI participants; the control group's distribution lies to the left of the YTSI group's distribution. Figure 4.7 contains a similar result for males with YTSII experience. Hence, it follows that matches produced using the YTSI and YTSII male datasets with the nearest neighbour algorithm may produce some poor matches for treatment individuals, as there are insufficient controls of high YTS propensity when the treatment and control groups are of equal size. Our pre-match control group is larger than the pre-match treatment group, but is this sufficient to overcome the problem?

Figure 4.8 contains the CDF of males with a spell on YT against that of those without any kind of YTS experience. Here we see that the two CDFs cross. This indicates that there are sufficient control persons of similar YTS propensity to produce a matched dataset, which will not suffer from degradation in match quality as the algorithm pro-

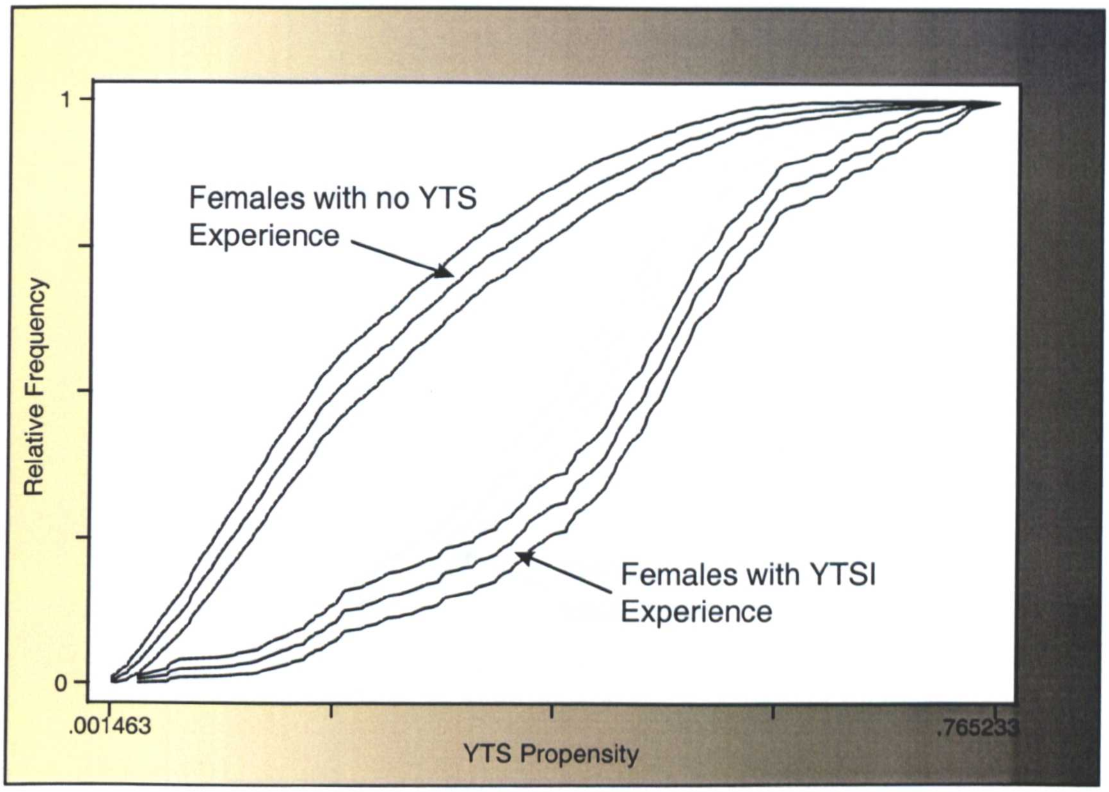

Figure 4.9: Cumulative Distributions of Females with YTSI Experience and those without any YTS scheme experience

gresses.

Figures 4.9, 4.10 and 4.11 contain CDFs for the various female YTS sub-groupings. As with the male samples the females with YT experience sub-grouping is the only dataset for which there are sufficient female controls to match, without degradation, using nearest neighbour methods and datasets of equal size. Again the question of whether the female control group is sufficiently large to offset this problem arises.

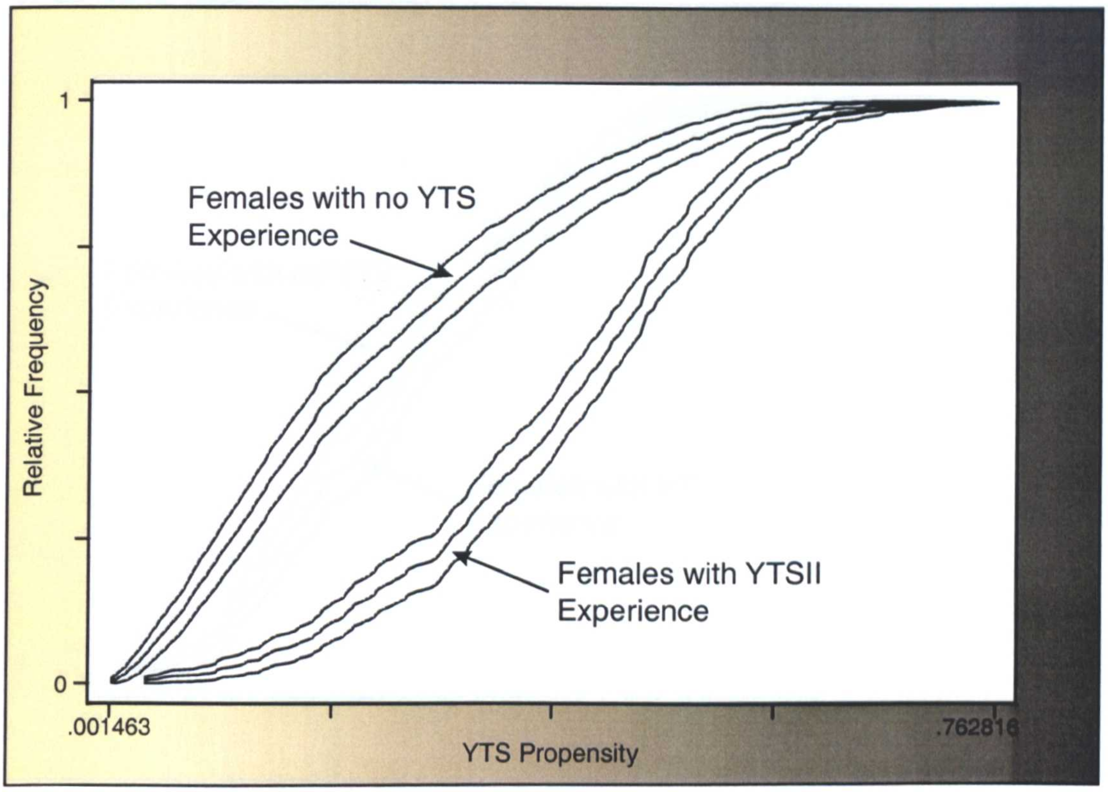

Figure 4.10: Cumulative Distributions of Females with YTSII Experience and those without any YTS scheme experience

4.3 Implementation of Kernel and Local Linear Matching Methods

Having uncovered a number of issues during our investigations concerning the performance of our nearest neighbour matching algorithm we now examine kernel based matching methods. To see a motivation for matching methods as well as an explanation of these methods in particular see sections 2.5 and 2.4. Since these methods require one to select the bandwidth parameter before attempting to match, it follows that their performance may be sensitive to this choice. As such this section will present our early kernel and local linear matching method results in terms of their performance and the ramifications for our understanding of the YTS treatment effect.

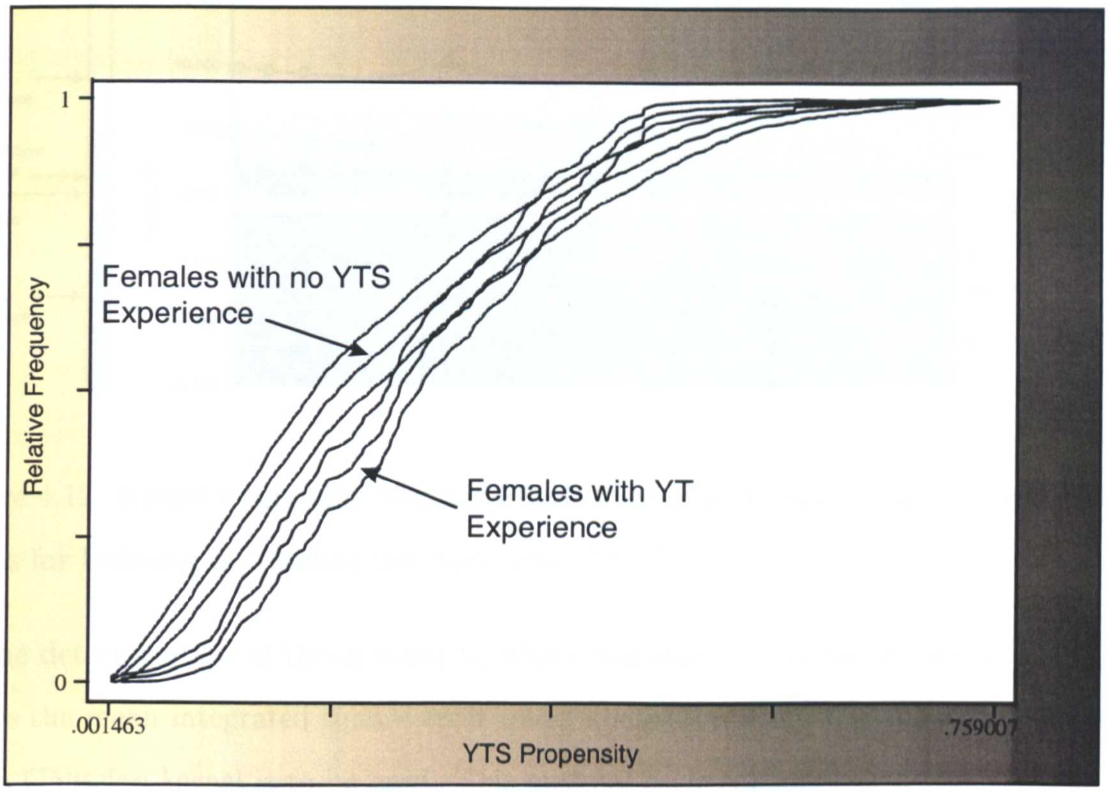

Figure 4.11: Cumulative Distributions of Females with YT Experience and those without any YTS scheme experience

4.3.1 Kernel Regression Matching Performance

The main strength of the kernel regression matching estimator which we described in subsection 2.5.2 is that it makes use of all available control group persons (from $j = 1, \ldots, n_c$) and weights them in accordance with their similarity to the treatment person i, as measured by their propensity scores. Unlike the "nearest neighbour" method, the pool of control individuals does not erode as the algorithm progressively creates matches for the treatment persons from $i = 1$ to n_t.

During our explanation of the kernel regression matching estimator we introduced the concept of the bandwidth parameter h_n, which the user of the algorithm must select for themselves and which can influence the accuracy of the subsequent matches produced. Many people have proposed methods for choosing h_n. Silverman (1986) presents a method

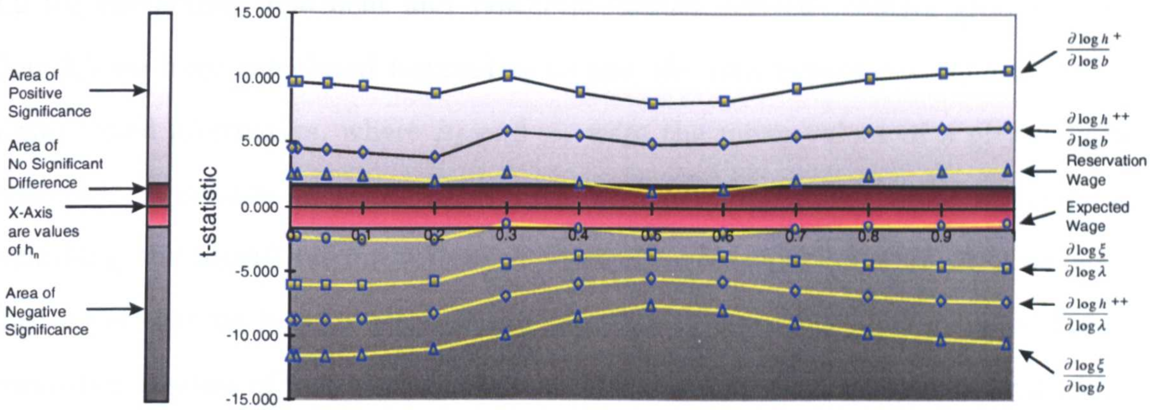

Figure 4.12: Kernel Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Males) See Appendix C.2

for the determination of the optimal h_n where you select h_n to be the width which minimises the mean integrated square error under the assumption that the data is Gaussian and a Gaussian kernel is to be used. This optimal h_n is calculated via

$$h_n = 1.06 s n_t^{-1/5} \tag{4.4}$$

a more robust version of which is

$$= \min\left(s, \frac{IQR}{1.34}\right) n_t^{-1/5} \tag{4.5}$$

However due to the assumptions, such an "optimal" value for h_n is not to be considered as a global solution. For multi-modal highly skewed densities, this width is usually too wide and can over smooth the density. Also we have chosen to use the biweight kernel, as such the above assumption does not hold. Despite this, Silverman (1986) method still represents a valuable tool when attempting to estimate a value for h_n.

Before attempting to use a plug in optimal h_n we present t-test results of significant treatment effects for any form of YTS schemes by gender without regard for YTS heterogeneity, for various h_n. This specification mirrors that of the early investigations in chapter 3. These include estimates of the search elasticities as described in chapter

2.2 for the actual treatment and synthetic (kernel derived) control groups. As in section 3.3 we have calculated t-statistics to test the null hypothesis $H_0 : \mu_1 = \mu_2$ verses a two tailed alternative, where μ_1 and μ_2 were the mean values of a given elasticity or expected/reservation wage for those with YTS experience and those without. It is by rerunning the algorithm for various h_n and then carrying out our tests on wages and elasticities that we hope to capture the variation in matching performance. This process generated a series of t-statistics. Figure C.1 compares kernel regression matched males with and without any form of YTS experience. In it we present a plot of these t-statistics as h_n was varied between 0 and 1. The x-axis represents the variation in h_n from 0 to 1. The y-axis depicts the t-statistic values for the differences between treatment and synthetic controls of wages/elasticities produced using the values of h_n. Therefore, the yellow plot lines represent the movement of these t-statistics for each wage/elasticity pair as h_n increases. The wages/elasticities to which the yellow lines belong are indicated to the right of the figure. The red area represents the area of no significant difference between a synthetic control person and their treatment counterpart. Hence yellow plots, which reside beyond this region, indicate the presence of a treatment effect. Those plots, which move from the area of significant difference to that of no significant difference or vis-versa, are difficult to interpret. Even those plots which remain significant over the range of h_n may experience a change in the magnitude of the significance as h_n increases. Hence the size of the treatment effect can be hard to discern.

Notice that both the treatment effects on the reservation and expected wages move in and out of significance as h_n increases. Overall though it is interesting to note that none of the plots deviates alarmingly over the range of h_n. Hence if one was to take a value of h_n and from it generate a matched sample from which one could create a table such as Table 3.4, section 2.5 which was produced using the nearest neighbour algorithm, one would draw broadly similar conclusions to those of Table 3.4. If you recall, the expected and reservation wage effects on this male sub-sample were shown to be insignificant after nearest neighbour matching, whilst the elasticities of the reservation wage and hazard

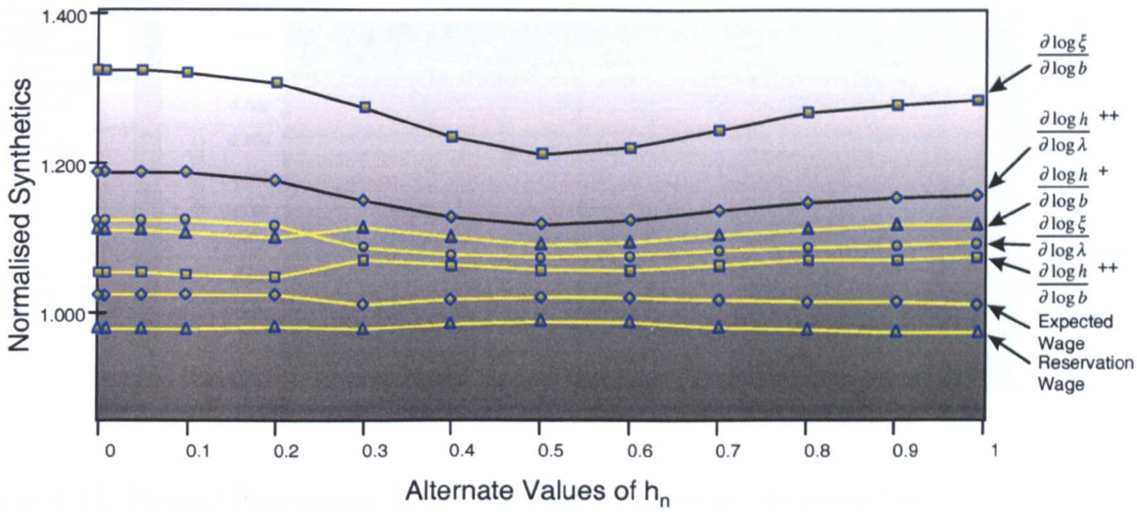

Figure 4.13: Kernel Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Males) See Appendix C.2

with respect to a change in the benefit level were shown to exhibit an effect in line with that which we see in Figure C.1 for all values of h_n after kernel regression matching. The elasticity of the hazard with respect to a change in the arrival rate of job offers (exponential wage offer distributional assumption) also moves in line with the nearest neighbour result for males. These results are evidence of the robustness of our results to differing matching methods.

Figure C.2 presents the way in which the magnitudes of the differences between treatment and synthetic wages/elasticities vary over h_n for the same male sub-sample as Figure C.1 using kernel regression matching. Here we have standardised the treatment wages/elasticities to 1. Hence, a synthetic control wage/elasticity with a value greater than 1 indicates that the wage/elasticity increased after treatment. Once gain there are no large fluctuations in any of the male plots over the range of h_n.

Figure 4.14 contains the female sub-sample counterpart to Figure C.1. The effects of all variations of YTS combined on this female sub-sample would appear to be far more

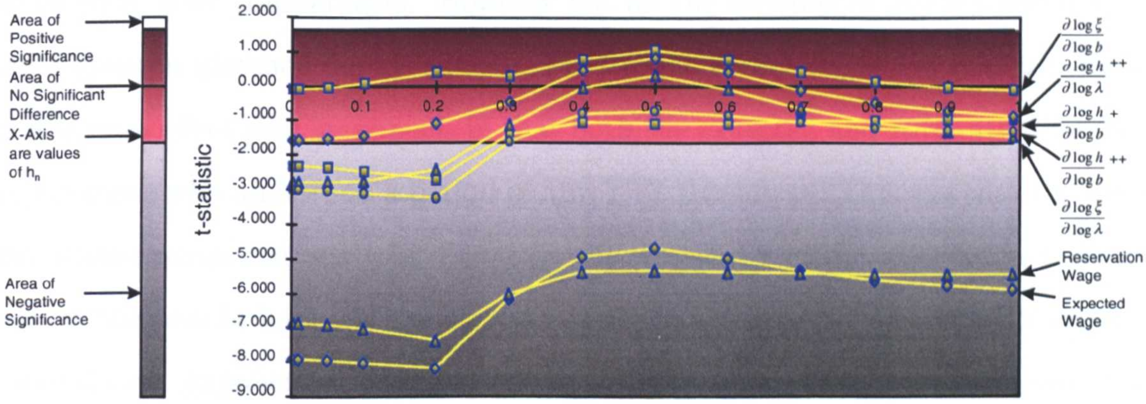

Figure 4.14: Kernel Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Females) See Appendix C.2

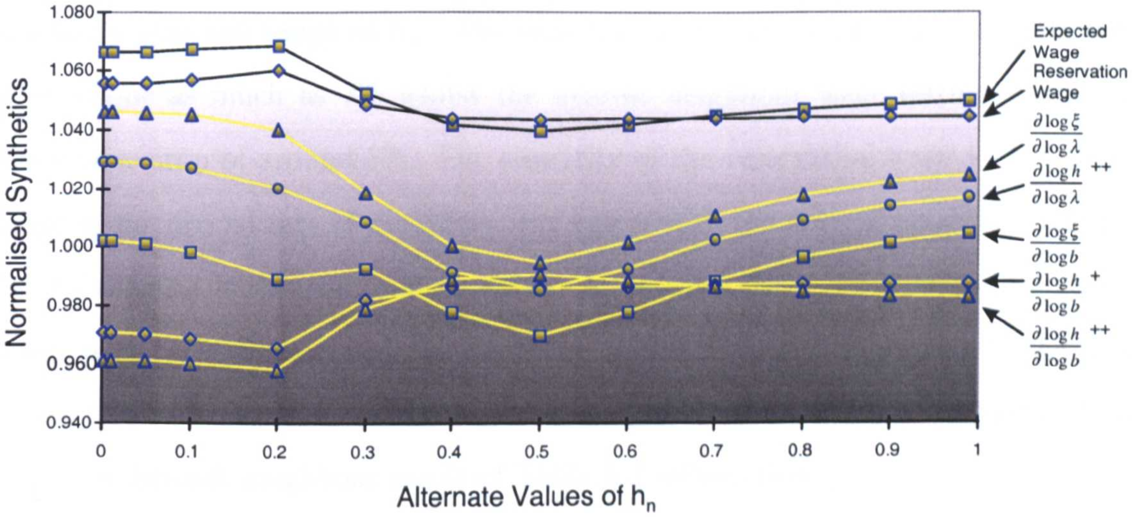

Figure 4.15: Kernel Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Females) See Appendix C.2

sensitive to our choice of h_n than was the case for the male sub-sample. A number of the elasticities move into the region of no significance for values of $h_n > 0.2$. Again a comparison with the findings of Table 3.4 is of interest.

Our nearest neighbour matched females produced a set of results, which differed from those of their male counterparts. However not all the differences are replicated by the kernel regression matched female sample. Firstly, the kernel regression matched females exhibit a wage effect post treatment, in that both the expected and reservation wages are lower for those who undertook a period of any YTS treatment. This was also the finding of the nearest neighbour analysis. However, the elasticities of the reservation wage with respect to the benefit level and the arrival rate of job offers or the hazard with respect to the arrival rate (exponential wage offer distributional assumption) failed to move in line with the nearest neighbour results. A look at the results for the local linear regression matched female sub-sample may help to reconcile these findings. As with Figure C.2, Figure 4.15 highlights the need to uncover some evidence as to what the correct value of h_n should be; the magnitude and significance of the measured treatment effects vary substantially over the range of h_n. The expected and reservation wages are effected by as little 4% or as much as 7% whilst the nearest neighbour wage estimate pairs both experienced a drop of around 3%. The elasticity of the reservation wage with respect to a change in the arrival rate of job offers also agrees with the nearest neighbour algorithm result of Table 3.4 with a percentage change of between 0% – 5% over the range of h_n. However, as seen in Figure 4.14 the elasticity of the hazard with respect to the benefit level under both wage offer distributional assumptions also indicates a treatment effect although the nearest neighbour result of Table 3.4 refutes this.

4.3.2 Local Linear Regression Matching Performance

As with the kernel regression matching estimator, the local linear regression matching estimator weights everyone in the non-participating sample by their similarity to each treatment person. However the local linear estimator has a number of advantages over the previous method. Subsection 2.5.3 provides a description of these advantages.

Figure 4.16 contains the t-test results for the same male sub-sample as that used for Figure C.1. We performed local linear regression matching for various values of h_n. As with our analysis of the kernel density matching results a comparison with the nearest

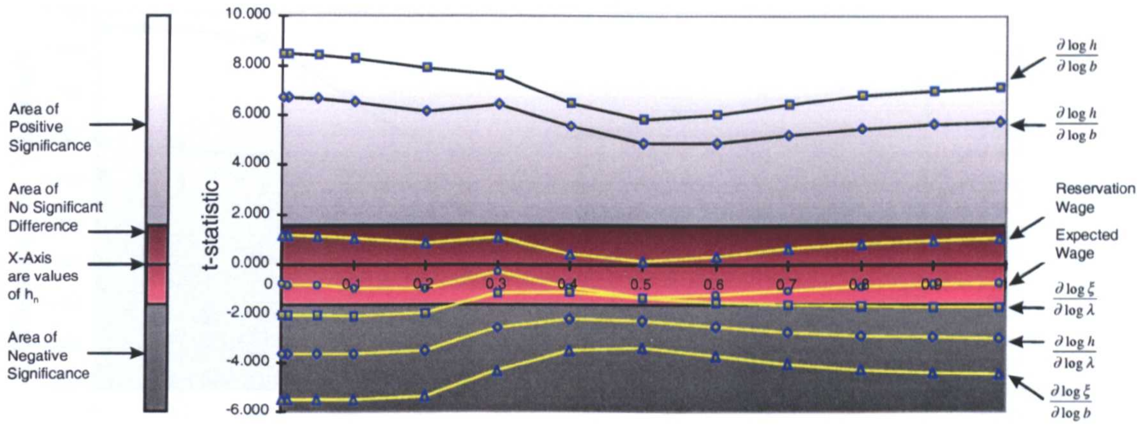

Figure 4.16: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Males) See Appendix C.2

neighbour matching results of Table 3.4 is of interest.

Notice how for all values of h_n the local linear method revealed no significant treatment effects on the reservation and expected wages of males. This result mirrors that of Table 3.4 in which we saw no significant wage effects for men who experienced any form of YTS. The elasticities of the hazard with respect to the benefit level under both the Pareto and Exponential wage offer distributional assumptions remained highly significant for all values of h_n. The result of Table 3.4 also indicated the presence of a hazard effect following a benefit change for either wage offer distributional assumption. The elasticity of the reservation wage with respect to the benefit level and the elasticity of the hazard with respect to the arrival rate of job offers remain significant for all values of h_n. These results were also seen in Table 3.4. However the elasticity of the reservation wage with respect to the arrival rate is only significant for values of h_n such that $0.25 < h_n > 0.65$. Interestingly, this elasticity was not significant in the analysis of Table 3.4. Perhaps this suggests that the correct bandwidth, h_n, lies in the region $0.25 > h_n < 0.65$. This result highlights the importance of the selection of h_n. Even if we put aside the movement of this one elasticity in and out of significance, it remains the case that the magnitude of the significance of all the parameters we have estimated varies with h_n and as a result

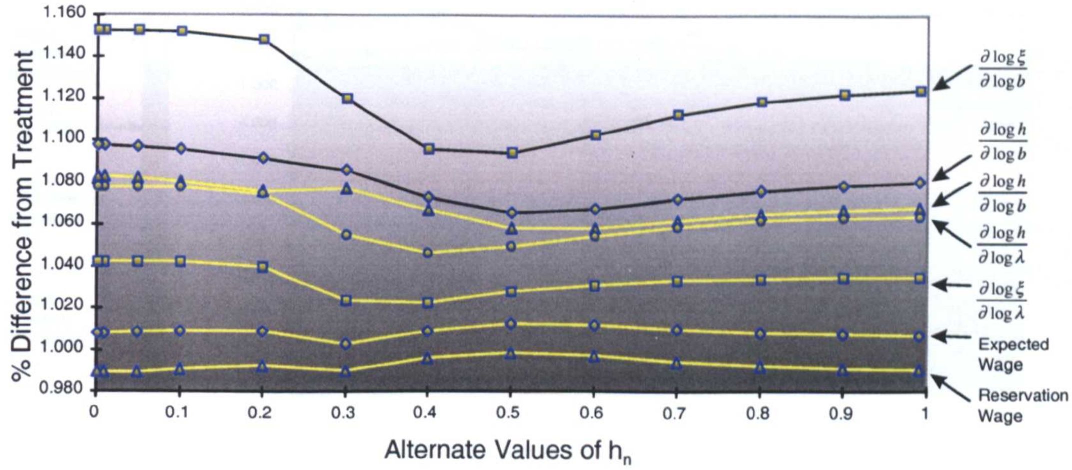

Figure 4.17: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Males) See Appendix C.2

it is difficult for us to arrive at a prediction of the size of the treatment effect on each parameter.

We can again examine the way in which the magnitudes of the treatment effects vary by standardising the estimated synthetic parameters by their treatment parameter counterparts. Figure 4.17 contains plots of the predicted synthetic search elasticities and reservation/expected wages for the male sub samples after standardisation. The reservation and expected wage rates of the male sample remain fairly stable over the range of h_n; they are close to 1 as was indicated by their t-statistic plots of figure 4.16. The elasticity of the reservation wage with respect to the arrival rate of job offers for persons with treatment is 4%(approx.) greater for all h_n. The elasticities of the hazard with respect to the benefit level, or the arrival rate under both the Pareto and exponential wage offer distributional assumptions experience a treatment effect of around 6% – 10% over the range of h_n after treatment as opposed to the kernel density plots of Figure C.2 which moved by 5% – 10% and the nearest neighbour estimates of Table 3.4, which exhibited a treatment effect of around 5% for both these elasticities. Here we see the

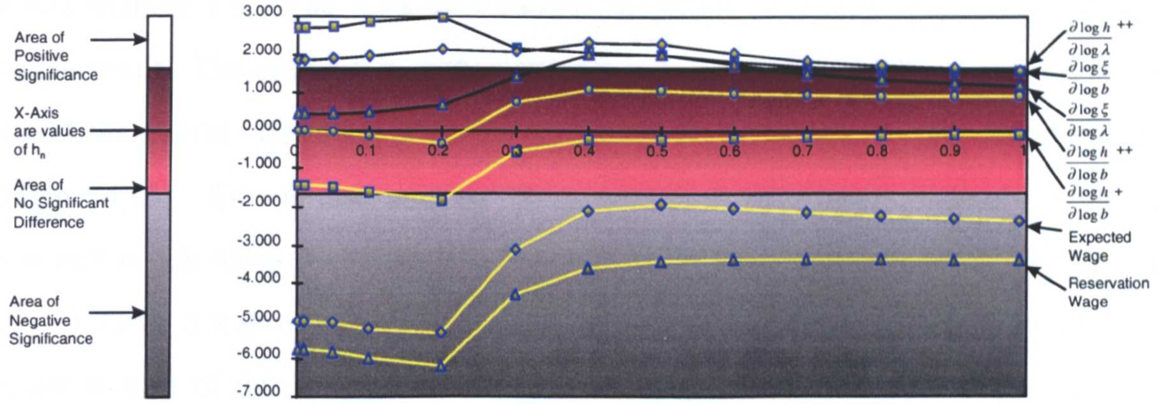

Figure 4.18: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Females) See Appendix C.2

elasticity of the reservation wage with respect to the benefit level exhibits a treatment effect of around 11% – 15%. This compares with the kernel density plots of Figure C.2 that suggested the effect between 20% – 30%. The nearest neighbour method produced a pair of elasticities, which indicated a reduction of the responsiveness of the reservation wage by around 7.4%.

As the reader will recall, our comparison between the t-test results for the nearest neighbour matched females by all types of YTS participation against those obtained following a match performed using the kernel density matching algorithm suggested that alternative matching methods could result in differing conclusions. This result contradicted the evidence from the male analysis, where both algorithms produced results of a similar nature. Our local linear matched male sample also produced t-test results akin to those of our previous analysis. What of the female local linear results?

Figure 4.18 contains the t-test results for females with and without any form of YTS experience, after local linear matching. Once more we compare our results to those of Table 3.4. Within Figure 4.18 we observe a wage effect for persons with any YTS experience. Both expected and reservation wages are reduced for women with such experience. As with the kernel density matching analysis, t-tests between the wage estimates of women with and without a spell on YTS after nearest neighbour matching (Table 3.4) produced the same result. The elasticities of the reservation wage with respect to both the benefit level and the arrival rate of job offers are significant for some values of h_n (see their plots on Figure 4.18). This is also true for the elasticity of the hazard with respect to the arrival rate of job offers (exponential wage offer distributional assumption). The female results of Table 3.4 show these three elasticities to be significantly altered in the same direction as part of this local linear matching analysis. Hence the local linear regression matching algorithm produces elasticities and wage estimates for synthetic control persons that are significantly different from their YTS treatment person counterparts in line with the results we saw in the nearest neighbour case. This result contradicts some of the t-test evidence of the female kernel density matched sub sample. In Figure 4.14 we saw the elasticities of the hazard with respect to the benefit level under both wage offer distributional assumptions and that of the reservation wage with respect to the arrival rate of job offers were significant for values of $h_n > 0.25$. Only the kernel density analysis indicated a significant change for the hazards with respect to the benefit level.

Figure 4.19 depicts the female elasticity/wage plots standardised as before. Both the reservation and expected wages experience a treatment effect of around $2\% - 5\%$. This compares to the kernel density plots of Figure 4.15, where we saw a change of around $4\%-7\%$ and the point wage estimates of the nearest neighbour matching algorithm (Table 3.4), where there was an effect of around $2\% - 4\%$. The elasticity of the reservation wage with respect to a change in the benefit level (Figure 4.18) varies by $5\% - 8\%$, whilst the kernel density analysis (Figure 4.15 not demonstrate the presence of a significant treatment effect and the nearest neighbour estimates of Table 3.4 indicated an effect of around 9%. The elasticity of the hazard with respect to the arrival rate of job offers (exponential wage offer distributional assumption) after local linear matching experienced a change of $3\% - 4\%$, no significant effect under the kernel analysis of Figure 4.15 and around 6% following a nearest neighbour matching procedure, Table 3.4.

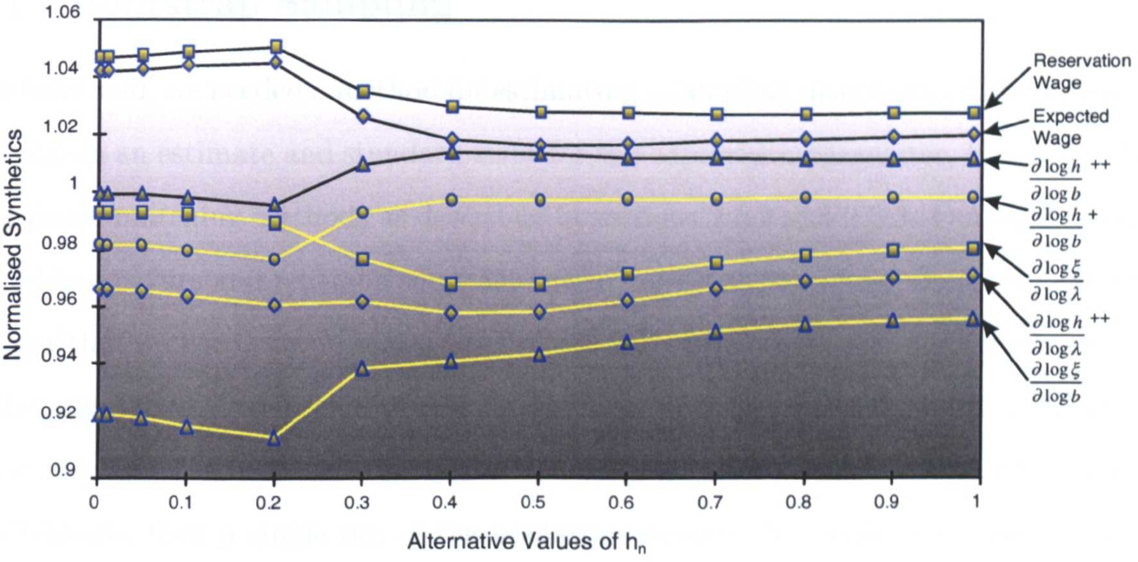

Figure 4.19: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Females) See Appendix C.2

4.4 Choice of the Bandwidth Parameter

As has been highlighted by the various figures for t-statistic and standardised elasticity/wage estimates, the selection of the kernel bandwidth parameter h_n can have an important bearing on the nature and magnitude of any YTS treatment effects we uncover using the kernel density or local linear matching methods. It is therefore of great importance that we arrive at a method of choosing a reasonable value for h_n given the dataset upon which we are to match. The unique nature of our analysis complicates this task. It is not possible to generate a value of h_n prior to running the algorithm (kernel or local linear). Hence we propose a bootstrapped iterative procedure as a solution to this problem.

4.4.1 Bootstrap Sampling

As we have said, we needed a method for estimating a sampling distribution from which we could obtain an estimate and standard error for the bandwidth parameter, h_n, used in the kernel based matching methods as described in sections 2.5.2 and 2.5.3. Our solution was of an ad-hoc nature and relied on some aspects of bootstrap sampling. Before we describe the algorithm we used, we present a brief discussion of bootstrap sampling methods.

The population distribution of any statistic can be approximated by the distribution of that statistic derived from the bootstrap samples. If we have a dataset containing N individuals, then a single run of the bootstrap samples N people from the N person dataset with replacement. It follows that some of the original people may appear on multiple occasions whilst others not at all. Hence, each bootstrap sample contains a sub sample of sample population individuals, many of whom appear more than once.

Having generated this new dataset we can estimate statistics of interest, in this case the job search elasticities and wage rates. This procedure is can then be repeated, each time generating a new set of estimates. Finally, we will arrive at a dataset of estimated statistics.

Bootstrapping relies on the validity of the asumption that the observed distribution is a good approximation of the population distribution. For this procedure to produce good estimates, this assumption must hold. The accuracy of this estimate of the sampling distribution is a function of the number of repetitions. As a result, some experimentation is required to identify the number of repetitions needed.

4.4.2 Estimation of the Optimal Bandwidth Parameter $\mathbf{h_n}$

Our solution to the problem of choosing a value for h_n is to take the basic local linear matching algorithm and apply it iteratively to the dataset until the difference between a generated h_n and the previous one is sufficiently small. The algorithm starts with an initial value for h_n [4].

From this starting point the algorithm generated a matched sample from the dataset, containing synthetic control persons and their elasticity and wage estimates constructed from the weights placed on control individuals using the formula of Equation (2.41). It then used these estimated synthetic elasticity and wage controls to estimate a new value for h_n using an approach akin to that of Silverman (1986), see Equation (4.4). The process was repeated, except that in subsequent iterations the control group on which Equation (2.41) placed a set of weights was the synthetic control group[5] that the weights from the previous iteration had generated. This continued until the difference between the last two iteration estimates of h_n was sufficiently small[6].

The main program only passed a random sub sample of size one thousand, of the sample population of treatment and control persons to the above subroutines. This process was repeated until it had generated one thousand estimates of h_n and their corresponding synthetic treatment and control elasticities/wages from one thousand intital random sub samples. An example of the bootstrap iterative code, which produced these estimates, is given in Appendix B.2.

The choice of the bootstrap sample size (number of iterations) was made after an analysis of the evolution of our estimates for h_n as the bootstrap sample size increased. Table 4.1 presents this analysis for males with YTSI experience. Figure 4.20 contains the corresponding plots of the estimates of h_n, which each of the one thousand bootstrapped iterations of the local linear regression algorithm produced. Notice that the majority of the values lie around the overall bootstrapped estimate of 0.041. At the same time the standard deviation of the estimates settles around 0.013. Plots of the bootstrapped iterative estimates of h_n for all the other versions of the YTS scheme, both male and

[4]During our analyses we chose to employ an initial value of $h_n = 0.5$. The choice of which was entirely arbitrary. In tests, the algorithm was not found to be sensitive to alternate initial values for h_n

[5]Specifically the Synthetic control group propensities

[6]When the difference between two successive estimates of $h_n \leq 0.005$ the iterative procedure was terminated and the current values for the synthetic estimates of the elasticities and wages were recorded.

female were of a similar form, with most iteration estimates of h_n around the overall estimates for each sub sample and are not reproduced here.

	Size of Bootstrap Sample									
	5	10	25	50	75	100	150	200	500	1000
Est. h_n	0.037	0.047	0.049	0.044	0.044	0.042	0.042	0.042	0.041	0.041
Sd(h_n)	0.007	0.016	0.022	0.013	0.013	0.013	0.013	0.014	0.013	0.013

Table 4.1: The Evolution of Bootstrapped Estimates of h_n as the Size of the Bootstrap Sample Increases for Males with YTSI Experience

For every variable estimated we now possessed a distribution. As the bootstrap sample size was increased so we were able to observe the shape of these distributions form. Figures C.7 to C.9 contain the distributions for the synthetic control elasticities and wages as the sample size increased from 5 to 500 (Males with YTSI experience). One can observe the distributions settling down to a bell curve shape as the number of iterations rises to 500.

The male treatment group elasticity/wage distributions by YTS type are reproduced in Figure C.10. The corresponding distributions for matched male controls by YTS type (bootstrap sample of 1000 iterations) are presented in Figure C.11. Notice that the tails of all the distributions for synthetic matched male controls are longer than those of the male treatment group elasticity and wage distributions. This result arises because bootstrap samples are taken with replacement, therefore the proportion of outliers in the bootstrap sample may be higher than in the original one. Hence bootstrap distributions may have heavy tails, which can result in inflated variance estimators. This property may be manifest in the standard deviations reported in Table 4.1. Two similar sets of distributions for the female treatment and control groups by YTS type are reproduced in Figures C.12 and C.13. As with the male sample, the tail ends of the female synthetic control distributions are longer than those of the female treatment group elasticity and wage distributions.

For each YTS treatment type and for both sexes we had generated a bootstrapped

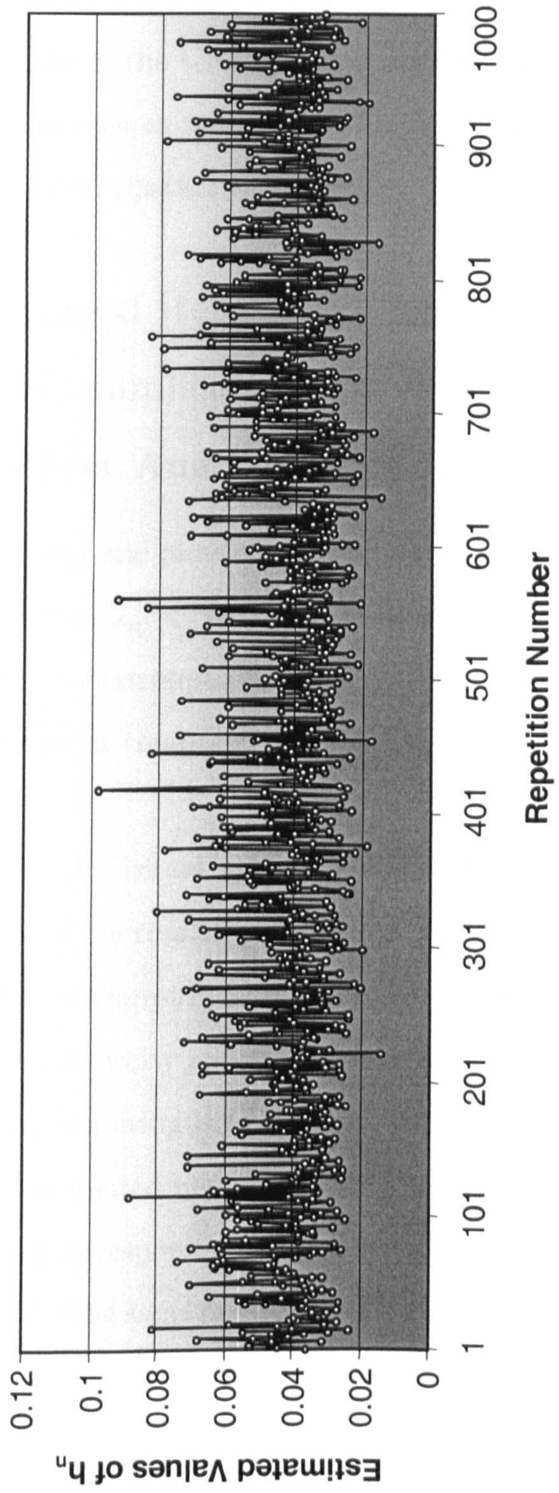

Figure 4.20: Estimates for the Value of h_n Using Bootstrapped Iterative Local Linear Regression Matching (Males YTSI)

sample of size one thousand for h_n. This allowed us to draw the resulting bootstrapped average value for h_n on graphs of the t-statistic and normalised elasticity/wage estimates and hence draw some conclusions as to the direction and magnitude of YTS treatment effects on the various job search elasticities/wages.

4.4.3 The Bootstrapped Iterative Optimal Bandwidth Parameter, $\mathbf{h_n}$ and it's Implications for Our Interpretation of the Treatment Effect Analysis (Males)

Figure 4.21 contains the t-statistic plots of elasticities/wages for YTSI males and their synthetic controls, where $0 < h_n < 1$. It also includes a vertical yellow line which indicates the position of the bootstrapped iterative estimate of h_n for this dataset. Notice the difference in the plots when compared to those for all YTS types which we saw in Figure (4.16).

As with our analysis of the kernel and local linear results for the combined YTS samples, the interpretation of the results produced by the matched samples of each YTS type is aided by a return to the nearest neighbour results obtained using the same datasets. Table 4.2 contains the mean elasticity and wage estimates which are calculated from their distributions which we obtained using the bootstrap iterative procedure when estimating h_n. T-test results are also presented here and test for significant differences between the means of the treatment and corresponding control groups. Table 3.5 contains the nearest neighbour matched results for the same male datasets. Figure 4.21 contains the local linear t-test plots from the male YTSI sample and may be compared to the plots of Figures C.1 and 4.16. The difference is striking. Whereas the plots of the later two figures were broadly comparable, those of the former exhibit some marked differences. This highlights the heterogeneity, which was hypothesised and then evidenced by the results of Chapter 3 subsection 3.4.1. Firstly we see that Figure 4.21 and Table 4.2 columns 1 and 2 suggest the presence of a wage effect for males with YTSI experience. Both the reservation wage

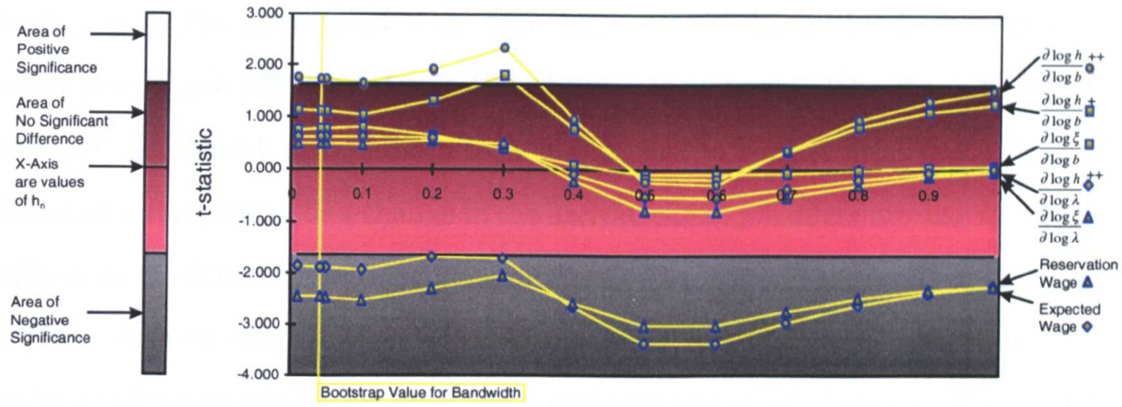

Figure 4.21: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Males YTSI) See Appendix C.2

and expected wages of these individuals undergo an effect, which sees them lowered as a result of the scheme. The results of Table 3.5 indicated that there may be some evidence for a wage effect, in that the reservation wage, post treatment was significantly lower to that for those without treatment at the 10% level. Table 3.5 also indicated that YTSI had no affect on any of the job search elasticities of interest. This result is largely repeated in this local linear analysis. Table 4.2 suggests that only the elasticity of the hazard with respect to the benefit level under the exponential wage offer distributional assumption exhibits a significant effect post treatment. However, an examination of the t-test plot for this elasticity in Figure 4.21 shows that the test statistic lies near the area of no significance. This combined with the fact that the elasticity of the hazard with respect to the benefit level under the Pareto distributional assumption is insignificantly different post treatment suggest that results of the male YTSI analysis are similar to those of the nearest neighbour study of Chapter 3.

Figure 4.22 contains the t-test plots for the elasticities and wages of treatment/control persons following local linear regression matching on the male YTSII dataset. A look to the values of the t-test statistics calculated at the bootstrap iterative optimal estimate of the bandwidth, h_n, shows evidence of more substantial treatment effects for those with YTSII experience. All five elasticities of interest are affected by a spell on YTSII. The elasticities of the reservation wage with respect to a change in the benefit level or the arrival rate of job offers are lower for scheme participants. The magnitudes of the elasticities of the hazard with respect to a change in the benefit level (both wage offer distributional assumptions) or a change in the arrival rate of job offers are also reduced. All these effects are well within the area of significant difference. As with our analysis of the male YTSI treatment effects, Table 4.2, columns 3 and 4 contain the mean elasticity and wage values and t-test results for both the treatment and control groups after matching at the bootstrap iterative optimal value for h_n. These results are comparable to those of Table 3.5 columns 3 and 4 following a nearest neighbour matching procedure. Recall that the results of Table 3.5 had suggested that there did indeed exist a treatment effect for these elasticities. All were lowered by a similar order of magnitude.

Again there appears to be a consistency in the results regardless of the matching procedure used. However the results of Table 3.5 also led us to conclude that there was no evidence to support the suggestion of a wage effect post YTSII treatment. In contrast, Figure 4.22 and Table 4.2 do indicate that there may be a reservation wage effect. Here we see reservation wages rise for those with YTSII experience. Although the plot of Figure 4.22 for the t-test statistic of the reservation wage effect does cross into the area of no significance for values of $0.4 > h_n < 0.7$. This result (as are others) is therefore reliant on the assumption that our optimal value for h_n is a correct one.

Figure 4.23 presents the t-test plot results for the male YT dataset after local linear matching for various values of h_n. Both the expected and reservation wage t-test plots remain insignificant over the whole range of h_n and are therefore insignificant at the optimal value for h_n. Table 4.2, columns 5 and 6 reflect this result with the mean wage values calculated after local linear regression matching with a bandwidth of h_n producing insignificant t-test statistics. This corresponds to the nearest neighbour matching results of Table 3.5, columns 5 and 6, where we saw no evidence for the existence of a wage effect following a spell on YT.

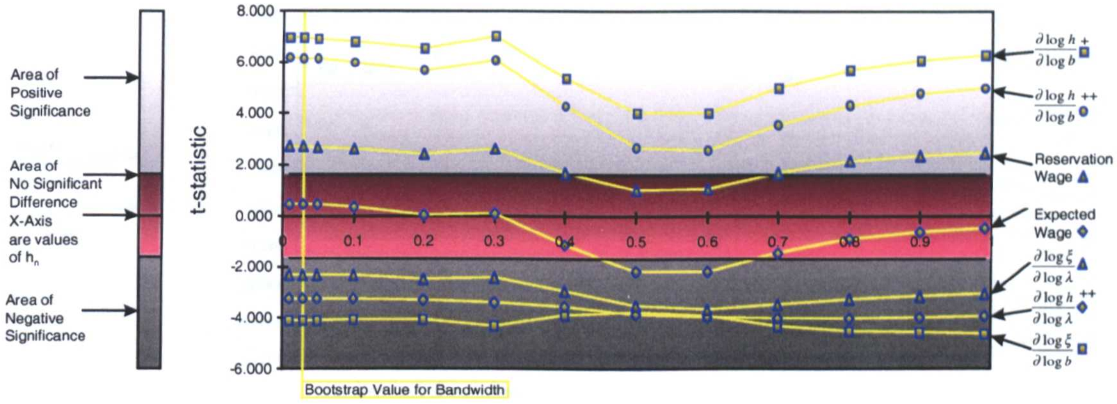

Figure 4.22: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Males YTSII) See Appendix C.2

The t-test plots for the elasticities of the reservation wage with respect to the benefit level and the hazard with respect to the arrival rate of job offers Figure 4.23 indicate a large treatment effect for YT participants, which produces the result of lowering these elasticities for the treated (see also Table 4.2). This is in contrast to the results of Table 3.5 (nearest neighbour matching), which gave no evidence for a YT treatment effect for these job search elasticities. The t-test and mean value estimates after optimal h_n local linear regression matching for both elasticities of the hazard with respect to the benefit level (Figure 4.23) indicate that these elasticities experience a reduction in magnitude following a spell on YT. The result of Table 3.5 after nearest neighbour matching mirrors these findings. However, the elasticity under the exponential assumption was only significant at the 10% level. The plot of Figure 4.23 for the elasticity crosses into the area of no significance for $0.4 > h_n < 0.7$. Perhaps this can help to explain the nearest neighbour result.

Lastly, we see that the t-test plot for the elasticity of the hazard with respect to a change in the arrival rate of job offers (exponential wage offer distributional assumption) is significant at the optimal value for h_n. The elasticities' mean values (treatment and control) indicate a reduction in this elasticity post YT treatment. The result of Table

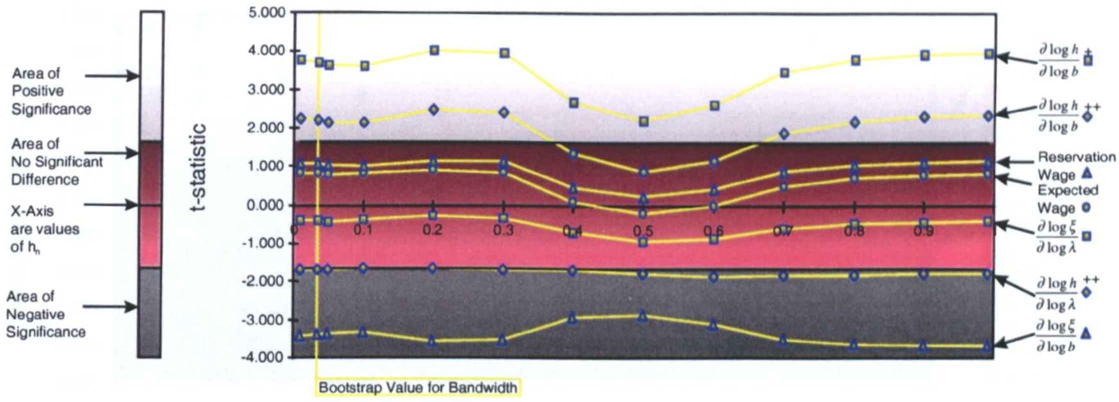

Figure 4.23: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Males YT) See Appendix C.2

3.5 following a nearest neighbour matching procedure led us to conclude that there was insufficient evidence for the existence of a treatment effect for this elasticity post YT. How can these two results be reconciled? As with the analysis of both elasticities of the hazard with respect to the benefit level, we see that the plot for the t-test statistics of the elasticity (Figure 4.23) lie on the boundary of the area of no significance and in fact cross this boundary for some values of h_n. This might explain the ambiguity between the results following different matching procedures.

We now turn our attention to the magnitude of the treatment effects upon the job search elasticities and wages following spells on the various versions of YTS. Figure 4.24 contains the standardised plots of treatment to control elasticites/wages following the local linear regression matching of Figure 4.21 and Table 4.2, columns 1 and 2 (males with YTSI experience). Again the vertical yellow line represents the point of the bootstrap iterative optimal h_n. Compare Figure 4.24 with the plots of Figure 4.16 to see the different results we obtain once we allow for YTS scheme heterogeneity. The presence of the optimal value for h_n allows us to compare the magnitude of various YTSI treatment effects at this point against those of the nearest neighbour analysis (Table 3.5. Table 4.3, columns 1 and 2 contain the significant elasticity and wage differences pre and post treatment

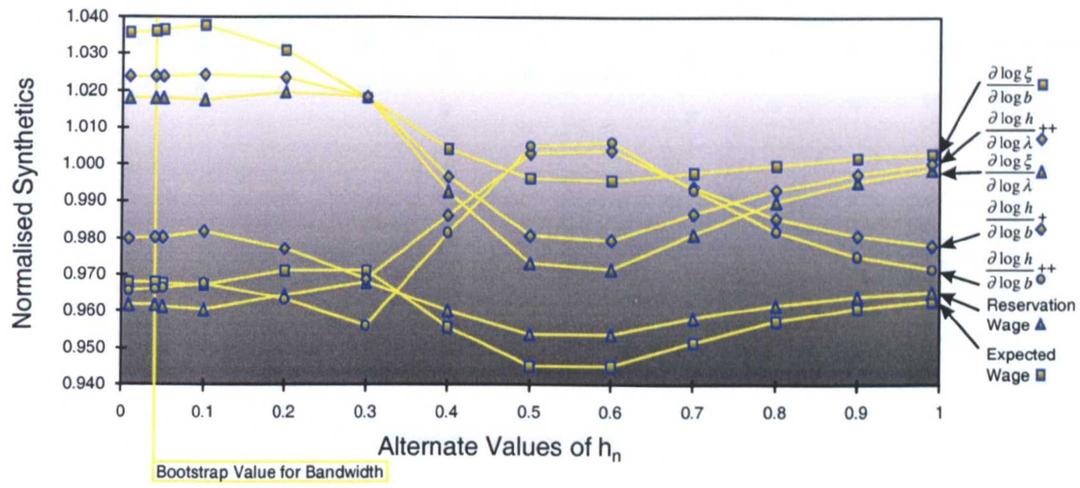

Figure 4.24: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Males YTSI) See Appendix C.2

following a spell on YTSI as estimated from nearest neighbour and local linear regression matched data respectively. Hence we see that the treatment effect on the reservation wage after a spell on YTSI is such that males reduce their reservation wages by around 2.7% (nearest neighbour result) or 3.8% (local linear result). The local linear regression male YTSI subset also indicated a treatment effect of around 3.3% for both the expected wage and the elasticity of the hazard with respect to the benefit level (exponential wage offer distributional assumption).

Figure 4.25 and Table 4.3, columns 3 and 4 allow us to compare the magnitudes of the treatment effects after a spell on YTSII. This scheme affects all the elasticities and there is also evidence of a wage effect. These results are present in both the nearest neighbour and local linear regression analysis. The evidence suggests that the elasticity of the reservation wage with respect to the benefit level for males with a spell on YTSII is lowered by between 16% (nearest neighbour result) and 15% (local linear result). This is a marked effect, which exists well beyond the area of no significant difference. Both the elasticities of the hazard

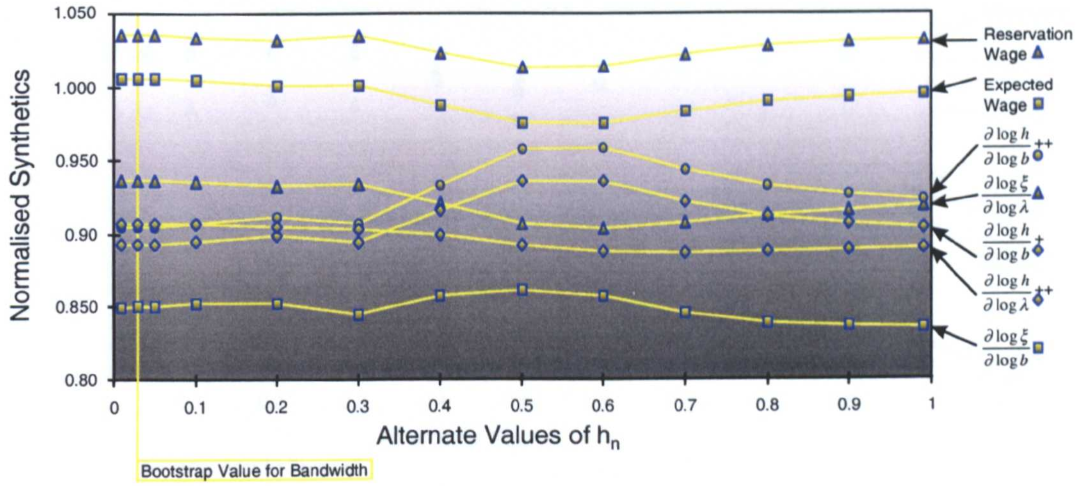

Figure 4.25: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Males YTSII) See Appendix C.2

with respect to the benefit level experience a treatment effect of 7.2% (Pareto wage offer distributional assumption) and 4.7% (exponential wage offer distributional assumption) for the nearest neighbour analysis. Whilst the same elasticities experience effects of 10.7% and 9.5% respectively with the local linear regression analysis. Again these are highly significant results. The elasticity of the reservation wage with respect to the arrival rate of job offers underwent an effect of around 7.7% under the nearest neighbour analysis and 6.4% for the local linear procedure. Finally, we see the elasticity of the hazard with respect to a change in the arrival rate of job offers reduced by 10.6% (nearest neighbour) and 9.3% (local linear). On the whole we have observed a consistency of both the significant effects, which the two matching methods have uncovered, and the magnitudes of these effects. All the significant effects were of a similar magnitude when compared with the corresponding results for the alternate matching procedure.

Figure 4.26 and Table 4.3, columns 5 and 6 contain evidence for the magnitude of the significant treatment effects for males with YT experience using the local linear matching

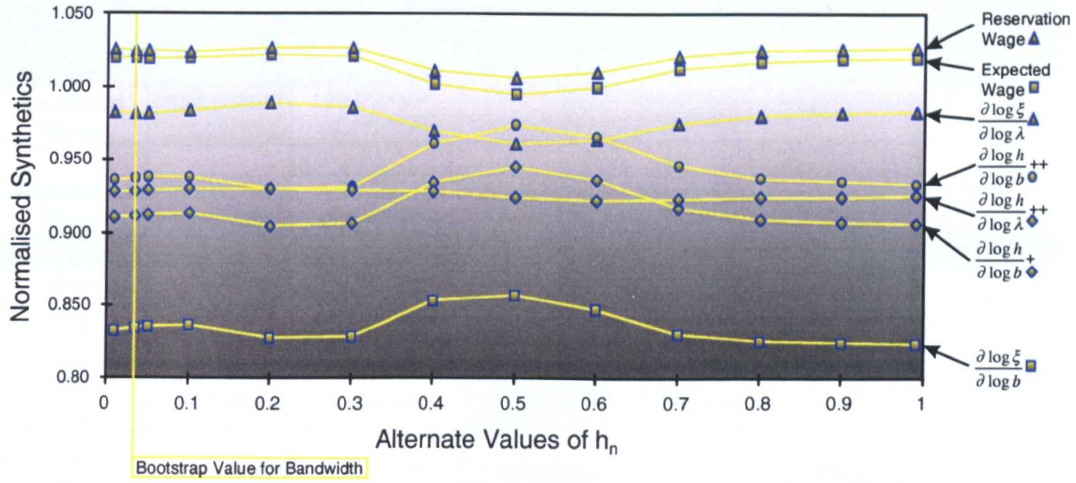

Figure 4.26: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Males YT) See Appendix C.2

algorithm. The plots within Figure 4.26 remain broadly constant over the whole range of h_n. We again concentrate on the effects as calculated at the estimated optimal value for h_n. We see more significant effects than we uncovered using the nearest neighbour algorithm (Table 3.6. The nearest neighbour analysis suggested a significant treatment effect for the elasticity of the hazard with respect to the benefit level (either wage offer distributional assumption) of 7.5% and 8.2% respectively. Whilst our local linear regression analysis suggests effects of 8.8% and 6.3% respectively. As with our previous analysis of the effects for males with spells on YTSI or YTSII these effects are of similar magnitude regardless of the method of matching used.

Means	Men YTSI=1	Men YTSI=0	Men YTSII=1	Men YTSII=0	Men YT=1	Men YT=0
			Samples			
Expected Wage	**122.837661** **(2.098533)**	126.902658 (0.415958)	127.111638 (1.540591)	126.386840 (0.228079)	130.416443 (3.027463)	127.967939 (0.211531)
Reservation Wage	**95.927513** **(1.523901)**	99.755042 (0.245034)	**101.446871** **(1.261891)**	98.009447 (0.157873)	98.454618 (2.218150)	96.094063 (0.162472)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	0.087658 (0.003931)	0.084608 (0.000669)	**0.079458** **(0.003357)**	0.093471 (0.000548)	**0.091758** **(0.005297)**	0.110008 (0.000687)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.177005 (0.006259)	0.173905 (0.000935)	**0.168128** **(0.004934)**	0.179581 (0.000628)	0.191412 (0.008316)	0.194890 (0.000593)
$\frac{\partial \log h}{\partial \log b}$ +	-0.412825 (0.007270)	-0.421132 (0.001629)	**-0.389081** **(0.006586)**	-0.435743 (0.001397)	**-0.409242** **(0.010636)**	-0.448803 (0.001615)
$\frac{\partial \log h}{\partial \log b}$ ++	**-0.325168** **(0.006330)**	-0.336524 (0.001670)	**-0.309623** **(0.005186)**	-0.342272 (0.001247)	**-0.317484** **(0.009659)**	-0.338796 (0.001137)
$\frac{\partial \log h}{\partial \log \lambda}$ ++	0.264662 (0.009778)	0.258513 (0.001505)	**0.247587** **0.007825**	0.273052 (0.001065)	**0.283170** **(0.012800)**	0.304897 (0.001043)
Sample Size	**340**	**340**	**476**	**476**	**220**	**220**
Btstrap h_n	**0.041185**		**0.029433**		**0.034939**	
+ Pareto Assumption, ++ Exponential Assumption						

Table 4.2: Elasticities for Matched Men Using the Local Linear Algorithm with Bootstrapped Estimate for h_n

Means	Samples					
	Nearest Neighbour Men YTSI	Local Linear Bootstrap Men YTSI	Nearest Neighbour Men YTSII	Local Linear Bootstrap Men YTSII	Nearest Neighbour Men YT	Local Linear Bootstrap Men YT
Expected Wage	-	3.2%	-	-	-	-
Reservation Wage	2.7%	3.8%	-	3.5%	-	-
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	-	-	16.1%	15.0%	-	16.6%
$\frac{\partial \log \xi}{\partial \log \lambda}$	-	-	7.7%	6.4%	-	-
$\frac{\partial \log h}{\partial \log b}$ +	-	-	7.2%	10.7%	7.5%	8.8%
$\frac{\partial \log h}{\partial \log b}$ ++	-	3.4%	4.7%	9.5%	8.2%	6.3%
$\frac{\partial \log h}{\partial \log \lambda}$ ++	-	-	10.6%	9.3%	-	7.1%
+ Pareto Assumption, ++ Exponential Assumption						

Table 4.3: Comparative Magnitudes of Significant Treatment Effects by Various YTS Types for Matched Men Using Both the Nearest Neighbour and Local Linear Algorithm with Bootstrapped Estimate for h_n

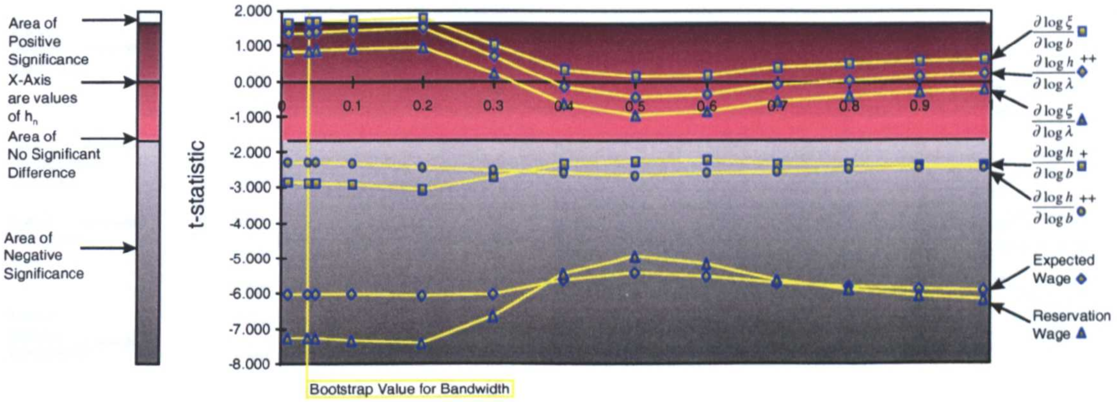

Figure 4.27: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Females YTSI) See Appendix C.2

4.4.4 The Bootstrapped Iterative Optimal Bandwidth Parameter, $\mathbf{h_n}$ and it's Implications for Our Interpretation of the Treatment Effect Analysis (Females) See Appendix C.2

We now move on to investigate the treatment effects, which the various forms of YTS had on the women within our sample. Figure 4.27 contains the t-test statistic plots for the elasticities and wages of females with and without YTSI experience. These plots should be compared to those of Figure 4.18 (all YTS types combined).

Once more, the vertical line represents the point at which the plots cross the optimal bootstrap iterative h_n. Table 4.4, columns 1 and 2 contain estimates of h_n and compares to the nearest neighbour results of Table 3.6 suggested that the effect of a spell on YTSI for the female subgroup was purely a wage reduction. The local linear regression matching technique also uncovers this wage effect. However, we now observe a set of elasticity effects. Firstly, the elasticity of the reservation wage with respect to the benefit level rises following a spell on YTSI this is also true of the elasticities of the hazard with respect to the benefit level (both wage offer distributional assumptions), which increase in magnitude post treatment. As does the elasticity of the hazard with respect to the arrival rate of

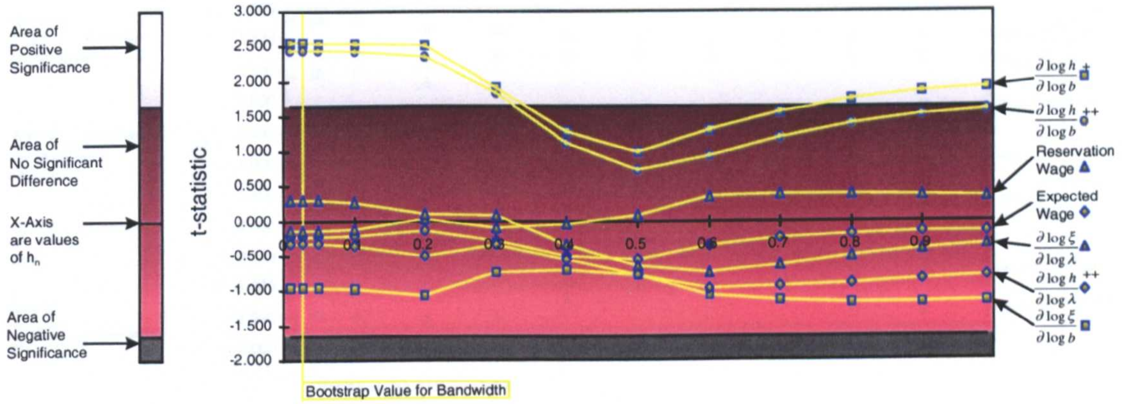

Figure 4.28: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Females YTSII) See Appendix C.2

job offers. All of these elasticity effects are only uncovered using the nearest neighbour algorithm.

Figure 4.28 contains the t-test statistic plots for females on YTSII for the range of values of h_n, having employed a local linear regression matching algorithm. Only the plots for both the elasticities of the hazard with respect to a change in benefit level lie beyond the area of no significance. The results of Table 4.4, columns 3 and 4 reveal that the evidence points to a reduction in these two elasticities when we employ a local linear regression approach. Previously, the results of Table 3.6, following the nearest neighbour algorithm had failed to produce any evidence of a YTSII treatment effect for female scheme participants.

Turning to look at YT scheme effects for females, we see that the t-test statistic plots of Figure 4.29 remain fairly constant for the plots, which exist beyond the area of no significance. Table 4.4, columns 5 and 6 present the estimated mean elasticities and wages as calculated at the yellow vertical line of Figure 4.29 and compares to the results of Table 3.6, columns 5 and 6. The nearest neighbour analysis of Table 3.6 pointed to the existence of a number of YT treatment effects on the female sub sample. We saw a wage effect in that the reservation wage was seen to fall after a spell on YT. This result is repeated at

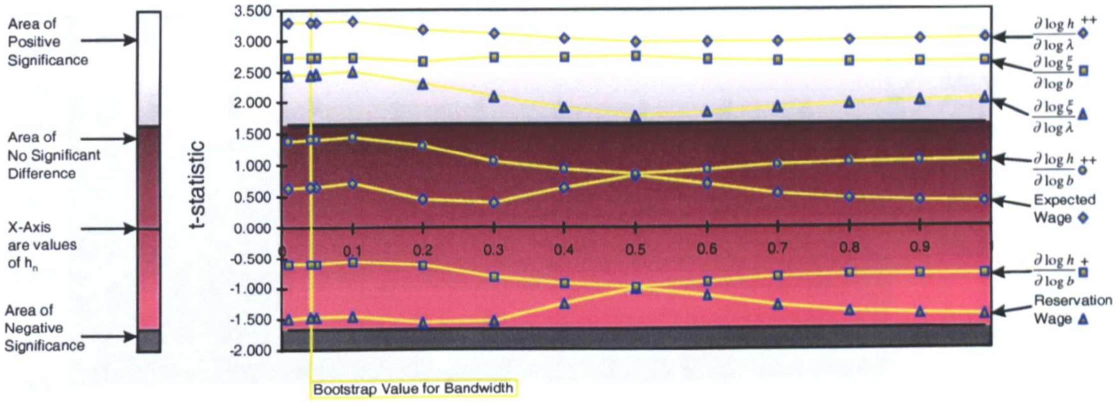

Figure 4.29: Local Linear Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Females YT) See Appendix C.2

the (10% level) within the results of Table 4.4. We had also discovered evidence for the existence of a rise in the elasticity of the reservation wage following a spell on YT. -It was seen to rise. This too is repeated within the local linear regression setting. Using local linear regression matching we also see a rise in the elasticity of the reservation wage with respect to the arrival rate of job offers after a spell on YT. -This we also saw for the nearest neighbour matched female dataset. The evidence for a YT effect on the elasticity of the hazard with respect to the benefit level is mixed. The local linear matching results reveal a drop in the elasticity (exponential wage offer distributional assumption), but this is only significant at the 10% level. In contrast the nearest neighbour analysis produced an effect on the elasticity (Pareto wage offer distributional assumption) that seemed to indicate a rise post treatment. These results conflict one another and are difficult to reconcile. Lastly, we observe (Table 4.4) a rise in the elasticity of the hazard with respect to the arrival rate of job offers. The nearest neighbour findings (Table 3.6) mirror this result.

We now examine the magnitude of the various YTS treatment effects we have uncovered. Figure 4.30 contains the standardised treatment effect plots over the range of h_n for the YTSI female sub sample compares to that of the combined YTS plots of Figure 4.18.

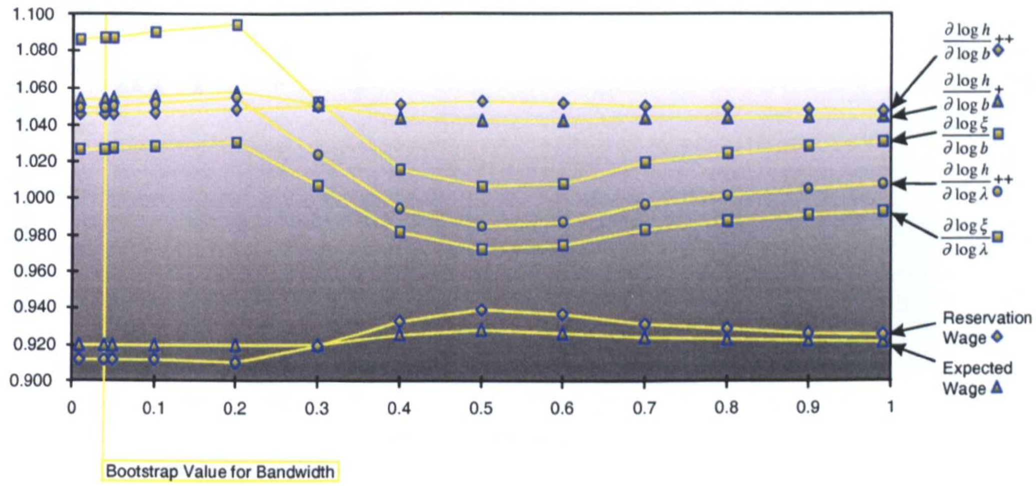

Figure 4.30: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Females YTSI) See Appendix C.2

The approximate size of the significant treatment effects for this YTSI female dataset as calculated at the yellow vertical line of Figure 4.30 are presented within Table 4.5, columns 1 and 2. The nearest neighbour algorithm produced a modified dataset from which the treatment effect on the expected and reservation wages was a reduction of 7.6% and 8.6% respectively. Whilst matched data from the local linear procedure suggests that these wages were reduced by 8% and 8.8% respectively. The local linear dataset analysis also contains evidence for the existence of some treatment effects for the elasticities of interest. The elasticity of the reservation wage with respect to the benefit level underwent an increase of 8.6%. The two versions of the elasticity of the hazard with respect to the benefit level increase in magnitude by 5.4% (Pareto) and 4.5% (exponential). The elasticity of the hazard with respect to the arrival rate of job offers also increased by 4.9%.

Figure 4.31 contains the standardised treatment effects plots for women following a spell on YTSII. The approximate treatment effects as evaluated at the optimal bandwidth, h_n, are reproduced within Table 4.5, columns 3 and 4. The nearest neighbour results of

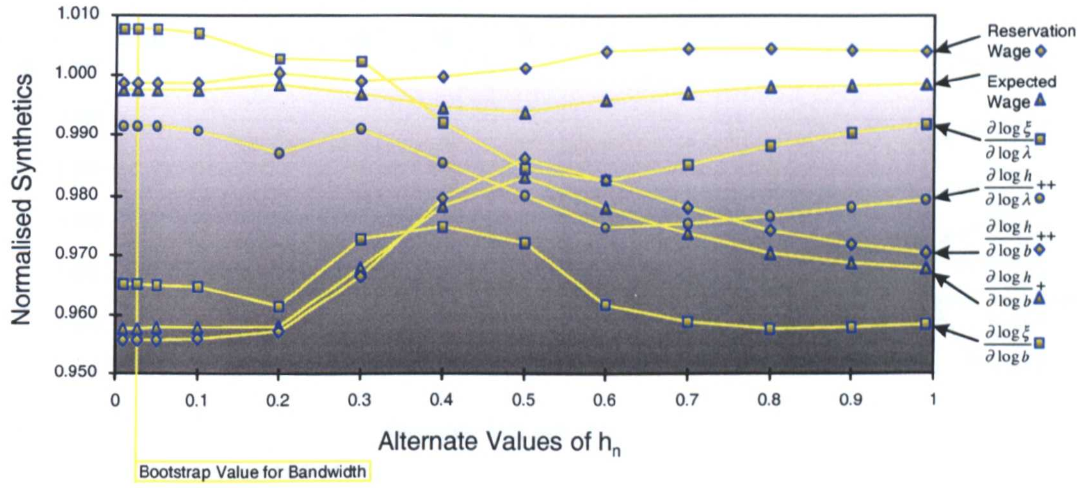

Figure 4.31: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Females YTSII) See Appendix C.2

Table 3.6 contained no evidence for the existence of any effects on the job search elasticities and wages. However, these results conflict with those we observe within the local linear regression setting to the extent that we see an effect on both versions of the elasticity of the hazard with respect to a change in the benefit levels.

Figure 4.32 contains the standardised plots of the treatment effect upon the elasticities and wages of interest for females with YT experience. Table 4.5, columns 5 and 6 contain the magnitudes of those effects, which we found to be significant (as calculated at the optimal value for h_n). We uncover some evidence for the existence of a wage effect. The reservation wage was subject to a reduction of around 2.8% following a spell on YT. However, this result was only significant at the 10% level. Analysis of the nearest neighbour matched dataset Table 3.6, columns 5 and 6, had revealed a reservation wage effect of 6.9%. The elasticity of the reservation wage with respect to a change in the benefit level is subject to a large treatment effect of around 56.4%. The nearest neighbour analysis had indicated an effect of around 22.5%.

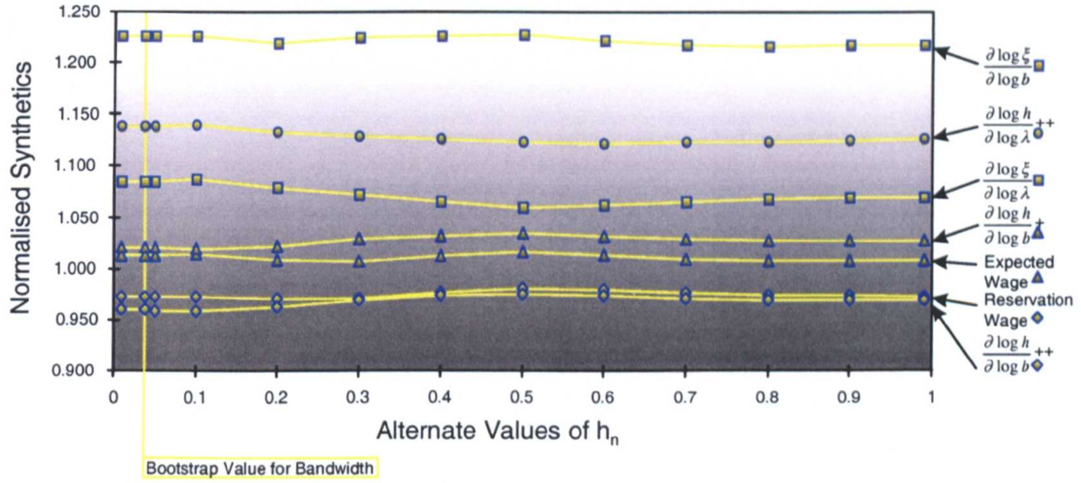

Figure 4.32: Local Linear Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Females YT) See Appendix C.2

The elasticity of the reservation wage with respect to a change in the arrival rate of job offers leapt by around 8.4% following a spell on YT (local linear regression matched dataset) and 17.5% (nearest neighbour matched dataset). Again the direction, if not the magnitude of the treatment effect was consistent irrespective of the matching method employed. As we saw during our analysis of the results for the differing specifications of the elasticities of the hazard with respect to a change in the benefit level are not consistent across the alternate matching methods. The elasticity under the Pareto wage offer distributional assumption seems to experience no treatment effects when using local linear regression matching. Yet the analysis of the nearest neighbour matched data suggests a rise of around 13.2%. In contrast, the local linear regression analysis suggests a fall in the elasticity under the exponential wage offer distributional assumption of 4.1%. There is no evidence for such an effect within the nearest neighbour analysis. Finally, the elasticity of the hazard with respect to a change in the arrival rate of job offers seems to rise by around 13.8% within the local linear regression matching analysis and 31% within the nearest neighbour matching analysis.

	Samples					
Means	Women YTSI=1	Women YTSI=0	Women YTSII=1	Women YTSII=0	Women YT=1	Women YT=0
Expected Wage	**104.805738** **(1.495201)**	113.933877 (0.229393)	114.968104 (1.341568)	115.261601 (0.130305)	119.141543 (2.230638)	117.721374 (0.223070)
Reservation Wage	**81.907391** **(1.071004)**	89.797707 (0.169267)	89.613785 (1.003444)	89.745478 (0.078049)	**88.109916** **(1.707600)**	90.643027 (0.104934)
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	**0.108035** **(0.005163)**	0.099460 (0.000473)	0.101906 (0.003819)	0.105594 (0.000377)	**0.136456** **(0.009209)**	0.111408 (0.000384)
$\frac{\partial \log \xi}{\partial \log \lambda}$	0.171995 (0.005269)	0.167523 (0.000583)	0.174638 (0.004213)	0.173309 (0.000542)	**0.196269** **(0.006241)**	0.181023 (0.000522)
$\frac{\partial \log h}{\partial \log b}$ +	**-0.506171** **(0.008997)**	-0.480369 (0.000884)	**-0.465501** **(0.008105)**	-0.486059 (0.000426)	-0.490360 (0.016360)	-0.480461 (0.000673)
$\frac{\partial \log h}{\partial \log b}$ ++	**-0.398137** **(0.007434)**	-0.380909 (0.000911)	**-0.363595** **(0.006908)**	-0.380466 (0.000669)	**-0.353904** **(0.010827)**	-0.369053 (0.000806)
$\frac{\partial \log h}{\partial \log \lambda}$ ++	**0.280030** **(0.009556)**	0.266983 (0.001038)	0.276544 (0.007378)	0.278903 (0.000908)	**0.332725** **(0.012258)**	0.292431 (0.000880)
Sample Size	**362**	**362**	**629**	**629**	**299**	**299**
Btstrap h_n	**0.038565**		**0.026783**		**0.036960**	

+ Pareto Assumption, ++ Exponential Assumption

Table 4.4: Elasticities for Matched Women Using the Local Linear Algorithm with Bootstrapped Estimate for h_n

	Samples					
Means	**Nearest Neighbour Women YTSI**	**Local Linear Bootstrap Women YTSI**	**Nearest Neighbour Women YTSII**	**Local Linear Bootstrap Women YTSII**	**Nearest Neighbour Women YT**	**Local Linear Bootstrap Women YT**
Expected Wage	7.6%	8.0%	-	-	-	-
Reservation Wage	8.6%	8.8%	-	-	2.8%	6.9%
Elasticities						
$\frac{\partial \log \xi}{\partial \log b}$	-	8.6%	-	-	22.5	56.4%
$\frac{\partial \log \xi}{\partial \log \lambda}$	-	-	-	-	8.4%	17.5%
$\frac{\partial \log h}{\partial \log b}$ +	-	5.4%	-	4.2%	-	13.2%
$\frac{\partial \log h}{\partial \log b}$ ++	-	4.5%	-	4.4%	4.1%	-
$\frac{\partial \log h}{\partial \log \lambda}$ ++	-	4.9%	-	-	13.8%	31.0%

+ Pareto Assumption, ++ Exponential Assumption

Table 4.5: Comparative Magnitudes of Significant Treatment Effects by Various YTS Types for Matched Women Using Both the Nearest Neighbour and Local Linear Algorithm with Bootstrapped Estimate for h_n

4.5 Context and Implications of Results

This chapter began with some thoughts on the ideal pre-match dataset. Using a pair of stylised distributions (Figure 4.1) we suggested that our pre-matched dataset might in fact resemble the situation we outlined in Figure 4.1 Case 2. If Case 2 did approximate our dataset, then we hypothesised that a nearest neighbour matching method would produce progressively poorer quality matches as the algorithm proceeded. Our analysis of Figures 4.2 and 4.4, where we presented the actual propensity score distributions for the male and female combined YTS datasets concluded that there was a subset of high YTS propensity treatment individuals for whom there were insufficient corresponding high propensity control persons. Further to this regard, we presented a pair of plots (Figures 4.3 and 4.5) depicting the average distances between nearest neighbour matched pairs as the algorithm progressed[7]. Emphasis was placed on the implications of these results for any conclusions, which we might draw from any analysis on a matched dataset of this type.

We called into question the practice of dropping treatment persons from the sample non-randomly, for whom there were no controls in possession of sufficiently similar YTS propensity scores. We concluded that rejecting anybody from our dataset on a non-random basis would lead to a situation where an effort to remove the bias caused by a lack of high quality matches would in itself induce a secondary bias through the removal of mainly high propensity treatment persons. It was also noted that these high propensity persons were exactly the kinds of individuals, which the schemes were designed to benefit. As such the treatment effects on this subset of YTS participants was of great interest to us.

Section 4.2.1 touched on the concept of stochastic dominance and how an examination of the cumulative distributions of treatment and control groups' YTS propensity scores could allow us to predict whether a dataset would produce a good matched dataset with few poor matches (in terms of distance between YTS propensity scores within matched

[7]See Chapter 2 section 2.5.1 and Appendix C, section C.1 for a full exposition of the matching algorithm used.

pairs). There was a brief examination of the CDFs for the subsets of interest for our study. Only the male and female YT datasets proved to have sufficient high propensity controls from which to construct a complete nearest neighbour matched dataset without degradation in quality.

4.5.1 Results

Section 4.3.1 introduced some early evidence for the existence of treatment effects on the job search elasticities and wages for men and women who had any kind of YTS scheme experience using a kernel regression matched dataset. We commended the kernel regression algorithm for its use of the whole control group when synthesising each new control person. However, we tempered our optimism for the performance of the algorithm with a cautionary look at the nature of the bandwidth parameter, h_n, which needed to be chosen before running the algorithm. In an effort to demonstrate the problems of bandwidth selection we began to present t-test results for the significance of treatment effects over the range of $0 < h_n < 1$ (see Figures such as C.1). These plots showed the need for some method of selecting a value for h_n for each dataset which would allow the algorithm to generate a matched sample from which to draw some conclusions as to YTS treatment effects. The Silverman (1986) optimal method of h_n selection was discussed.

During the analysis of t-test results we returned to the equivalent nearest neighbour matching results, Table 3.4 Section 3.4, to consider whether these two matching schemes had produced synthetic datasets on which testing indicated the presence of similar treatment effects. We concluded that for the male dataset the results of t-tests even without a value of h_n on which to settle, were broadly in line with those which we saw within the nearest neighbour setting. However kernel regression matching, using the female dataset, over the range of h_n led to a set of t-test plots which were harder to interpret. We saw large fluctuation in the t-test statistics especially for values of h_n in the range $0.2 - 0.4$. We again returned to the nearest neighbour matching results of Table 3.4, Section 3.4 in an effort to discover whether the pattern of treatment effect was consistent across matching methods for the female dataset. We consistently uncovered evidence of a wage effect on the treated, irrespective of the algorithm used; both methods suggested the presence of a reduction in wage expectations post scheme, for females. However, evidence for a treatment effect on the elasticities of job search was inconsistent and therefore difficult to interpret. It was suggested that a local linear regression analysis might help to uncover the truth and allow us to form a consistent picture.

We also took time to consider the magnitudes of any significant treatment effects, which we found. Again we had to acknowledge that without a method to select a value for h_n we would have to present plots for the movement of treatment effects as h_n varied. Figures C.2 and 4.15 presented the ways in which the magnitudes of the treatment effects for males and females varied over the range of h_n.

Section 4.3.2 then examined the t-test results for the same combined YTS datasets after having performed local linear regression matching. A comparison between the male t-test plots of Figure 4.16 and the nearest neighbour t-test results of Table 3.4 concluded that there was a consistency of treatment effect for all matching algorithms we employed. However the elasticity of the reservation wage with respect to a change in the benefit level crossed beyond the area of no significant difference for some values of values of h_n. If h_n did reside within the area where this elasticity was significant, then this result would conflict with that of the nearest neighbour analysis. As had been suggested during our analysis of the results for the female kernel regression matched dataset, the equivalent female local linear regression matched dataset produced t-test results which were broadly in line with the results of the nearest neighbour analysis of Table 3.4. This cast doubt on the validity of the kernel regression results.

Again we examined the ways in which significant YTS treatment effects varied over the range of h_n. Figures 4.17 and 4.19 contained plots of the magnitudes of the treatment effects on the various elasticities/wages, as revealed using local linear matching. These highlighted the need for a single optimal value of h_n.

Section 4.4.2 introduced a solution to our bandwidth selection problems. We formulated the idea of a bootstrap iterative optimal value for h_n, with the method of Silverman (1986) at its heart. We demonstrated the convergence of the bootstrap estimates (Male YTSI sub sample). As the size of the bootstrap sample increased, the estimated mean of the bandwidth parameter settled to a given value whilst the bootstrap standard deviation tended to zero. Plots of the sample distributions for the various elasticities and wages were also shown to converge.

Subsection 4.4.3 saw us concentrate all further analyses solely on test results produced using local linear regression matching. No more use was made of the kernel regression algorithm. We now re-introduced the notion of presenting results by YTS type. Again, there was evidence to support the hypothesis of YTS scheme heterogeneity.

Figure 4.21 presented the t-test statistic plots for the YTSI male sub sample after local linear regression matching. All further plots now contained a vertical yellow line to indicate the location of the bootstrap iterative value for the optimal h_n. Elasticity and wage treatment effects were considered using their bootstrap iterated sample distributions[8]. The results within Table 4.2 were comparable to those of Table 3.5, subsection 3.4.1. Table 4.3 contained the magnitudes of all significant treatment effects for males with YTSI experience, as estimated using both the nearest neighbour and local linear regression matched datasets.

Results suggested that males with YTSI experience lowered their reservation wages by between 2.7% and 3.8% (depending on the matching algorithm used) after exiting to another labour market state. This indicates that YTSI produced the wage effect for males which some advocates of this initial incarnation of the scheme had hoped. The scheme appeared to cause participants to reappraise their earnings expectations and as such price themselves into work. The local linear analysis also suggested a similar effect on the expected wage, in which this important employment prospects indicator was also reduced. Furthermore the local linear regression results contained some evidence for the existence of a reduction in the elasticity of the hazard with respect to a change

[8]See Appendix C, Figures C.10 to C.13

in the benefit level (exponential wage offer distributional assumption) for males with YTSI experience. Again, this was good news in some political quarters as this result indicated that scheme participants' job search strategies became less sensitive to benefit level rises post treatment. These results seem to reflect some of the historical opinion regarding the nature of the overall YTSI scheme effect. Recall from our introduction in Chapter 1, that YTSI grew out of a previous Government training scheme, YOP (Youth Opportunity Program), which was an attempt to address what was generally considered to be a temporary excess supply of youth labour. However, as the 1980s began, the numbers of unemployed youngsters continued to climb. This led to the introduction of YTSI, which encouraged participants to learn general transferable skills, which employers might find of use. If YTSI succeeded in improving the skills base of its participants then we might expect to see some of the elasticity effects which we have discovered evidence of. At the same time the scheme also attempted to suppress reservation wages and there is evidence that it was successful in this regard.

The local linear regression algorithm produced a male YTSII matched dataset with some interesting t-test results. Unlike YTSI, YTSII seemed to lead to an upward appraisal of male participants reservation wages. This result suggests that this second incarnation of the scheme had a wage effect which was the antithesis of one of the original aims of the scheme (to lower wage expectations post participation). The nearest neighbour analysis had failed to uncover such an effect.

Results of t-tests for the possible existence of treatment effects on elasticities caused us to conclude that males with YTSII experience possessed a set of elasticities whose magnitude was reduced as a result of the scheme. These results mirrored those of the nearest neighbour analysis. Males with a smaller elasticity of the reservation wage with respect to a change in the benefit level will as a consequence be less sensitive to a rise in the benefit level. Allowing policy makers to raise the benefit level, thereby improving claimants' standards of living without so much of a rise in the numbers of claimants. A fall in the magnitude of the elasticity of the hazard with respect to a change in the benefit level would also have contributed to this effect.

The way in which one interprets the reductions in the elasticities of the reservation wage and hazard with respect to a change in the arrival rate of job offers depends on whether the economy was successful in generating increasing numbers of jobs during the period in which males began to exit YTSII. If the arrival rate of job offers was rising during this period, then YTSII would be regarded as having a detrimental effect as it depressed males' receptiveness to the increasing arrival rate of offers. Whereas if the rate of arrival fell then the treatment effect would ensure that the hazard was slow to fall. Conversely, it is desirable for the reservation wage to fall rapidly when the arrival rate of jobs falls and rise slowly when it climbs. Interestingly, this schemes introduction coincided with a fall in the numbers of young people leaving school and an economic boom[9]; the YTSII male wage effect would have acted to counter the upward pressure these labour market conditions would have placed on reservation wages. The depreciation in the elasticity of the reservation wage with respect to the arrival rate of job offers would also have acted to counter these conditions and maintain scheme participants competitiveness within the job market. However, the falling elasticity of the hazard of exiting unemployment would, under these conditions, lead to a smaller rise in response to an increase in the arrival rate of job offers.

Evidence for the existence of a male YT treatment effect was more pronounced within the local linear regression matching algorithm. The elasticities of the reservation wage and hazard with respect to a change in the benefit level and that of the hazard with respect to the arrival rate of job offers all experienced a fall in magnitude following a spell on YT. As YT matured the economy lurch back into recession, which in turn had a negative effect on the job prospects for school leavers. Under such conditions within the labour market, YT would have caused the hazard of exiting unemployment to decline at a reduced rate for its male participants. This may be interpreted as a positive result, given

[9] Appendix C, Figure C.5 contains plots of the evolution of several key labour market indicators over the time period of all YTS incarnations

the Governments wish that the scheme should aid the unemployment to work transition process.

Discussion then moved to the effects of each YTS incarnation on the female scheme participants. As we had discovered during the nearest neighbour matched data analysis of Section 2.5 the YTSI, YTSII and YT treatment effects for women uncovered using local linear regression matched datasets were very different to those uncovered for the corresponding male samples.

As with the male analysis, test results for the female YTSI sub sample using both nearest neighbour and local linear regression matched data revealed the presence of a lowering in both expected and reservation wages post participation. Again YTSI had some success at pricing the young into jobs. However, results for the treatment effects on the job search elasticities indicated that YTSI female participants had search elasticities of greater magnitude. Hence, female participants sensitivity to a change in the benefit level rose after treatment. Given that the economy was in recession as YTSI matured and those exiting the scheme were facing reduced job prospects and in the light of these findings any rise in the benefit level would have caused those females with YTSI experience to become uncompetitive within the job market at an increased rate to non-participants in the scheme.

Results for the treatment effects of YTSII showed that females were affected only through a reduction in the magnitude of both specifications of the elasticity of the hazard with respect to a change in the benefit level. The probability of participants exiting unemployment during this period would therefore be less responsive to any rises in the benefit level. There was no evidence for a wage effect. These results where found to be in line with those of the nearest neighbour analysis of Section 2.5.

Finally, we investigated the effect of YT on females within our dataset. There was some evidence for a wage effect post treatment. The reservation wage was seen to fall. -The last incarnation of YTS did appear to price women into work. Evidence for the effect of YT on elasticities was mixed. The elasticities of the reservation wage with respect to changes in either the benefit level or arrival rate of job offers increased post treatment. Given the unhealthy state of the labour market at the time when those who undertook a period of YT training were finishing their training it is reassuring to uncover evidence of increased sensitivity of the reservation wage with respect to the arrival rate of job offers as those with scheme experience would have acted quicker to adjust their wage expectations downward as conditions deteriorated. Evidence for a rise in the elasticity of the hazard with respect to a change in the arrival rate of job offers was also uncovered. With the arrival rate falling, with probability females with YT experience exiting from unemployment would have fallen at a greater rate than that of non-participants.

4.5.2 Random Exclusions: A Solution?

Figures 4.3 and 4.5 demonstrated that our fears for the performance of the nearest neighbour algorithm were well founded. As a run of the algorithm progressed the average distance between matched propensity scores began to rise. Interestingly this rise was not gradual; the average distance between propensity score matches for both male and female matching runs rose sharply towards the end of the process.

Conventional nearest neighbour matching methods would not have led to a degradation in match quality. However, this high level of match quality would have been maintained by excluding people from the treatment group for whom there was no control person left whose propensity was sufficiently similar. We cannot countenance the use of such methods. The nature of the dataset, which we uncovered during our prematch examination of stochastic dominance of treatment and control group propensity score distributions indicated that there might be a lack of support in the right tail of these distributions. Figures 4.3 and 4.5 did not reveal whether the poor matches were distributed throughout the matched treatment persons or concentrated in one area (Most likely to the right of the propensity score distribution). The plot in Figure 4.33 depicts the actual distance between the propensity score of matched male couples. The distances are ordered along the x-axis by the size of the propensity score of the treatment person in the matched

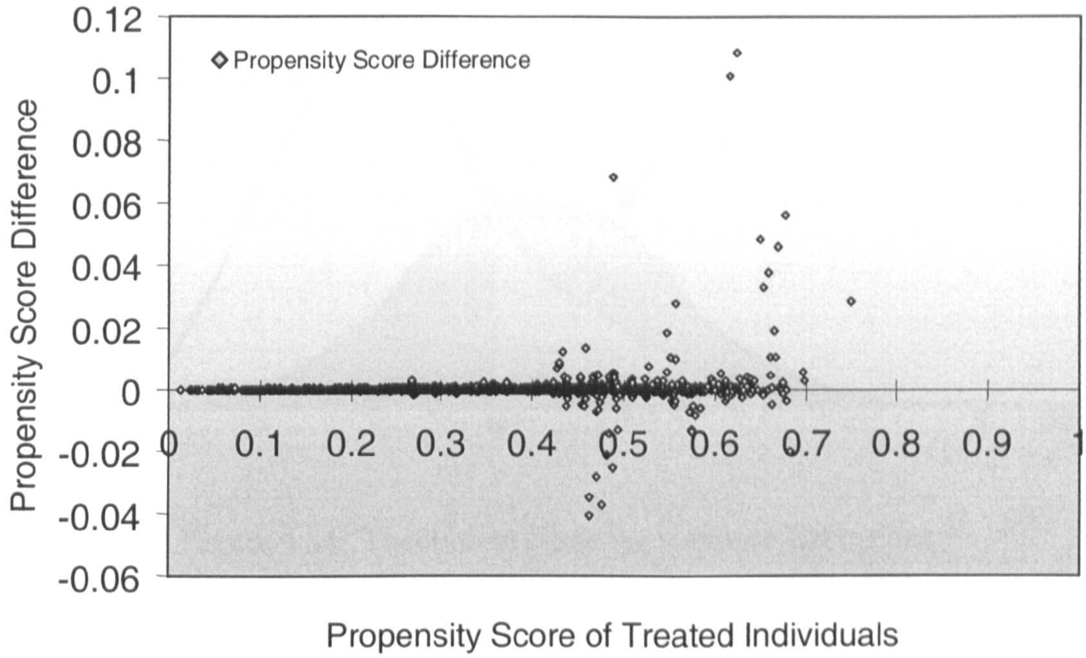

Figure 4.33: Matching Performance from Lowest Treatment Propensity to Highest Treatment Propensity (Males)

couple. Notice how the quality of the matches remains high for treatment persons with propensity scores under 0.4. However, as we had feared it is amongst the high propensity (0.5 and greater) treatment people that most of the poor matches occur.

Our study seeks to uncover the true nature of the YTS treatment effect. Beyond this goal we are interested in the effect of the scheme on high propensity persons, since the literature suggests that it was these young, low skilled individuals[10] for whom the scheme was introduced to help[11]. As such we were concerned that traditional matching methods might introduce a bias which could prevent us from uncovering the true YTS treatment effect. If our nearest neighbour algorithm had been adapted to include caliper exclusions, the majority of rejections would have occurred at the high propensity end of the pool of

[10]The Probit models which produced the propensity scores for both males and females in Table C.4 indicate a relationship between academic underachievement at 16 and higher propensity to participate in YTS.

[11]See Chapter 1 for an outline of YTS inception.

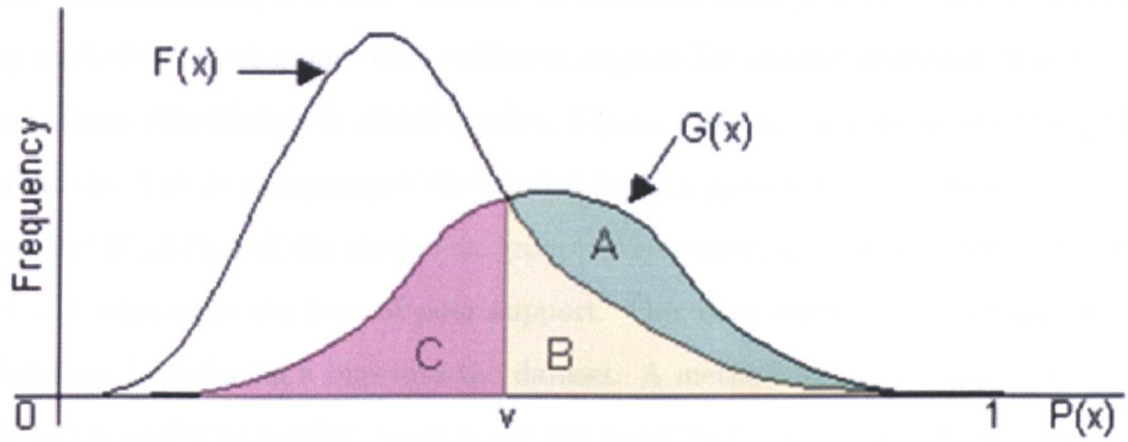

Figure 4.34: Theoretical Basis for Random Exclusions

treatment people.

We sought a solution to the above problem. There are variations of the nearest neighbour algorithm, which include returns of controls to the unmatched pool from the matched pool to counteract the lack of support. These methods raise a number of questions. Firstly, how many times should a single control person be available for matching? Twice, three, ten times? If no limit is placed on the number of possible rematches then analyses on the matched sample may lead to conclusions, which are over reliant on a handful of multiple matched high propensity control persons. Secondly, what is the criteria for the return of matched controls to the unmatched control pool. Should they be automatically returned to the pool after matching with a treatment person, or should they be excluded until such time as there are no control persons in the unmatched pool with a propensity, which lies within the caliper of a treatment person, upon which the full pool of matched and unmatched controls would become available for matching? Beyond these examples lies a large array of variations. The implications for the nature of the matched dataset, which might result are difficult to predict. We have chosen to examine the performance of just one example of the nearest neighbour algorithm with a view to a comparison with the results from kernel density and local linear regression matching methods.

We suggest that rather than allow the algorithm to drop treatment persons or resample controls non-randomly, it is more sensible to randomly drop persons from the treatment group until the control group offers sufficient support for nearest neighbour matching to occur without degradation in match quality. Figure 4.34 contains two stylised frequency distributions. Let $F(x)$ represents the control group's propensity score frequency distribution and $G(x)$ that of the treatment group. We breakup $G(x)$ into three areas, A, B and C. A represents the area of poor support. Our twin goals are to minimise region A whilst not introducing a bias into the dataset. A method based on caliper exclusions would be successful in the first requirement but would fail to meet the second. A random reduction in numbers from the combined areas under the curve, $G(x)$, would see the size of all three areas fall, without altering the ratios of each area to the others. Therefore, we would be minimising area A whilst adhering to the requirement that no new bias should enter the dataset. Of course there is a downside to such an approach. Random reductions in the size of the treatment group mean that not only are people dropped from area A (as with caliper exclusions) but also from areas B and C. As such, standard errors will rise with the result that it becomes harder to identify specific treatment effects. As an example, if the ratio of area A to the combined area of B and C was 1 to 2, then for every 1 person which caliper exclusions would drop from area A, a random exclusion method would need to drop a similar person from area A, that person plus another 2 from the combined area of B and C.

It should be noted at this point that there are those who do not consider non random exclusions to be a problem. They argue rightly, that the Rosenbaum and Rubin (1983) result does not require us to maintain the original "shape" of the treatment group in order to remain valid. This is indeed the case. Since all that it gives us is a framework of assumptions in which the propensity scores of the treatment group and control group persons are accurate predictors of the degree of similarity between any two persons plucked from the two groups. As such, even a dataset with a treatment group, the numbers of which have been reduced non-randomly to account for a lack of support, will still produce a well matched dataset. This also is true. However, any further tests on this dataset will only identify the treatment effect on those treatment people contained within it. this effect cannot be interpreted as the average treatment effect on the treated.

We are presented with a trade off. Caliper exclusions reduce poor matches but introduce bias into the dataset. Whilst random exclusions could also reduce poor matches whilst not introducing bias but the smaller size of the resulting matched dataset might make it difficult to uncover treatment effects.

4.6 Conclusions

The alternative to trying to develop the nearest neighbour algorithm further was to step sideways and employ a kernel based matching method. Such schemes take advantage of the whole dataset when generating each synthetic control person. Sample size is maintained, since there is no need to drop persons from the treatment group. However we saw evidence of the weaknesses of such methods. Firstly, the strength gained from the use of all of the control group can also develop into a weakness when most of the contribution to the synthetic controls comes from a handful of prematch controls. Secondly, the choice of the bandwidth parameter was shown to be crucial to the performance of the matching process. We produced results which showed variations in perceived YTS treatment effects when the bandwidth was varied between 0 and 1. Comparisons with elasticity and wage rate results from the nearest neighbour analyses suggested that conclusions drawn from datasets produced using differing matching algorithms can vary.

We suggest that the ability of the nearest neighbour and kernel based methods to produce well matched datasets is jointly dependent on the structure of the unmatched dataset and the inherent strengths and weaknesses of the algorithms themselves. It is not enough to simply perform these algorithms and present results as though the act of matching has removed the problem of self-selection whilst not introducing a secondary bias. The researcher must ask himself, what is the question I wish to answer? And will this procedure produce a dataset which might allow me to address it?

Chapter 5

Labour Market State Transition Processes

In this section we present a modified version of the Burdett *et al.* (1984), Kiefer and Neumann (1979), Mortensen (1986) and Neumann (1984) paper in which we model individuals' search and decision problems when faced with moving between unemployment, YTS training and work. Their formal model may be modified to consider all these states simultaneously although it requires some careful adaptation and reinterpretation. First we outline the model then we review the implications of the model for our empirical data problems.

We proceed to estimate the unemployment to work transition process using unmatched, nearest neighbour matched and local linear matched datasets. We examine the consistency of transition model results and find some evidence to support our earlier findings as to the sensitivity of kernel matching methods to the choice of bandwidth.

5.1 A Three State Model of Training, Unemployment and Work Transitions

Let U_j be utility in state $j = N, T, U$ where N denotes the state of employment, T denotes the state of training and U denotes the state of unemployment. Assume that each state has associated with it a utility which may be determined as follows.

$$\begin{aligned} U_N &= z + \epsilon_1 \\ U_T &= a + \epsilon_2 \end{aligned} \tag{5.1}$$

and

$$U_U = b$$

where z is the unconditional mean wage and individual attributes exogenously determine ϵ_1, a is a training allowance paid by employers to trainees and b is the unemployment benefit level (which is fixed).

This model is a reasonable approximation to what may be observed empirically in the labour market. Young people get jobs which will have wage levels influenced by their level of human capital. However this process is stochastic and may depend greatly on extraneous factors, unobservables or even luck. Likewise the determination of the allowance paid by an employer when training is partly fixed in the sense that there is a government minimum level of such payments. (See Data Appendix for details). However, in practice it is left up to the employer to pay what they deem appropriate to trainees. We can see from the distribution of this training allowance in our data graphed in Figure C.4 that a substantial minority of employers choose to pay more than the statutory minimum. From the viewpoint of the individual this discretionary component of the training allowance is stochastic.

Assume individuals get random draws of wage offers and training allowance offers from a cumulative probability distribution function $F(\epsilon_1, \epsilon_2)$ which is known to workers. In addition assume that the arrival of such offers follow a Poisson process with parameters δ_N and δ_T respectively. Burdett *et al.* (1984) show that given state j, future optimal decisions are characterised by

$$V_j(z+\epsilon_1, a+\epsilon_2) = \frac{U_j + \delta_j E[max_{k=N,T,U} V_k(z+\epsilon_1, a+\epsilon_2)]}{\delta_j + r}. \tag{5.2}$$

The analogue of the reservation wage property in this model is in terms of acceptance sets of doubletons (ϵ_1, ϵ_2). The position can be represented in Figure 5.1. This figure shows the areas of acceptance sets A_j associated with the three labour market states of unemployment, training and work. Conditional on the individual's draw on ϵ_1 and ϵ_2 this will determine which region in the diagram they fall into and hence the choice of the optimal state. The borders in the diagram provide the threshold levels between the different labour market states. Burdett *et al.* (1984) fully explore the analogue of this model for the states of unemployment, work and out of the labour force. They describe the comparative statics associated with changing levels of human capital on the acceptance sets which carry over to our modified structure.

Burdett *et al.* (1984) use the model to motivate a reduced form estimation of state to state transitions in terms of cause specific hazards. However, in their data they have no information on reservation wages and they ignore the potential simultaneity between unemployment duration and reservation wages. This study does not have these limitations. The next section makes use of our data on reservation wages in the reduced form estimation of the transition equation of moves between unemployment and work.

5.2 Analysis of Duration Data using Hazard Functions

A person who positions themselves within the labour market may at some point during their working life experience a period of unemployment. These periods are referred to as spells. A single spell can be thought of as the length of time from the entry into

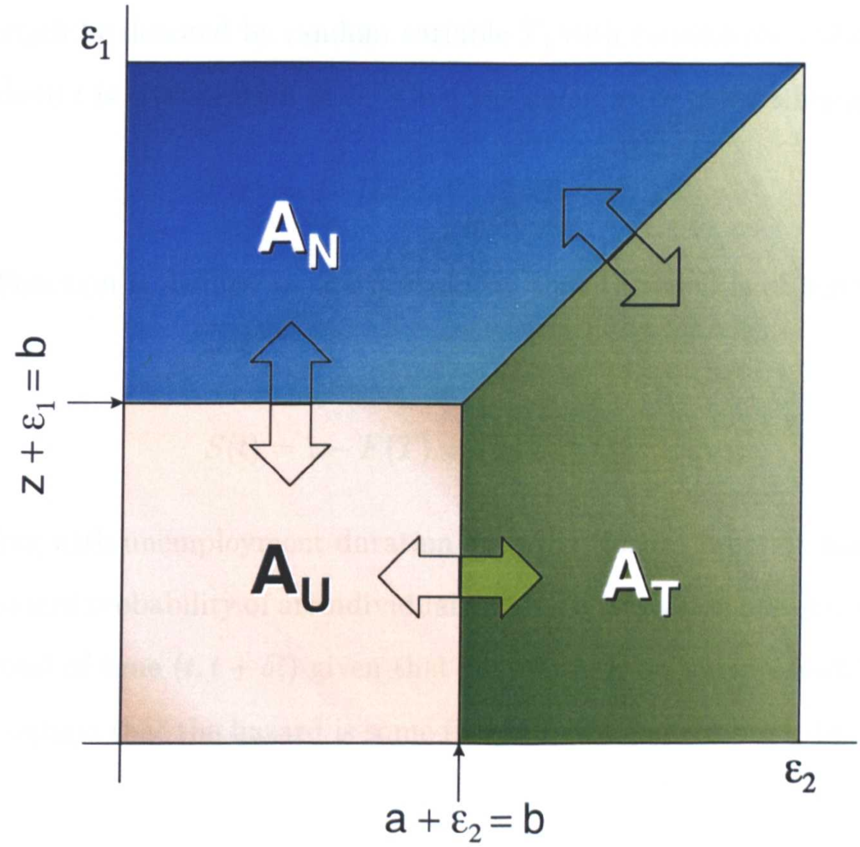

Figure 5.1: The 3-State Model of Labour Force Participation

unemployment until the person exits to another labour market state or the measurement is taken. Measurement of spell length can occur before the individual has exited.

During the course of this section we will highlight the use of Hazard functions when attempting to analyse survival data. Data for the duration of each YTS participants' spell of unemployment can be thought of as a set of survival data. An individual survives in the state of unemployment for the duration of their spell before exiting to a new state, such as employment. The hazard is defined to be the probability of a person exiting to such a state. If measurement occurs before an exit is achieved then we have a censored dataset. The YCS data which we will be investigating contains many individuals who were still unemployed at the time of the survey. All such people have a spell which continues up to the end of the survey and perhaps beyond. Any model should explicitly allow for the fact that the data are censored.

Let spell length be denoted by random variable T, with continuous probability distribution $f(t)$, where t is a realisation of T. Then the cumulative probability is

$$F(t) = \int_0^t f(s)ds = Pr(T \leq t). \tag{5.3}$$

The Survival Function is defined as the probability that the spell is of length at least t, that is

$$S(t) = 1 - F(T) = Pr(T \geq t). \tag{5.4}$$

When dealing with unemployment duration data the Hazard function maybe thought of as the conditional probability of an individual making a transition out of unemployment in a small interval of time $(t, t + \delta t)$ given that they have been unemployed for at least t periods. If we assume that the hazard is some function of t and denote it by $\lambda(t)$ then we have

$$h(t) = \lim_{\delta t \to 0} \frac{Pr(t \leq T < t + \delta t | T \geq t)}{\delta t}. \tag{5.5}$$

From the definition of conditional probability, we have

$$\begin{aligned} Pr(t \leq T < t + \delta t | T \geq t) &= \frac{Pr[(t \leq T \leq t + \delta t) \cap (T \geq t)]}{Pr(T \geq t)} \\ &= \frac{Pr[(t \leq T \leq t + \delta t)]}{Pr(T \geq t)} \\ &= \frac{F(t + \delta t) - F(t)}{Pr(T \geq t)}. \end{aligned}$$

Therefore,

$$h(t) = \lim_{\delta t \to 0} \frac{F(t + \delta t) - F(t)}{\delta t} \frac{1}{Pr(T \geq t)} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}. \tag{5.6}$$

There exists a direct relationship between the hazard and the probability distribution of T. This can be demonstrated via use of the Integrated hazard function

$$H(t) = \int_0^t \lambda(u)du = \int_0^t \frac{f(u)}{1 - F(u)} du. \tag{5.7}$$

Rearranging we arrive at

$$1 - F(t) = e^{-\int_0^t \lambda(u)du}.$$

Parametric models of the above have the advantage of simplicity but the restrictions imposed via their structural implications leads to possible distortions of the estimated hazard rates. Cox (1972) proposes a semi-parametric Proportional Hazard (PH) model which has the advantage of fewer restrictions. The model specifies that

$$\lambda(t_i) = e^{-\beta' x_i} \lambda_0(t_i) \tag{5.8}$$

where λ_0 is the baseline hazard and captures individual heterogeneity. This can be viewed as an individual specific constant. Hence the proportional hazard model does not include a constant term. The partial likelihood estimator allows us to estimate β without requiring estimation of λ_0. The conditioning operation is used to remove any heterogeneity. Central to the proportional hazards framework is this notion that all direct dependence of the hazard on duration is captured in the baseline hazard and so differences across individuals at duration point t depend on t only through regressor variation over time. If the regressors remaining constant over time, then so does the factor of proportionality. The direct dependence on t encapsulated by the baseline hazard can be thought of as a measure of any economic processes whose actions are not contained within the regressor function.

Suppose the sample consists of K exit times, $T_1, \ldots, T_K$. For any time T_i, we define the risk set R_i to consist of all individuals whose exit time from unemployment is at least T_i. The risk set with respect to any time T consists of all those individuals who remain unemployed prior to that time. For individual j in risk set $R_i, t_j \geq T_i$. Then the probability that an individual exits at time T_i given that one individual exits at time is

$$Pr(t_j = T_i | R_i) = \frac{e^{\beta' x_i}}{\sum_{j \in R_i} e^{\beta' x_j}}. \tag{5.9}$$

The conditioning has removed the heterogeneity. Under the simplest situation of one exit per period and no censoring, the partial log-likelihood is

$$\ln L = \sum_{i=1}^{K} \left\{ \beta' x_i - \log \left(\sum_{j \in R_i} e^{\beta' x_j} \right) \right\}. \tag{5.10}$$

When m_i individuals exit at time t_i, the contribution to the log-likelihood is just the sum of the terms for each of these individuals.

In the absence of any information on the baseline hazard, only the order of the durations provides us with information about our unknown coefficients. The partial likelihood framework allows us to easily accommodate our censored YCS dataset.

The job search/unemployment duration literature contains many variations on the basic unemployment duration model described above. An earlier study of the Restart Survey data by Dolton and O'Neill (1995) used a multiple exit framework. They distinguish between exits to employment, training and those into a state of non-participation in the labour force. Under a competing risks framework the exit time t_i is characterised by both a censoring indicator and an exit type indicator. K mutually exclusive and exhaustive exit types let C represent the exit type. At any given time the exit process in terms of K transition intensities is defined to be

$$h_k(t) = \lim_{\delta t \to 0} \frac{Pr(t \leq T < t + \delta t, C = k | T \geq t)}{\delta t}. \tag{5.11}$$

The total hazard rate $h(t)$ then equals the sum of all K transition intensities at time t, that is $h(t) = \sum_{k=1}^{K} h_k(t)$.

5.3 The Unemployment to Work Transition Process for the YCS Data

Research which makes use of empirical Two Stage Least Squares methods to analyse unemployment durations is limited by the analysis of only uncensored durations. Using the three state model introduced above in section 5.1, a more flexible empirical approach of the kind adopted by Burdett *et al.* (1984) and Gritz (1993) is possible. More specifically it is feasible to estimate transition functions between any pair of states in the model outlined in the previous section.

The approach taken by Burdett *et al.* (1984) is to estimate such a transition between the states U and N using assumed exogenous regressors and predicted wages. The predicted wage regressor is included to capture the possible effects of changing human capital or labour market employment prospects. Burdett *et al.* (1984) did not have reservation wage or expected wage information in their data (and it is unclear how they may have used it had it been available). In addition, the Burdett *et al.* (1984) paper was not concerned with evaluation of a government training programme on the unemployment to work transition.

5.3.1 Evidence from Cox Proportional Hazard Models Using an Unmatched YCS Dataset

As with previous chapters we again concentrate on the dual goals of uncovering the nature of the YTS treatment effect, whilst offering a critical assessment of the methods which we employ. As such we are only interested in datasets which include both treated and control/untreated observations. A dataset which excludes either one of these is lacks the necessary data to evaluate the counterfactual, which is vital to the determination of the YTS effect. Although the 3-state model allows for us to estimate transition equations for the any state to any other, we shall concentrate on the unemployment to work transition process and attempt to uncover any changes in this mechanism for those who have experienced a spell on YTS. The YTS to unemployment/work transition process is of less concern to us as it does not represent an opportunity to investigate the YTS treatment effect since there are no persons without YTS experience within such a dataset.

Again, we consider the YTS treatment effect to be different for the sexes. Hence, as with earlier work we present results for males and females separately. Table 5.1 contains proportional hazard model results estimated using male and female unmatched samples. The model specification allows us to concentrate on the YTS treatment effect whilst at the same time examine the robustness of the results to the dataset and matching procedure being used. Only individuals whose labour market state diary contained a period of unemployment which was characterised by an exit to work or censored were to be considered. With this in mind, the male sample consists of 1303 individuals and the female 1142 sample.

Looking at the results for ethnicity first; we included a number of dummy variables to represent the various categories of ethnicity from which, persons responded that they originated. The category 'White' was used as the reference group. We find no evidence to support the hypothesis that an individuals ethnic origin has an effect on the U to N transition process. Results for the effect of education suggest that, for both sexes, the probability of exiting from unemployment into work is enhanced for those with a higher level of educational attainment at 16. Of course, this represents our prior beliefs, since we would expect that, other things being equal, those with higher previous attainment will be more attractive to prospective employers. The size of the effect seems to be greater for females. Perhaps this is a reflection of the different kinds of jobs which the two sexes were likely to do during the period of study. It seems to suggest the females were more likely to take work which required the individual to have attained a higher level of educational qualification. Large numbers of males still had the option to undertake unskilled work, requiring no previous qualifications.

The next two effects we observe are only present for the male sample. LEA unemployment is seen to be negatively significant for males making the U to N transition. The lack of a significant effect for the female sample is perhaps another reflection of the area of the job market into which they were likely to enter. Women have traditionally found work in the service sector, and it is this area of the economy which emerged the healthier after the recession of the early 1980s. Expansion of the sector, combined with lower labour force participation rates amongst women could have acted to reduce the effect of unemployment rates on the female U to N transition process. Those males who had previously applied for a job are more likely to make the transition.

As with the models for the propensity score and the reservation wage we include a number of regional and cohort dummies. The reference groups being the "North" and cohort 1 respectively. A number of effects were observed in the estimated model.

The natural extension to the model outlined by Burdett *et al.* (1984) is to explicitly include a term in this reduced form estimation which attempt to capture the 'threshold utility' levels that a person in unemployment might have for entering the states of work or training. According to the 3-state Burdett model described above such reservation levels on the rewards for entering these states will explicitly influence the size of the acceptance sets for entering one of these other two states. Hence this logic suggests that the predicted reservation wage could be used to proxy for these reservation utility levels. Clearly it is necessary to use predicted levels of these reservation levels, since not every person in the data had occasion to report each of these values. These predictions are based on the reduced form estimations reported in Table C.5 in the Data Appendix. This variable is included in the unmatched dataset as a regressor in the estimated equations reported in Table 5.1. We see from this that the coefficient of the reservation wage variable is insignificantly different from zero in both the U to N transition equations.

Lastly, we consider the YTS treatment effect as measured in unmatched samples. Unlike in Chapter 3, we make immediate allowance for the heterogeneity of YTS. This seems reasonable given the evidence we uncovered for the differing effects of the three YTS incarnations during the work that followed. The months prior to the beginning of each individuals unemployment spell were scanned for the presence of any YTS spells. Once a spell was found we checked its' start date to determine which version of the scheme the participant had undertook and then set a dummy variable for that version of the scheme to 1, or 0 otherwise. The reference group were those individuals who had not experienced YTS.

Results indicate that males who undertook a period of YTSI were less likely to find work in any subsequent unemployment spell within the period of study. There was no evidence for such an effect amongst the female sample. This may be the result of a lack of females who undertook YTSI. Evidence for the effect of YTSII is consistent across the sexes. Both males and females are less likely to make the transition from U to N during a period of unemployment if they had previously experienced a spell of YTSII. Neither of the coefficients of YT experience for males or females are seen to be significant.

All of these results seem to support anecdotal evidence for the stigmatisation of those who had participated in both YTSI and YTSII by prospective employers. Whilst perhaps the lack of a negative effect for YT is the result of an improvement in both the scheme and employers perception of it following the introduction of vocational qualifications to the programme. However, much as with the early results for Chapter 3 these results were obtained using a dataset which had not been adjusted to make allowance for self-selection into YTS. We propose to examine the robustness of these findings to adjustments made using alternative methods of matching.

5.3.2 Evidence from Cox Proportional Hazard Models Using a Nearest Neighbour Matched YCS Dataset

We now move on to consider Cox proportional hazard models, estimated using matched datasets. We just run a Cox partial ML estimator on a single dataset containing n_t treatment cases and the same number of matched controls.

Before the matching process could begin the correct subsample of the dataset needed to be selected. As with the estimation of the U to N transition model using unmatched samples, only persons whose labour market state diary contained a spell of unemployment which was either censored or exited to work were considered. These individuals were separated by sex. The resulting two datasets were then cut to only include those persons with a complete set of U to N transition model dependent variables. Nearest neighbour Mahalanobis matching of the type already used in Chapter 3[1] was then performed. This process culminated in the production of male and female matched datasets. The number of treatment and control individuals now present were equal within each gender. Hence, we now had a male dataset which contained a treatment group of 572 persons and a matched control group consisting of a single individual for each of these, for a grand total of 1144 men. The female matched sample now consisted of 540 treatment and 540 control

[1] See section 2.5 for an explanation of the method and its assumptions.

persons for a total dataset of 1080 women.

Results for the estimation of the U to N transition using the matched datasets are presented in Table 5.2. As with the results for unmatched samples we find no evidence for the existence of an effect on the transition of those males and females who responded that they were from an ethnic background other than white.

In contrast to the results for the unmatched male sample we now observe a positive effect for the number of YTS training places. This may indicate that in areas where YTS provision was high the number of high quality placements was correspondingly higher and as such many males participants in YTS were effectively apprentices and may have been unaware of their YTS status. The lack of a effect for YTS provision for the female U to N transition again suggests that the mechanisms by which YTS affected males and females are different. It would be helpful to know the breakdown of YTS provision by employment sector. It could be that in areas with high YTS provision, the majority of that provision was in sectors where the work force was traditionally dominated by male workers. We could also postulate that YTS provision was of less use to the female labour force. This appears to run parallel to the initial purpose of the YTS scheme which was, at least in part, a measure designed to fill the gap in young male training vacated by the declining manufacturing sector and its' system of apprenticeships.

Our a priori belief for the existence of a negative unemployment effect for both males and females was only partially realised when we estimated the U to N transition models for the unmatched dataset. LEA unemployment proved to be negatively significant for the male transition process. We observe this effect for the nearest neighbour matched dataset with a similar magnitude. We now also see weak evidence for the existence of a regional youth unemployment rate effect. The coefficient of this variable is negatively significant at the 10% level for the male sample. Once again there is no evidence to support the hypothesis of a female unemployment rate effect. Once again we see a positively significant coefficient for those males who had previously applied for a job, the magnitude of which corresponds to that seen using unmatched samples.

Although all of the significant regional and cohort effects which we previously observed appear here, with the same signs and similar magnitudes, a number of new effect are also present. Cohort dummies 4 and 5 are now significant at the 10% level. An examination of the unmatched sample shows these two coefficients to have been bordering on this level of significance previously. Interestingly the coefficient of 'Yorkshire and Humberside' for males is now significant at the 5% level. There was no evidence for such an effect when using unmatched samples.

Once more reservation wages fail to capture the "threshold utility" levels that a person in unemployment might have for entering the state of work. As such, we find no evidence for an effect on the U to N transition for males or females when using the reservation wage as a proxy for individuals "threshold utility".

Evidence for the YTSI, YTSII and YT treatment effects remain consistent with the findings for the unmatched dataset. Again, we see a negative effect for males who experienced YTSI, although this is now only significant at the 10% level. A look back to the same coefficient in Table 5.1 for males reveals that the previous result was borderline significant at the 5% level, indicating that this result is still consistent. All three significant effects are of similar magnitudes to those of the unmatched sample.

5.3.3 Evidence from Cox Proportional Hazard Models Using a Local Linear Matched YCS Dataset

We now present results from of Cox proportional hazrd models using local linear matched datasets. Hence the male sample contains n_t males from the treatment group and n_t "averaged" matched control people. These datasets were created using the local linear method as described in subsection 2.5.3.

We now propose to take the concept of the synthetic control group a stage further. Before we could attempt an analysis using a local linear regression matched dataset, we needed to work through the problems which arise as a result of the lack of covariate data following a conventional local linear matching procedure. In the same way that we had generated synthetic elasticities and reservation/expected wages[2] using the weights as calculated using Equation (2.41) we were now able to generate synthetic values for all the covariates which we would need to proceed with a Cox Proportional Hazard Models analysis of a local linear regression matched dataset. In essence we had created a set, of size n_t (the number of persons in the treatment group), of synthetic control persons who "possessed" synthetic values for their covariates calculated from a combination of the unmatched control group covariates and the local linear regression weights matrix. The variables were now ready for use in our Cox Proportional Hazard Models analysis.

As with the presentation of the local linear matching results for the YTS treatment effects on job search elasticities and wages we suggest that results can be sensitive to the choice of the kernel bandwidth parameter, h_n. In an attempt to observe any fluctuations in the magnitude and significance of the coefficients we present our results over the range of h_n. Chapter 4 contained a series of plots for the migration of job search elasticities and wages as the bandwidth parameter, h_n, varied. We now present our results for the local linear matched dataset U to N transition model in a series of plots (Figures 5.2 to 5.7) for each of the key variable coefficients over the range $0.1 \leq h_n < 1$. Each of these plots is constructed such that h_n is represented by the x-axis. There are then two y-axes; the left-hand y-axis measures fluctuations of the coefficient, whilst the right-hand y-axis measures fluctuations in the t-stat for that coefficient. The blue plot line represents the coefficient and the red plot line represents it's t-stat. The range of the t-stat axis which we report covers only the region 0 to 0.1, since this corresponds to the area in which a coefficient is significant. Hence it does not appear on the graph in Figure 5.2 if it falls outside of this range. Regions of the y-axis over which the red line is missing correspond to regions of no significance. At the same time we present Table 5.3, a model evaluated using a dataset matched at the iterative optimal bandwidth for the purposes of comparison with Tables 5.1 and 5.2.

Examination of the results in Table 5.3, for the male local linear matched sample

[2]See subsection 4.4.2 for an explanation of how this was done.

reveals that, as with those for the unmatched and nearest neighbour datasets, there is no evidence for the existence of any ethnicity effects. Plots of the coefficients for the variables "Black", "Asian" and "Other Ethnic Origin", over the range of h_n are presented in Figure 5.2. All three reveal that although the magnitudes of the coefficients vary over this range, at no point do any of them become significant. Results for the female dataset contrast with those seen using the two previous female samples. Unlike the results for the unmatched and nearest neighbour matched female datasets, the U to N transition model estimated at the optimal bandwidth using the female local linear matched sample contains evidence suggesting that female Asians and those who placed themselves in the category "Other Ethnic Origin" are less likely to move from unemployment to work. Figure 5.5 contains the plots for the coefficients of these two variables over the range of h_n. Notice just how sensitive some of these results can be to the choice of h_n. Taking the plot for the coefficient of "Asian" first we observe that the negatively significant result in Table 5.3 is only significant over the approximate range $0 < h_n < 0.2$, after which, a movement of the coefficient causes it to enter the region of no significance, where it remains for the approximate range $0.2 < h_n < 1$. The sudden shift in the red t-stat line between $h_n = 0.2$ and $h_n = 0.3$ is testament to the sensitivity of this coefficient to the kernel bandwidth.

Both the unmatched and nearest neighbour matched male and female datasets produced models with significant positive coefficients for the education score at 16. Table 5.3 shows that the coefficient of the variable 'Education Score at 16' for local linear matched males, calculated at the optimal bandwidth parameter, is insignificant. The plot of this coefficient in Figure 5.2 shows that although it's magnitude varies over the range of h_n, at no point does it become significant. In contrast, the results for the female education score variable in Table 5.3 are consistent with previous findings. Unemployed females with higher educational attainment do appear to find it easier to make the transition to work. The plots of this coefficient and it's t-stat contained in Figure 5.5 are interesting in that the variable remains significant over the range of h_n, with the value of the t-stat remaining close to 0 at all times. However, we do observe some fluctuation in the magnitude of the coefficient around the region $h_n = 0.2$ to 0.5.

Results for the local labour market variables within the male local linear matched sample are stronger than those seen when estimating the two previous male U to N transition models. The coefficient of the variable for regional YT provision within the unmatched sample was insignificantly different from zero. The same coefficient, when estimated using the nearest neighbour male sample was positively significant. Once again we observe a positive significance for this variable. The plot for this coefficient and it's t-stat is contained within Figure 5.3. The red t-stat line is present for low values of the bandwidth before moving outside of the range of significance as the magnitude of the coefficient begins to fall. Higher values for the bandwidth lead to a matched dataset which produces a model in which magnitude of this coefficient begins to rise again. As this happens we see the red t stat line return to the area of significance. Any conclusions as to the effect of regional YT placement provision on the rate of males making the transition from unemployment to work made using these results would be dependent on the choice of the kernel bandwidth.

Neither of the two previous datasets produced a model which contained a significant result for the effect of regional YT provision on the female U to N transition. Results for the female local linear matched dataset both support and contradict these findings. Again this is due to fluctuation in the magnitude and in consequence the t-stat of the coefficient of this variable. Figure 5.3 contains the plots for this variable. The coefficient is insignificant until $h_n > 0.35$, at which point, the red t-stat line appears and remains until $h_n = 0.5$ It the reappears for $h_n > 0.65$. Clearly both the male and female local linear matched samples are extremely sensitive matching specification.

Evidence from models estimated using the unmatched and nearest neighbour datasets had suggested a relationship between male LEA unemployment rates and the rate of exit from unemployment to work. Results for females had failed to uncover any such connection. The coefficient of LEA unemployment for male and female U to N transition models are now seen to be negatively significant at the optimal bandwidth. A look at both the male and female plots for the coefficients and t-stats of this variable, in Figures 5.3 and 5.6, show that both are significant over the entire range of h_n. Notice also that the form of the variation in the magnitude of these coefficients for both males and females is the same; most of the variation occurs around the $0.2 < h_n < 0.4$.

Previous results had not indicated a link between youth unemployment rates and the transition of unemployed males and females into work. The male transition model using local linear matched data now contains a negatively significant coefficient for this variable. Figure 5.3 contains the plot for this coefficient. As h_n increases, so the magnitude of the coefficient falls until around $h_n = 0.3$, when it ceases to be significant. Results for the female local linear matched dataset were in line with those using the two previous datasets.

The variable "Ever Applied for a Job" had shown mixed results for the unmatched and nearest neighbour matched datasets. It's coefficient was positive for males using both. Neither model had suggested such a relationship for female job seekers. Figure 5.3 contains the plot for the male coefficient of this variable. At no point does the red t-stat line appear. In contrast, the female coefficient plot of Figure 5.6 shows the coefficient become negatively significant around $h_n = 0.3$.

We do not present a full investigation of the migration of the coefficients of the regional and cohort dummy variables since these were only included to control for variations in the dataset resulting from differing local labour market conditions and temporal changes in the economic environment into which, young school leavers were exposed.

As was stated at the beginning of this whole section, we included a predictor of reservation wages with the hope that it might capture the 'threshold utility' levels that a person in unemployment might have for entering the states of work or training. The male and female populations, from all three datasets used, produced U to N transition models with insignificant coefficients for this variable. However a look at the plot for the male local linear matched sample, in Figure 5.4, does reveal a move to positive significance for the coefficient when matching was performed using values of $h_n > 0.45$. At the same time the coefficient almost doubles in magnitude. The equivalent plot for the female dataset,

Figure 5.7, contains no regions of significance.

We now consider the effect of all three versions of YTS on the U to N transition. Taking the coefficient of YTSI first, Table 5.3 shows this to be negatively significant for males and insignificant for females at the optimal bandwidth. These results reflect those for the unmatched and nearest neighbour samples. The plot for the male YTSI coefficient, Figures 5.4, indicates that the effect is strongly negatively significant over the entire range of h_n. However, it's magnitude does decline as h_n rises. The female YTSI coefficient plot contains no evidence for a significant female YTSI effect.

As with YTSI, results for YTSII are in line with previous findings. Results suggest that the unemployed of either sex, with YTSII experience, were less likely to exit to work. In contrast to the male YTSI coefficients, the magnitude of the effect grows as h_n rises. The female effect reduces in magnitude as h_n rises.

Finally, we consider the effect of YT on the U to N transition. None of the three male datasets uncovered an effect for YT. Figure 5.4 contains the plot for the coefficient of this variable reveals that this remains the case over the whole range of h_n. The unmatched and nearest neighbour models did not suggest a relationship between female YT experience and the time to exit from unemployment into work. Table 5.3 suggests that females with a spell on YT maybe less likely to exit. The plot for this coefficient contains reveals that, as with several of the results in this subsection, this result is sensitive to the choice of h_n. We do not observe the red t-stat line around the region of $0.25 < h_n < 0.45$.

Regressors	Males		Females	
	U → N	Standard Error	U → N	Standard Error
Black	-0.3010644	0.4655724	-0.2622033	0.2825738
Asian	0.0991228	0.2796618	-0.1381936	0.2774568
Other Ethnic Origin	0.1086571	0.1736851	-0.2241836	0.1844621
Education Score at 16	**0.0175384**	0.0062921	**0.0239475**	0.0060667
Regional YT Places	7.177268	5.012926	4.351977	5.279642
LEA Unemployment	**-0.0321242**	0.0115167	-0.0113013	0.011341
Regional Youth Unemployment	-1.92945	1.69107	-1.778768	1.809986
Ever Applied for Job	**0.1863709**	0.0781305	0.0335759	0.0719602
Yorkshire & Humberside	0.1959552	0.1370854	0.1474534	0.1369818
East Midlands	**-0.270313**	0.1174874	**0.2945252**	0.1173168
East Anglia	-0.0621847	0.1978639	-0.0019859	0.2173162
Greater London	0.2214141	0.1863598	0.1502816	0.2081107
South East	-0.1656823	0.2522709	0.3690366	0.2470798
South West	**0.6610613**	0.3731564	**0.8151292**	0.3753839
West Midlands	0.0876336	0.252259	**0.4720825**	0.2615014
North West	0.0728373	0.2121285	**0.4301338**	0.2295347
Wales	-0.0167016	0.1376586	**0.2759921**	0.1586849
Cohort 2	0.0171707	0.1443004	-0.0230942	0.1514039
Cohort 3	0.0279374	0.1697276	0.138981	0.1867366
Cohort 4	0.303183	0.1902795	0.3244546	0.2078507
Cohort 5	0.2839207	0.251272	0.4290856	0.2696632
Cohort 6	-0.2695125	0.2106865	0.2074527	0.1957924
Predicted log of Wage	-0.1231428	0.6758956	0.0048022	0.7800348
YTS I before Unemployment Spell	**-0.3204607**	0.1620362	-0.1159314	0.2008059
YTS II before Unemployment Spell	**-0.3412091**	0.1095644	**-0.3723779**	0.1198777
YT before Unemployment Spell	0.1821303	0.2574848	0.107904	0.2484242
χ^2	96.77		102.44	
Pseudo R^2	0.0059		0.0073	
Log Likelihood	-8149.1472		-6994.4895	
Sample Size	1303		1142	

Table 5.1: Cox Proportional Hazard Model for the U to N Male and Female Transition Process (Unmatched Dataset)

Regressors	Males U → N	Males Standard Error	Females U → N	Females Standard Error
Black	-0.3348777	0.4686046	-0.2061529	0.3401145
Asian	0.0989408	0.2809493	-0.178762	0.2779017
Other Ethnic Origin	0.0809268	0.1811238	-0.2476814	0.1873954
Education Score at 16	**0.0140513**	0.0066223	**0.0234838**	0.0062533
Regional YT Places	**11.31041**	5.712521	2.576579	5.412354
LEA Unemployment	**-0.0292295**	0.0124654	-0.0127602	0.0116797
Regional Youth Unemployment	**-3.381178**	1.985636	-0.545921	1.937689
Ever Applied for Job	**0.2071765**	0.0855468	0.0384625	0.0734508
Yorkshire & Humberside	**0.2881552**	0.1433614	0.1556714	0.1399598
East Midlands	**-0.3123568**	0.1285754	**0.3177136**	0.1204551
East Anglia	0.0146096	0.2180934	-0.0336035	0.2210286
Greater London	**0.3838412**	0.2098404	0.1570495	0.2121221
South East	0.033044	0.2785128	0.3473323	0.2533353
South West	**0.8652406**	0.4236607	**0.7421164**	0.3827568
West Midlands	0.2963857	0.286848	**0.4571354**	0.2651298
North West	0.2365539	0.2344751	**0.4081132**	0.2347613
Wales	-0.026711	0.1481314	**0.2927429**	0.1623819
Cohort 2	-0.0395245	0.150275	0.0067976	0.1531237
Cohort 3	0.0250025	0.1925307	0.183334	0.1916085
Cohort 4	0.3065541	0.2116407	**0.3669918**	0.2129891
Cohort 5	0.2722826	0.2592821	**0.5043142**	0.2751833
Cohort 6	-0.2860503	0.2244709	0.2338126	0.199947
Predicted log of Wage	-0.0072393	0.7043089	0.1485837	0.7982437
YTS I before Unemployment Spell	**-0.2888603**	0.1635088	-0.1216007	0.2013769
YTS II before Unemployment Spell	**-0.354268**	0.110578	**-0.3704267**	0.1203834
YT before Unemployment Spell	0.1674214	0.2668467	0.1245818	0.2494195
χ^2	97.71		96.00	
Pseudo R^2	0.0069		0.0073	
Log Likelihood	-7001.4393		-6555.4162	
Sample Size	1144		1080	

Table 5.2: Cox Proportional Hazard Model for the Male and Female U to N Transition Process (Nearest Neighbour Matched Dataset)

Regressors	Males		Females	
	U → N	Standard Error	U → N	Standard Error
Black	-0.6415866	0.4848389	0.0345284	0.5073386
Asian	0.538299	0.4043378	**-1.825117**	0.6301077
Other Ethnic Origin	-0.1539133	0.3232332	**-0.8306165**	0.3378005
Education Score at 16	-0.0114904	0.010881	**0.0542092**	0.01028
Regional YT Places	**14.22831**	6.238399	11.88717	8.663363
LEA Unemployment	**-0.0677621**	0.0179002	**-0.0684307**	0.0172047
Regional Youth Unemployment	**-6.988046**	2.752259	1.81104	2.913894
Ever Applied for Job	0.0822098	0.1376205	-0.0280611	0.142718
Yorkshire & Humberside	0.212276	0.1529093	**-0.5604344**	0.1900364
East Midlands	**-0.4488667**	0.195606	-0.255107	0.175739
East Anglia	0.150119	0.227893	-0.2014418	0.3239725
Greater London	0.1219509	0.1644689	0.0454197	0.3114626
South East	-0.5371308	0.3598346	**1.175755**	0.3863077
South West	0.631213	0.4044976	0.5719454	0.6021811
West Midlands	0.3146283	0.2142097	0.119334	0.4039506
North West	0.07504	0.2412238	0.2270781	0.3585024
Wales	0.0698121	0.1880825	-0.2758991	0.2592864
Cohort 2	**0.3719814**	0.1476243	**0.9178647**	0.2370152
Cohort 3	**-0.3499463**	0.1637918	**1.496143**	0.2713786
Cohort 4	0.1761424	0.2219662	**1.55833**	0.3213547
Cohort 5	-0.0492508	0.3314125	**1.657248**	0.3899189
Cohort 6	**-0.5645457**	0.2921989	**1.876665**	0.2847118
Predicted log of Wage	1.163237	0.8785303	0.000667	1.379391
YTS I before Unemployment Spell	**-0.685633**	0.1695052	-0.1007095	0.215254
YTS II before Unemployment Spell	**-0.6487357**	0.1299508	**-1.12893**	0.1427642
YT before Unemployment Spell	0.1497676	0.3258626	**-0.6500879**	0.3035491
χ^2	231.56			268.53
Pseudo R^2	0.0166			0.0206
Log Likelihood	-6843.473			-6374.1187
Sample Size	1144		1080	

Table 5.3: Cox Proportional Hazard Model for the Male and Female U to N Transition Process (Local Linear Matched Dataset)

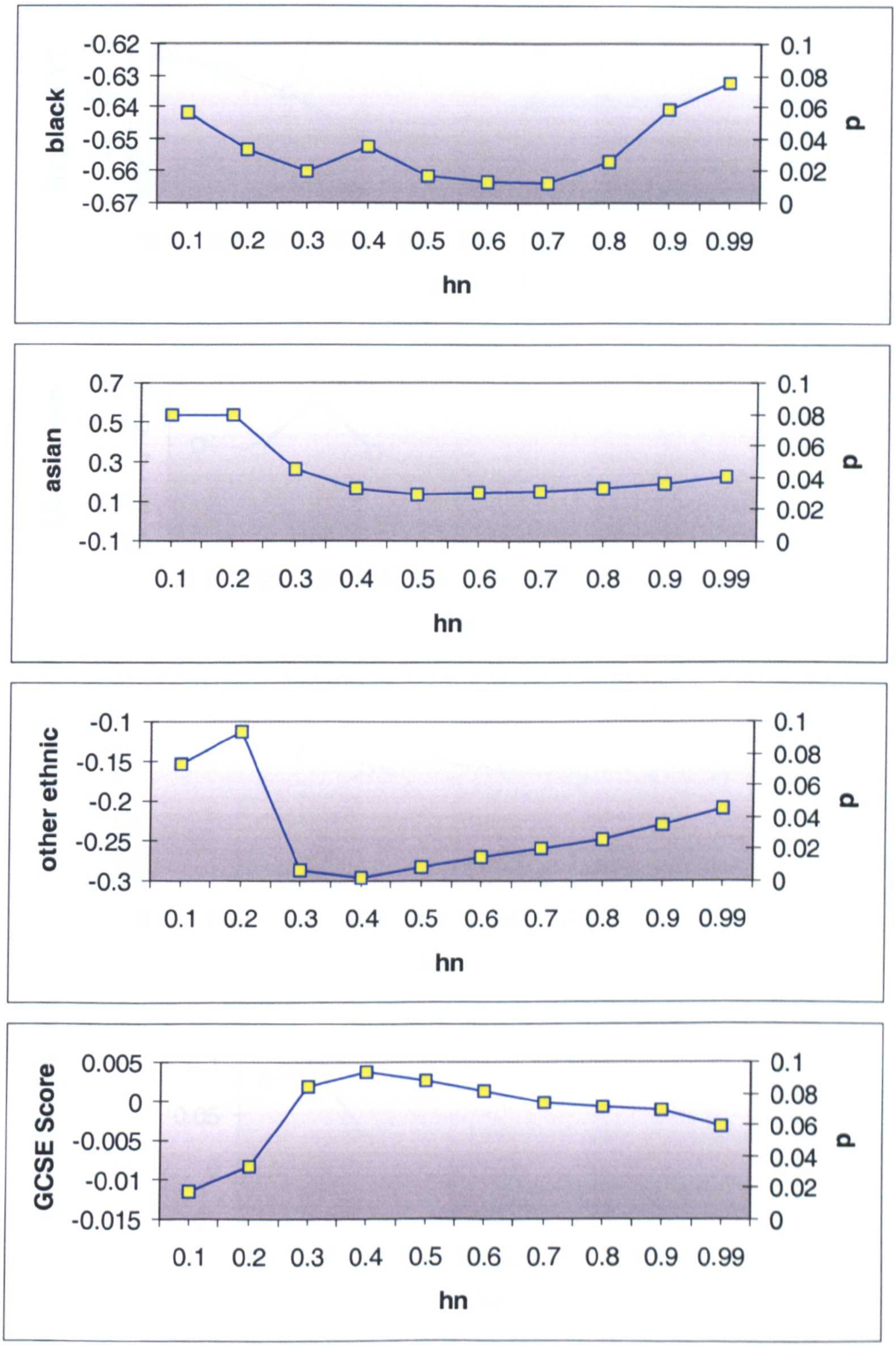

Figure 5.2: Coefficients of Independent Variables in the Cox Proportional Hazard Model for the Male U to N Transition Process (Local Linear Matched Dataset)

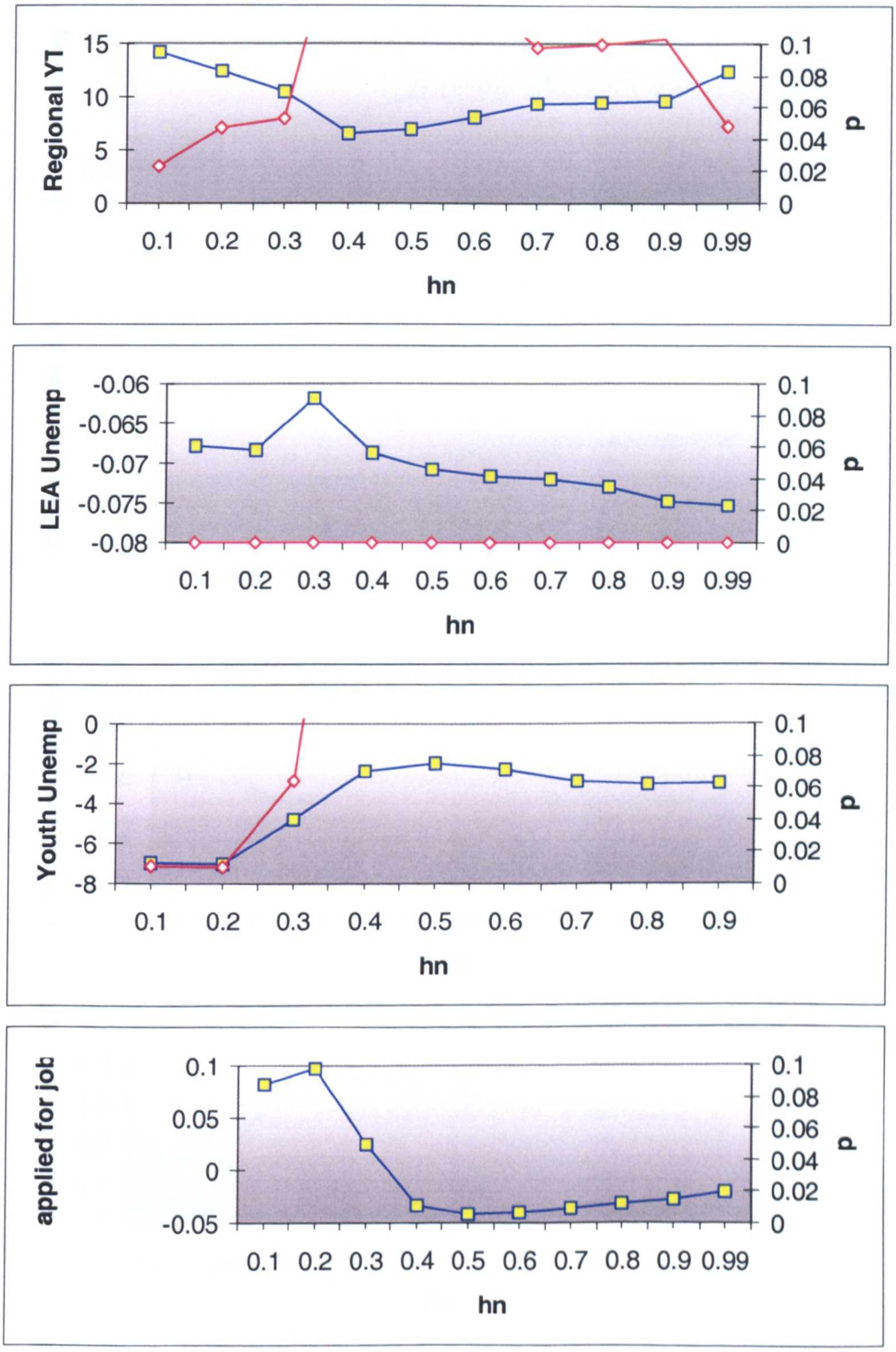

Figure 5.3: Coefficients of Independent Variables in the Cox Proportional Hazard Model for the Male U to N Transition Process (Local Linear Matched Dataset)

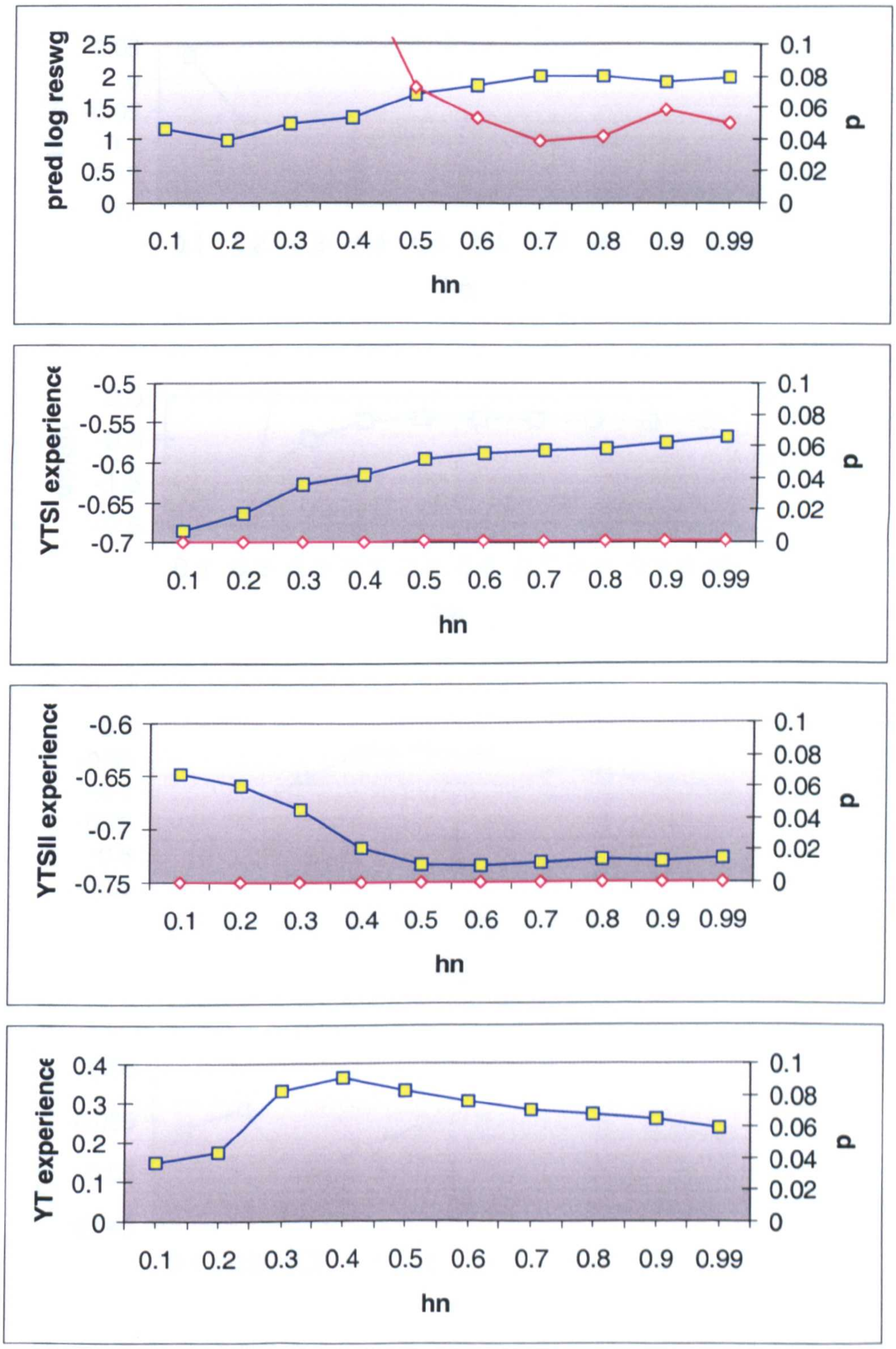

Figure 5.4: Coefficients of Independent Variables in the Cox Proportional Hazard Model for the Male U to N Transition Process (Local Linear Matched Dataset)

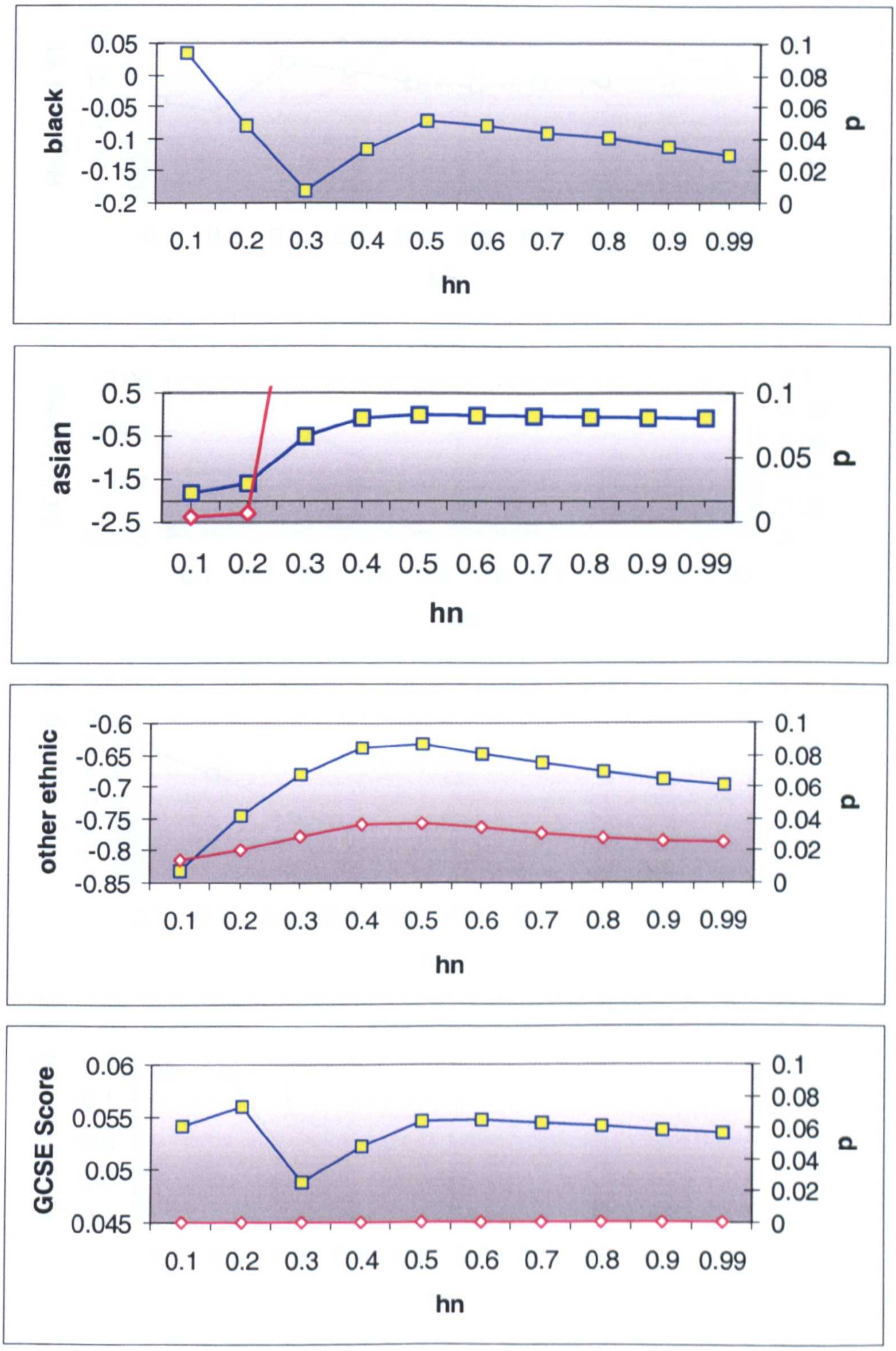

Figure 5.5: Coefficients of Independent Variables in the Cox Proportional Hazard Model for the Female U to N Transition Process (Local Linear Matched Dataset)

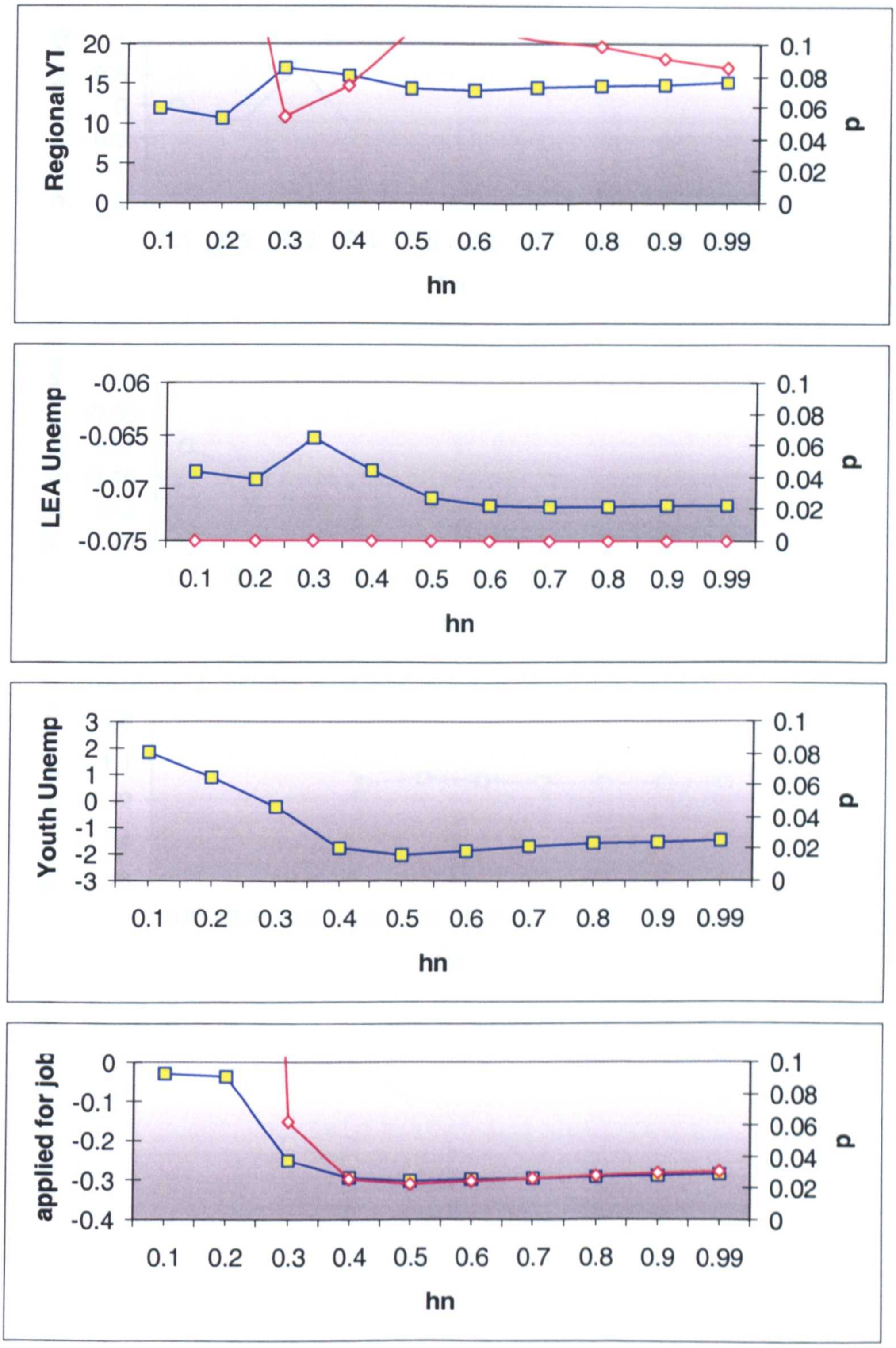

Figure 5.6: Coefficients of Independent Variables in the Cox Proportional Hazard Model for the Female U to N Transition Process (Local Linear Matched Dataset)

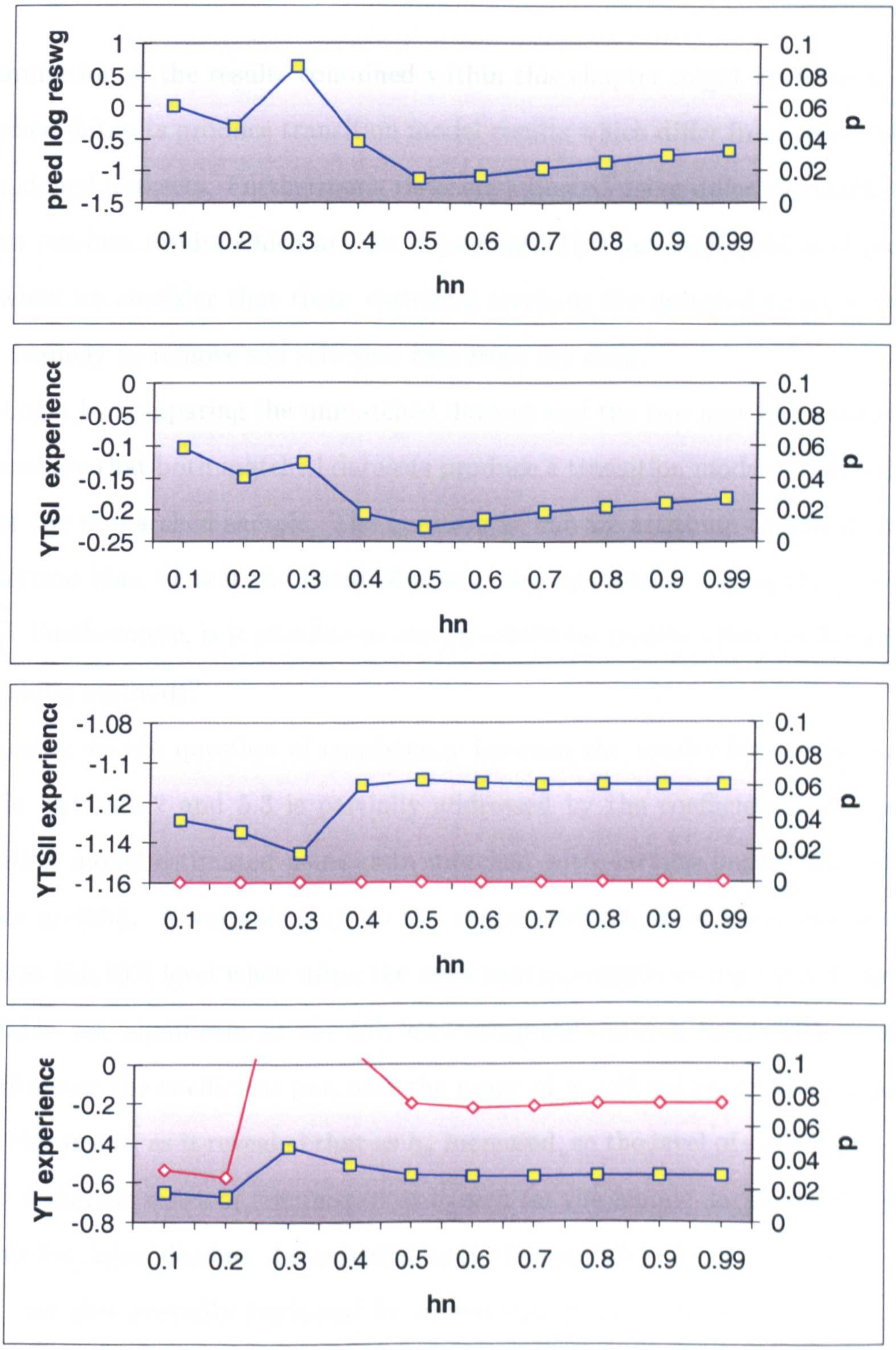

Figure 5.7: Coefficients of Independent Variables in the Cox Proportional Hazard Model for the Female U to N Transition Process (Local Linear Matched Dataset)

5.4 Conclusions

First examination of the results contained within this chapter might lead one to believe that matched datasets produce transition model results which differ from those estimated from unmatched datasets. Furthermore, datasets adjusted using differing matching algorithms can produce results which are not consistent. This last statement is of particular concern when we consider that these matching methods are designed to solve the same problem. Namely to remove self selection bias from the data.

If we begin by comparing the unmatched dataset and the two matched datasets, then we can conclude that both matched datasets produce a transition model which is different to that for the unmatched sample. The question is, can we attribute this to the removal of self selection bias, or is it the result of a new bias introduced during the processes of matching? Furthermore, is it possible to obtain consistent results when employing two or more matching methods?

The answer to the question of consistency between the results for the two matched datasets in Tables 5.2 and 5.3 is partially addressed by the coefficient/t-stat plots for the transition model estimated using data matched with varying bandwidths using the local linear method. A variable such as the youth unemployment rate, was seen to be significant at the 10% level when using the male nearest neighbour matched dataset. The same variable was significant at the 5% level using the optimal bandwidth local linear dataset. However the coefficient plot over the range of h_n offered evidence to explain the change in this result, as it revealed that as h_n increased, so the level of significance for this effect fell. A similar series of circumstances is seen for the female dataset with regard to the variable for Asian descent. Inconsistencies in the result for the variable "ever applied for a job" are also partially explained by movements in and out of significance for the female coefficient plot.

We included the predicted reservation wage with the hope that it would capture the threshold utility levels between the different labour market states. However neither table contained a significant coefficient for this variable. This however is not the end of the story. The male plot for this coefficient shows us that as the bandwidth rises so a positively significant coefficient emerges.

Finally, we see that the female nearest neighbour dataset produced a model which did not contain a significant result for the YT variable. This was in contrast to the findings for the model estimated using the optimal bandwidth local linear matched data. Yet again, an examination of the plot for this female coefficient reveals that it moves in and out of significance over the range of h_n.

When considering the nearest neighbour algorithm separately we conclude that wider callipers lead to convergence in the coefficients of models run using these matched datasets. This follows intuitively since as the caliper width approaches 1, the set of controls which can be matched approaches the total number of controls. Similarly, as the bandwidth of a kernel based method approaches 1, so the set of controls to be weighted for the creation of a synthetic match approaches the total number of controls.

A deeper examination of results from the two matching regimes has shown us that many of the their apparent differences are less visible as the bandwidth of the kernel rises. We would suggest that this is a reflection of our adjustment of the nearest neighbour algorithm to not exclude anyone from the treatment group, even if they lacked support, since this is akin to widening the bandwidth for the local linear algorithm Hence, results for the local linear dataset converge to those of the nearest neighbour data as h_n increases. In conclusion, we have evidence to suggest that given a suitable framework, differing matching algorithms of the type considered in this work, can produce results which are consistent with one another.

Chapter 6

Conclusions

In the introduction to this work we broached two points for discussion. Firstly, we asked what was the YTS treatment effect. We wondered whether the scheme, in its different versions, had succeeded in meeting the Governments requirements of it. Secondly, we acknowledged some of the difficulties inherent in an empirical study of this question and introduced the concept of matching as an answer to the problem of self selection into YTS. When using matching methods it is important to question of the performance of these algorithms and whether consistent results could be produced when using them.

6.1 On the Question of Matching Algorithm Performance and Consistency

If all matching methods produced the same matched dataset, there would be no need for researchers to use any but the simplest, most convenient method. In reality, since each method of matching gives rise to a different "quasi-control" group it is important to compare them. The methods which we have considered during this work all have their strengths and weaknesses. The nearest neighbour algorithm benefited from the fact that it matched real people to one another. As well as being intuitively simple, nearest neighbour methods produce matched samples containing individuals with full sets of covariates. A feature which greatly simplifies any analysis which might follow. However, with datasets containing areas of poor support and depending on whether we choose to match in spite of this or exclude those without support, such algorithms can produce either poor matches or biased datasets. Kernel based algorithms such as the kernel and local linear methods have the advantage that they reuse the dataset by weighting individuals within the control set by how similar they are to each person in the treatment set. The Achilles heal of such methods lies in the need for researchers to select the bandwidth parameter prior to matching. Silverman (1986) has suggested the use of an optimal bandwidth. However, the nature of the dataset which we employed made this difficult to estimate. Kernel methods also suffer from the need to synthesise the covariates which we wish to employ in our post match analyses.

The results which we have presented during the course of this work provide evidence to suggest that we cannot rely on matching methods as "magic box" solutions. The nature of each dataset can exacerbate the various weaknesses of each method. Rendering it unwise to perform a match with a given algorithm. It can never be enough to simply submit a dataset for input to an algorithm and then perform an econometric analysis on the output. Although the Rosenbaum and Rubin (1983) result of Chapter 2, section 2.4 provides a theoretical framework that justifies methods that match using propensity scores, the assumptions which underpin it, especially those which relate to the observation of a complete set of X variables, for the whole of the sample are limiting given the nature of real world data. Clearly, propensity score models, containing X variables which capture the nature of the propensity for treatment are needed. However, the greater the number of independent variables used, in a model for propensity, the smaller the sample available to us. This is especially true of exogenous variables such as local labour market data, which are difficult to accurately merge with the endogenous data obtained from survey respondents. The tradeoff which a researcher must face may determine the stress to which assumptions are subjected.

We have also highlighted the issue of "region of overlapping support". We tackled the question of whether a researcher must make choices about the omission of persons whose propensity scores lie outside of the region. We voiced our concern for results obtained using datasets matched with some treatment persons excluded. We demonstrated that given that exclusions for non-overlapping support were likely to occur without uniform probability throughout the group, it followed that any treatment effect subsequently uncovered would represent the effect for those treatment persons who remained in the sample and not the effect for the dataset as a whole. Such exclusions did not effect the validity of the Rosenbaum and Rubin (1983) result. However, in most cases the researcher must be aware that their actions lead to an exclusion bias, which can prevent them from answering the question they most frequently pose, 'What is the mean treatment effect on the treated?'

6.2 On the Question of YTS Performance

We began our study of the YTS treatment effect in Chapter 3, with the belief that the males and females YTS experience had been very different. In order that we might capture these differences all subsequent analysis was performed separately for males and females.

As was acknowledged during our early analysis of the YTS treatment effect grouping together of all YTS types obscures the differences between these schemes. We identified three incarnations of the scheme which we labelled YTSI, YTSII and YT, and proposed to treat them separately during our investigations into the YTS treatment effect. From that point on, we considered these to be three kinds of YTS treatment effect. This gave us the opportunity to chart the progress of each scheme to determine whether later incarnations were more or less successful than their predecessors.

Early work considered the treatment effect on reservation/expected wages and various job search elasticities. Results for the nearest neighbour matched datasets were less dramatic than those for the unmatched sample. Perhaps suggesting that there had been a self selection bias present in the unmatched dataset. We found strong evidence for the presence of a YTS treatment effect.

Chapter 4 began to question the ability of the nearest neighbour algorithm to produce consistently good matches throughout the evolution of a matching 'run'. We demonstrated the degradation in match quality as a match progressed. We introduced the kernel and local linear methods of matching and investigated the YTS treatment effect on the reservation/expected wages and various job search elasticities using datasets matched with these methods.

Finally, Chapter 5 introduced the 3-state labour market model as a theoretical framework for the labour market state transition process. Given the government's original desire to train the young so that they might become more attractive to employers seeking skilled workers, we choose to concentrate on the unemployment to work transition. Cox proportional hazard models were then estimated for this transition. The effects of YTSI, YTSII and YT were captured through the inclusion of dummies for whether individuals had experienced a spell on such a scheme prior to their spell of unemployment.

6.2.1 The YTSI Effect

Taking the results for reservation/expected wages and the various job search elasticities first, the unmatched male dataset suggested a strong YTSI treatment effect. YTS was seen to depress expected wages whilst having mixed effects on the magnitudes of the various elasticities. However, we suggest that these results were the product of the inherent pre-scheme differences between those who had participated in YTSI and those who did not.

The effect of YTSI on wages and job search elasticities, as revealed in male and female nearest neighbour and local linear matched samples, was minimal. We uncovered some evidence to suggest the presence of a downward effect on wage expectations for both genders. Other things being equal this would have increased the chances of YTSI participants finding work post scheme. Results for the local linear matched male sample also suggested that YTSI had increased the magnitude of several job search elasticities.

Results from the male unmatched and matched datasets, for the unemployment to work transition process, suggested that YTSI had a detrimental effect on the probability of exiting from unemployment into work. This suggests a failure of the scheme in it's attempt to improve the employment prospects of participants and is evidence for the "scaring" which some have suggested such schemes can lead to. With reference to the previous paragraph, "other things" were clearly not equal. Female transition models estimated using all three datasets offered no evidence for a YTSI effect on the probability of transition.

6.2.2 The YTSII Effect

The effect of a spell on YTSII for males and females as estimated using the unmatched sample and measured on wages and job search elasticities was, as with the YTSI effect, mixed. Male reservation wages were seen to rise whilst the elasticities of both genders were reduced for scheme participants. As with our discussion of the YTSI effect, some of these "effects" may be the result of self selection bias.

Results for the nearest neighbour and local linear matched male samples suggest that there are a number of YTSII treatment effects on job search elasticities. Some desirable and some not so. The elasticities of the reservation wage to the benefit level, arrival rate of job offers are lower for those with YTSII experience. Indicating that these males were less likely to price themselves out of the labour market following changes in benefits or job vacancies. Results for the elasticities of the hazard with respect to benefits indicated that males with YTSII experience were more willing to take a job than those without as benefits rose. However, results for the elasticity of the hazard with respect to vacancies, was adversely effected by YTSII. This indicates that even when the number of positions on the market rose, those with YTSII experience lost out to those without. Evidence that YTSII may also have "scared" its participants.

The female nearest neighbour dataset produced estimates of control persons wages and elasticities which were insignificantly different from those of participants. The local linear female sample produced job search elasticities for the hazard with respect to benefits which suggested that female YTSII participants were less likely to remain unemployed following a rise in benefits. This scheme seems to have reduced female participants dependence on benefit payments through a maintenance of their willingness to take on a job in the face of rising benefits.

Transition models for unemployment to work for males and females using unmatched and matched samples revealed that, as with YTSI, its successor reduced the prospect of future employment on its participants. It seems that the desirable movements in many of the elasticities which we observed were not enough to offset the reduction in the hazard with respect to the arrival rate of offers. Again, evidence points to "scaring" of those with YTSII experience leading to employers choosing not to employ those, of either gender, with a work history containing a period on the scheme.

6.2.3 The YT Effect

The final incarnation of YTS, YT was much vaunted due to it's emphasis on the attainment of vocational qualifications. As such it is interesting to see whether this recognition of new skills allowed those with experience of YT to demonstrate to employers the enhanced skill level which they obtained from the scheme.

The unmatched male wage and elasticity results indicated that men with YT experience had higher required higher wage rates and reduced magnitudes of the various elasticities. There was little evidence for a female YT treatment effect. As with both previous scheme analyses we considered there to be a self selection bias component to these results.

Nearest neighbour matched male samples produced evidence of a wage effect. It appears that males with YT experience had higher mean wage expectations. This result was not replicated using the local linear male sample. If present, then there was some evidence to suggest that this effect was offset by a reduction in the elasticity of the reservation wage with respect to the benefit level. Both matched male datasets produced results for the elasticities of the hazard with respect to benefits showing that the rate of reduction in the hazard was lower for those with YT experience. The results for the elasticity of the hazard with respect to vacancies differed in direction of significance across matching methods.

The female nearest neighbour matched dataset gave a value for reservation wages of those women with YT experience which was significantly lower than for those without. However, reservation wages are more sensitive to benefits and the number of job vacancies. Also, we uncovered some evidence to suggest that those women with YT experience were less likely to leave the state of unemployment as benefits rose. Finally, we see that the hazard with respect to the benefit level was is larger post scheme. Perhaps this is evidence of the value of the qualifications which YT participants could now achieve.

Turning to the YT effect on the unemployment to work transition process we find no evidence for a YT effect on the likelihood of males making the transition. This might help to explain the mixed results for the male elasticity of the hazard with respect to vacancies. Suggesting that YT succeeded in convincing employers of the value of a spell on YT. The evidence for the YT effect on females within the dataset suggests that they were less likely to exit the state of unemployment and into work following a period of YT. It seems that the rise in the elasticity of the hazard with respect to vacancies was not enough to offset the rise in reservation wages for scheme participants.

6.3 Final Remarks

Our concluding remarks of Chapter 3 voiced some reservations for the performance of matching methods. We noted that such procedures, when used correctly, could be a powerful tool. Much of the work that followed was concerned with the examination of these methods. It was hoped that by subjecting them to the full range of their specification we might identify a consistency in the YTS treatment effect. Results have shown us the danger of matching as a "magic box" solution and the importance of the careful use and appraisal of the different matching algorithms. Despite the problems highlighted using different matching methods, we believe that their use in a non-experimental setting did allow us to better understand the YTS effect and when used carefully could do the same for a range of econometric bias reduction problems.

Bibliography

Ashenfelter, O. (1978) Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*, 47–57.

Ashenfelter, O. and Card, D. (1985) Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *The Review of Economics and Statistics*, 648–660.

Bradley, S. (1995) The Youth Training Scheme: a critical review of the evaluation literature. *International Journal of Manpower*, **16**(4), 30–56.

Burdett, K., Kiefer, N., Mortenson, D. and Neumann, G. (1984) Earnings, Unemployment and the Allocation of Time over Time. *Review of Economic Studies*, **L1**, 559–578.

Card, D. and Sullivan, D. (1988) Measuring the Effect of Subsidised Training Programs on Movements in and out of Employment. *Econometrica*, 497–530.

Chapman, P. and Tooze, M. (1987) *The Youth Training Scheme in the United Kingdom.* Aldershot: Avebury.

Cochran, Z. and Rubin, Z. (1973) Controlling bias in observational studies. *Sankyha*, (35), 417–446.

Courtney, G. and McAleese, I. (1991) The Youth Cohort Study, report on Cohort III Sweep 3, 18-19 Year Olds in 1989. In *The Employment Department Research and Development Paper*, no. 66, Sheffield.

Crosslin, R. and Stevens, D. (1977) The Asking Wage -Duration of Unemployment Relation Revisited. *Southern Economic Journal*, **43**, 1298–1302.

Deakin, B. M. (1996) *The Youth Labour Market in Britain: The Role of Intervention.* Cambridge: University Press.

Dolton, P., Makepeace, G., Hutton, S. and Audas, R. (1999) *Making the Grade: Education, the Labor Market and Young People.* Work and Opportunity Series, Joseph Rowntree Foundation: Work and Opportunity Series.

Dolton, P., Makepeace, G. and Treble, J. (1992) The Wage Effect of YTS: Evidence from YCS. *Scottish Journal of Political Economy*, **41**(4), 444–453.

Dolton, P., Makepeace, G. and Treble, J. (1993) The Economics of Youth Training in Britain. *Economic Journal*, **103**, 1261–1278.

Dolton, P., Makepeace, G. and Treble, J. (1994) The Youth Training Scheme and the School to Work Transition. *Oxford Economic Papers*, **46**, 629–657.

Dolton, P. J. and O'Neill, D. (1995) Unemployment Duration and the Restart Effect: Some Experimental Evidence. *The Economic Journal*, **106**(435), 387–400.

Gorter, D. and Gorter, C. (1993) The Relation between Unemployment Benefits, the Reservation Wage and Search Duration. *Oxford Bulletin of Economics and Statistics*, **55**(2), 199–214.

Gritz, R. (1993) The Impact of Training on the Frequency and Duration of Employment. *Journal of Econometrics*, 21–51.

GSS (1996) *Youth Cohort Study: Trends in the Activities and Experiences of 16-18 Year Olds: England and Wales 1985-1994.* HMSO.

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1997) Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies*, **64**(4), 605–654.

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1998) Matching as an Econometric Evaluation Estimator. *Review of Economic Studies*, **65**(2), 261–294.

Jones, S. (1988) The Relationship between Unemployment Spells and Reservation Wage as a Test of Search Theory. *The Quarterly Journal of Economics*, **103**, 741–765.

Kiefer, N. and Neumann, G. (1979) An Empirical Job Search Model with a Test of the Constant Reservation Wage Hypothesis. *Journal of Political Economy*, **87**, 89–107.

Lancaster, T. and Chesher, A. (1983) An Econometric Analysis of Reservation Wages. *Econometrica*, **51**(6), 1661–1676.

Lechner, M. (1999) Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany After Unification. *Journal of Business & Economic Statistics*, **17**(1), 74–90.

Lynch, L. (1983) Job Search and Youth Unemployment. *Oxford Economic Papers*, **35**(4), 595–606.

Main, B. and Shelley, M. (1988) School Leavers and the Search for Employment. *Economic Papers*, **40**, 487–504.

Mealli, F., Pudney, S. and Thomas, J. (1996) Training Duration and Post-Training Outcomes: A Duration-Limited Competing Risks Model. *The Economic Journal*, **106**, 422–433.

Mortensen, D. (1986) Job search and Labour Market Analysis. In *Handbook of Labour Economics* (ed. R. Ashenfelter, O. & Layard), vol. II.

Narendranathan, W. and Elias, P. (1993) Influences of Past History on the Incidence of Youth Unemployment: Empirical findings for the UK. *Oxford Bulletin of Economics and Statistics*, 161–185.

Rosenbaum, P. and Rubin, D. (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrica*, **70**, 41–50.

Rubin, D. (1977) Assignment of Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*, **2**, 1–26.

Silverman, B. (1986) *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall.

Taylor, R. (2000) Here Today, Gone Tomorrow.. An Empirical Analysis of Attrition and Recal Bias in Labour Market Data. Ph.D. thesis, University of Newcastle.

Todd, P. (1999) A Practical Guide to Implementing Matching Estimators. Internet, Santiago, Chile. IABB.

Van Den Berg, G. (1994) The Effects of Changes of the Job Offer Arrival rate on the Duration of Unemployment. *Journal of Labor Economics*, **12**(3), 478–498.

Appendix A

A.1 Fortran Code for Extracting Unemployment Durations from Labour Market State Diary Data

```
c     Program to calculate unemployment durations for those
c     who exit to a full time job with laststate indicator
c
c
      double precision wage1,wage2
      integer id,sweep,x(116),i,j,ndur,lstytdur,m,lststate,ytstart
      open(7,file='trans3.txt')
      open(8,file='trans4e.res')
      do 10 i=1,1000000
         x(1)=-99
         read(7,*,end=11) id,wage1,wage2,sweep,(x(j),j=2,115)
         x(116)=-99
         lststate=-9
         m=1
         do 20 j=1,116
            call duration(x,m,ndur,lstytdur,lststate,ytstart)
            write(8,200) id,ndur,lstytdur,lststate,ytstart
            if (ndur.lt.0) goto 10
20       continue
10    continue
200   format(5i10)
11    stop
      end
```

```
      subroutine duration(x,m,ndur,lstytdur,lststate,ytstart)
      integer i,j,k,x(116),ndur,xx,xxx,lstytdur,lststate,m,ytstart
      lstytdur=0
      lststate=0
      ytstart=0
      do 10 i=1,115-m
         xx=x(m+i)
         if (((xx.eq.1).or.(xx.eq.2)).and.(x(m+i+1).eq.5)) goto 20
 10   continue
      ndur=-99
      return
 20   m=m+i
      do 30 j=1,m-1
         xx=x(m-j)
         if ((xx.ne.1).and.(xx.ne.2)) then
            lststate=xx
            if (xx.eq.8) then
               do 25 k=1,m-j-1
                  xxx=x(m-j-k)
                  if (xxx.ne.8) then
                     lstytdur=k
                     ytstart=m-j-k+1
                     goto 40
                  end if
 25            continue
            end if
            goto 40
         end if
 30   continue
 40   ndur=j
      return
      end
```

Appendix B

B.1 Fortran Code for Nearest Neighbour Matching Procedure

```
c      Program to match treatment and control group individuals
c      with similar covariate structures: First estimate a Probit model for YTS
c      participation (I did this in STATA). There is a switch in STATA which cau
c      ses the PROBIT command to also produce the covariance matrix from the mod
c      el. -You'll need this. Including the constant I had 26 variables in my
c      PROBIT
c
       double precision prpensty,covarmtx(26,26),data(26)
      & ,disc,lc(100000),uc(100000),dvar,betas(26)
      & ,datat(26),datac(26),lpt(100000),lpc(100000)
      & ,dist,mindist,eucldist,mineucl,rub2(100000)
       integer dopid(100000),ytsindic(100000),i,j,k,ytstart(100000)
      & ,totcalip,l,dopidt(100000),dopidc(100000),ytstartt(100000)
      & ,ytstartc(100000),nt,nc,exclude(100000),calipmtx(100000)
      & ,excludet(100000),excludec(100000),m,nearest,match(100000)
      & ,o,nearcalp,p,q,label(100000),rub1(100000),r
c      Covariance matrix must be complete (not just the bottom half)
c      Covariance matrix is then stored in covmatrix.txt
       open (7,file='covmatrix.txt')
       do 10 i=1,26
          read(7,*) (covarmtx(i,j),j=1,26)
          print *, 'hello',i
c      Covariance matrix is now read into the 2-D array, covarmtx(i,j)
 10    continue
       close(7)
c      The estimates of the betas from the PROBIT model are stored in
c      thebetas.txt. There are 26 of these. Just list them one after the other
       open (7,file='thebetas.txt')
```

```
      read(7,*) (betas(i),i=1,26)
      close(7)
      print *, (betas(j),j=1,26)
c     The betas vector is now read into the array, betas(j)
c
c     The variables for each person in turn are then read in using the file
c     thedata.txt. This file must list the variables for each person in the
c     same order as they are for the covariance matrix and betas.
c     I used the STATA command:
c     outfile <<dopid, propensity, yts participation indicator, ytstart date
c     (could just leave this set to 1 since the algorithm never used it.),
c     a variable for indicating when a person has been matched (Set exclude=0)
c     and names of variables in order used for PROBIT>> using thedata.txt
      open (7,file='thedata.txt')
      nt=0
      nc=0
      do 20 i=1,3905
         read(7,*) dopid(i),prpensty,ytsindic(i),ytstart(i)
     & ,exclude(i),(data(j),j=1,25)
c     Now read in the data
         data(26)=1d0
         print *, dopid(i),prpensty,ytsindic(i),ytstart(i)
     & ,exclude(i),(data(j),j=1,26)
c     The data is now read into the matrix data(j).
         if (ytsindic(i).eq.1) then
            nt=nt+1
            dopidt(nt)=dopid(i)
            ytstartt(nt)=ytstart(i)
            excludet(nt)=exclude(i)
            do 15 j=1,26
               datat(j)=data(j)
 15         continue
            disc=0d0
            dvar=0d0
            do 30 j=1,26
               disc=disc+betas(j)*datat(j)
               do 40 k=1,26
                  dvar=dvar+covarmtx(j,k)*datat(j)*datat(k)
 40            continue
 30         continue
            lpt(nt)=disc
            print *, disc,dvar
            lc(nt)=disc-1.645*dsqrt(dvar)
            uc(nt)=disc+1.645*dsqrt(dvar)
         else
            nc=nc+1
            dopidc(nc)=dopid(i)
            ytstartc(nc)=ytstart(i)
```

```
            excludec(nc)=exclude(i)
            do 16 j=1,26
               datac(j)=data(j)
16          continue
            disc=0d0
            do 31 j=1,26
               disc=disc+betas(j)*datac(j)
31          continue
            lpc(nc)=disc
            print *, disc
         end if
20    continue
      print *, nt,nc
      close(7)
      do 47 p=1,nt
         label(p)=p
47    continue
      ifail=0
      call g05ehf(label,nt,ifail)
      do 49 q=1,nt
         rub1(q)=dopidt(label(q))
         rub2(q)=lpt(label(q))
49    continue
         dopidt=rub1
         lpt=rub2
      do 52 p=1,nc
         label(p)=p
52    continue
      ifail=0
      call g05ehf(label,nc,ifail)
      do 54 r=1,nc
         rub1(r)=dopidc(label(r))
         rub2(r)=lpc(label(r))
54    continue
         dopidc=rub1
         lpc=rub2
      do 50 i=1,nt
         write(6,987) i
987      format('Person ',i3)
         totcalip=0
         do 60 l=1,nc
            if (excludec(l).eq.1) goto 60
            if ((lpc(l).ge.lc(i)).and.(lpc(l).le.uc(i))) then
               totcalip=totcalip+1
               calipmtx(totcalip)=l
            end if
60       continue
         print *, totcalip
```

```
        if (totcalip.eq.0) then
           mindist=1d10
           do 80 m=1,nc
              dist=(lpc(m)-lpt(i))*(lpc(m)-lpt(i))
              if ((dist.lt.mindist).and.(excludec(m).eq.0)) then
                 mindist=dist
                 nearest=m
              end if
 80        continue
           match(i)=nearest
           excludec(nearest)=1
        end if
        if (totcalip.eq.1) then
           nearest=calipmtx(1)
           match(i)=nearest
           excludec(nearest)=1
        end if
        if (totcalip.ge.2) then
           mineucl=1d10
           do 90 o=1,totcalip
              eucldist=(lpc(calipmtx(o))-lpt(i))
     & *(lpc(calipmtx(o))-lpt(i))
              print *, eucldist
              if (eucldist.lt.mineucl) then
                 mineucl=eucldist
                 nearcalp=o
              end if
 90        continue
        print *, nearcalp
        nearest=calipmtx(nearcalp)
        match(i)=nearest
        excludec(nearest)=1
        end if
      print *, nearest

 50   continue
c     Matches are now output
      open (7,file='trainmatch.res')
      open (8,file='contlmatch.res')
      open (9,file='bothmatch.res')
c     trainmatch.res contains the matched YTS participants
      write(7,*) (dopidt(j), j=1,nt)
c     contrlmatch.res contains the matched YTS non-participants
      write(8,*) (dopidc(match(j)),j=1,nt)
c     bothmatch.res contains the matched YTS participants followed by the
c     matched non-participants
c
c     The first dopid number in trainmatch.res has been matched to the first
```

```
c     dopid number in contrlmatch.res
      write(9,*) (dopidt(j), j=1,nt)
      write(9,*) (dopidc(match(j)),j=1,nt)
      close(7)
      close(8)
      close(9)
 200  format(i10,i4,3d25.15)
 11   stop
      end
```

B.2 Fortran Code for Local Linear Regression Bootstrap Iterative Procedure

```
c     Program to match treatment and control group individuals
c     with similar covariate structures (for men) using local
c     linear regression with t-test for treatment/control
c     differences. H0: m1.eq.m2 vs H1: m1.neq.m2
c
c     Takes random samples of 1000 individuals from all control
c     persons and YTSI treatment individuals. Repeats this process
c     1000 times and produces the Bootstrapped estimates of job search
c     elasticites from these samples.
c
c     The main program reads in the data and then generates the Boot-
c     strapped random samples. These are then passed to the subroutine
c     genh.
c
      double precision prpensty(5000),ebyb(5000),ebyl(5000)
     & ,bprpnsty(1000),bebyb(1000),bebyl(1000),bphbyb(1000),behbyb(1000)
     & ,behbyl(1000),bnewrw(1000),bnewew(1000),booth
     & ,phbyb(5000),ehbyb(5000),ehbyl(5000),newrw(5000),newew(5000)
     & ,totnewew,tsynewew,tnewewt,relnewew,synnewew,newewtpl
     & ,totnewrw,tsynewrw,tnewrwt,relnewrw,synnewrw,newrwtpl
     & ,totebyb,tsyebyb,tebybt,relebyb,synebyb,ebybtpl
     & ,totebyl,tsyebyl,tebylt,relebyl,synebyl,ebyltpl
     & ,totphbyb,tsyphbyb,tphbybt,relphbyb,synphbyb,phbybtpl
     & ,totehbyb,tsyehbyb,tehbybt,relehbyb,synehbyb,ehbybtpl
     & ,totehbyl,tsyehbyl,tehbylt,relehbyl,synehbyl,ehbyltpl
     & ,tsrnewew,tssnewew,sdrnewew,sdsnewew
     & ,tsrnewrw,tssnewrw,sdrnewrw,sdsnewrw
     & ,tsrebyb,tssebyb,sdrebyb,sdsebyb
     & ,tsrebyl,tssebyl,sdrebyl,sdsebyl
     & ,tsrphbyb,tssphbyb,sdrphbyb,sdsphbyb
     & ,tsrehbyb,tssehbyb,sdrehbyb,sdsehbyb
     & ,tsrehbyl,tssehbyl,sdrehbyl,sdsehbyl
      integer p,newid(4500),ytsindic(4500),nall,sampn,samp(1000)
     & ,bnewid(1000),bytsindc(1000),bootnumb,o
      open (7,file='kernelmytsi.txt')
      open (8,file='newrwewytsi.txt')
      bootnumb=1000
      booth=0.9
      sampn=1000
      nall=3384
      do 20 p=1,nall
         read(7,*) newid(p),prpensty(p),ytsindic(p),ebyb(p),ebyl(p)
```

```
     & ,phbyb(p),ehbyb(p),ehbyl(p)
         read(8,*) newrw(p),newew(p)
         print *, newid(p)
 20   continue
      close(7)
      close(8)
      open (3,file='booti.res')
      open (8,file='bootlocalytsim.res')
      open (7,file='bootlocalytsim2.res')
      call g05CCF
      do 24 o=1,bootnumb
      ifail=0
      call G05EJF(newid, nall, samp, sampn, ifail)
         do 23 p=1,sampn
            bnewid(p)=samp(p)
            bprpnsty(p)=prpensty(samp(p))
            bytsindc(p)=ytsindic(samp(p))
            bnewrw(p)=newrw(samp(p))
            bnewew(p)=newew(samp(p))
            bebyb(p)=ebyb(samp(p))
            bebyl(p)=ebyl(samp(p))
            bphbyb(p)=phbyb(samp(p))
            behbyb(p)=ehbyb(samp(p))
            behbyl(p)=ehbyl(samp(p))
 23      continue
      call genh(bnewid,bprpnsty,bytsindc,bebyb,bebyl
     & ,bphbyb,behbyb,behbyl,bnewrw,bnewew,sampn,booth
     & ,relnewew,synnewew,newewtpl,relnewrw,synnewrw,newrwtpl
     & ,relebyb,synebyb,ebybtpl,relebyl,synebyl,ebyltpl
     & ,relphbyb,synphbyb,phbybtpl,relehbyb,synehbyb,ehbybtpl
     & ,relehbyl,synehbyl,ehbyltpl
     & ,sdrnewew,sdsnewew,sdrnewrw,sdsnewrw
     & ,sdrebyb,sdsebyb,sdrebyl,sdsebyl
     & ,sdrphbyb,sdsphbyb,sdrehbyb,sdsehbyb
     & ,sdrehbyl,sdsehbyl)
      print *, booth
      totnewew=totnewew+relnewew
      tsynewew=tsynewew+synnewew
      tnewewt=tnewewt+newewtpl
      tsrnewew=tsrnewew+sdrnewew
      tssnewew=tssnewew+sdsnewew
c     *************************************
      totnewrw=totnewrw+relnewrw
      tsynewrw=tsynewrw+synnewrw
      tnewrwt=tnewrwt+newrwtpl
      tsrnewrw=tsrnewrw+sdrnewrw
      tssnewrw=tssnewrw+sdsnewrw
c     *************************************
```

```
      totebyb=totebyb+relebyb
      tsyebyb=tsyebyb+synebyb
      tebybt=tebybt+ebybtpl
      tsrebyb=tsrebyb+sdrebyb
      tssebyb=tssebyb+sdsebyb
c     ***********************************
      totebyl=totebyl+relebyl
      tsyebyl=tsyebyl+synebyl
      tebylt=tebylt+ebyltpl
      tsrebyl=tsrebyl+sdrebyl
      tssebyl=tssebyl+sdsebyl
c     ***********************************
      totphbyb=totphbyb+relphbyb
      tsyphbyb=tsyphbyb+synphbyb
      tphbybt=tphbybt+phbybtpl
      tsrphbyb=tsrphbyb+sdrphbyb
      tssphbyb=tssphbyb+sdsphbyb
c     ***********************************
      totehbyb=totehbyb+relehbyb
      tsyehbyb=tsyehbyb+synehbyb
      tehbybt=tehbybt+ehbybtpl
      tsrehbyb=tsrehbyb+sdrehbyb
      tssehbyb=tssehbyb+sdsehbyb
c     ***********************************
      totehbyl=totehbyl+relehbyl
      tsyehbyl=tsyehbyl+synehbyl
      tehbylt=tehbylt+ehbyltpl
      tsrehbyl=tsrehbyl+sdrehbyl
      tssehbyl=tssehbyl+sdsehbyl
c     ***********************************
      write (8,*) '***********'
      write (7,*) '***********'
 24   continue
      totnewew=totnewew/1000
      tsynewew=tsynewew/1000
      tnewewt=tnewewt/1000
      tsrnewew=tsrnewew/1000
      tssnewew=tssnewew/1000
c     ***********************************
      totnewrw=totnewrw/1000
      tsynewrw=tsynewrw/1000
      tnewrwt=tnewrwt/1000
      tsrnewrw=tsrnewrw/1000
      tssnewrw=tssnewrw/1000
c     ***********************************
      totebyb=totebyb/1000
      tsyebyb=tsyebyb/1000
      tebybt=tebybt/1000
```

```
      tsrebyb=tsrebyb/1000
      tssebyb=tssebyb/1000
c     *************************************
      totebyl=totebyl/1000
      tsyebyl=tsyebyl/1000
      tebylt=tebylt/1000
      tsrebyl=tsrebyl/1000
      tssebyl=tssebyl/1000
c     *************************************
      totphbyb=totphbyb/1000
      tsyphbyb=tsyphbyb/1000
      tphbybt=tphbybt/1000
      tsrphbyb=tsrphbyb/1000
      tssphbyb=tssphbyb/1000
c     *************************************
      totehbyb=totehbyb/1000
      tsyehbyb=tsyehbyb/1000
      tehbybt=tehbybt/1000
      tsrehbyb=tsrehbyb/1000
      tssehbyb=tssehbyb/1000
c     *************************************
      totehbyl=totehbyl/1000
      tsyehbyl=tsyehbyl/1000
      tehbylt=tehbylt/1000
      tsrehbyl=tsrehbyl/1000
      tssehbyl=tssehbyl/1000
c     *************************************
      write (3,*) totnewew,tsynewew,tnewewt
      write (3,*) tsrnewew,tssnewew
      write (3,*) totnewrw,tsynewrw,tnewrwt
      write (3,*) tsrnewrw,tssnewrw
      write (3,*) totebyb,tsyebyb,tebybt
      write (3,*) tsrebyb,tssebyb
      write (3,*) totebyl,tsyebyl,tebylt
      write (3,*) tsrebyl,tssebyl
      write (3,*) totphbyb,tsyphbyb,tphbybt
      write (3,*) tsrphbyb,tssphbyb
      write (3,*) totehbyb,tsyehbyb,tehbybt
      write (3,*) tsrehbyb,tssehbyb
      write (3,*) totehbyl,tsyehbyl,tehbylt
      write (3,*) tsrehbyl,tssehbyl
      close (3)
      close(8)
      close (7)
 11   stop
 250  format(1d25.10)
      end
```

```
c     The subroutine genh calls varoius other functions which
c     generate the values of h (the bandwidth parameter for
c     the kernel) and the matches themselves.

       subroutine genh(newid,prpensty,ytsindic,ebyb,ebyl
     & ,phbyb,ehbyb,ehbyl,newrw,newew,nall,booth
     & ,relnewew,synnewew,newewtpl,relnewrw,synnewrw,newrwtpl
     & ,relebyb,synebyb,ebybtpl,relebyl,synebyl,ebyltpl
     & ,relphbyb,synphbyb,phbybtpl,relehbyb,synehbyb,ehbybtpl
     & ,relehbyl,synehbyl,ehbyltpl
     & ,sdrnewew,sdsnewew,sdrnewrw,sdsnewrw
     & ,sdrebyb,sdsebyb,sdrebyl,sdsebyl
     & ,sdrphbyb,sdsphbyb,sdrehbyb,sdsehbyb
     & ,sdrehbyl,sdsehbyl)
       double precision prpnstyc(5000),prpensty(5000)
     & ,ebyb(5000),ebyl(5000),phbyb(5000),ehbyb(5000),ehbyl(5000)
     & ,newrw(5000),newew(5000)
     & ,ebybt(2000),ebylt(2000),prpnstyt(2000),s
     & ,phbybt(2000),ehbybt(2000),ehbylt(2000)
     & ,ebybc(4000),ebylc(4000),sumkernl,kernel,weighsum
     & ,phbybc(4000),ehbybc(4000),ehbylc(4000)
     & ,weight(500,3500),synebybc(4500)
     & ,synebylc(4500),syphbybc(4500),syehbybc(4500)
     & ,syehbylc(4500),smkpdfsq,smkpdiff,pdiff
     & ,kpdiffsq,kpdiff
     & ,newrwt(4500),newewt(4500),newrwc(4500),prpdif(500,3500)
     & ,newewc(4500),synewrwc(4500),synewewc(4500),sqprpdif(4500)
     & ,diffebyb(4500),diffebyl(4500),difphbyb(4500),hidiff,oldhi
     & ,difehbyb(4500),difehbyl(4500),difnewrw(4500),h,hsize
     & ,difnewew(4500),mdifebyb,diffsq(4500),sd,t,ntplus,sumdiff
     & ,sdifebyb,totdiff,sumdifsq,sdp,tplus,meandiff,meanact,meansynth
     & ,ssknl,ssknl2,sumnum,spls,sumkern,sumkern2,numerat(4500),booth
     & ,sumprpi,meanprpi,sumsqpdf,stdevdfi,hi,synprpyc(4500),samp(1000)
     & ,relnewew,synnewew,newewtpl,relnewrw,synnewrw,newrwtpl
     & ,relebyb,synebyb,ebybtpl,relebyl,synebyl,ebyltpl
     & ,relphbyb,synphbyb,phbybtpl,relehbyb,synehbyb,ehbybtpl
     & ,relehbyl,synehbyl,ehbyltpl,stderact,stdersyn
     & ,sdrnewew,sdsnewew,sdrnewrw,sdsnewrw
     & ,sdrebyb,sdsebyb,sdrebyl,sdsebyl
     & ,sdrphbyb,sdsphbyb,sdrehbyb,sdsehbyb
     & ,sdrehbyl,sdsehbyl
       integer i,p,j,k,newid(4500),ytsindic(4500),nall
     & ,newidt(4500),newidc(4500),iter
     & ,nt,nc
       h=0.001d0
```

```
      nt=0
      nc=0
      do 21 p=1,nall
         if (ytsindic(p).eq.1) then
            nt=nt+1
            newidt(nt)=newid(p)
            prpnstyt(nt)=prpensty(p)
            newrwt(nt)=newrw(p)
            newewt(nt)=newew(p)
            ebybt(nt)=ebyb(p)
            ebylt(nt)=ebyl(p)
            phbybt(nt)=phbyb(p)
            ehbybt(nt)=ehbyb(p)
            ehbylt(nt)=ehbyl(p)
         else
            nc=nc+1
            newidc(nc)=newid(p)
            prpnstyc(nc)=prpensty(p)
            newrwc(nc)=newrw(p)
            newewc(nc)=newew(p)
            ebybc(nc)=ebyb(p)
            ebylc(nc)=ebyl(p)
            phbybc(nc)=phbyb(p)
            ehbybc(nc)=ehbyb(p)
            ehbylc(nc)=ehbyl(p)
         end if
21    continue
      print *, nt,nc,h
      do 30 i=1,nt
         do 32 j=1,nc
            prpdif(i,j)=prpnstyt(i)-prpnstyc(j)
32       continue
30    continue
         sumprpi=0d0
         do 37 j=1,nc
            sumprpi=sumprpi+prpnstyc(j)
37       continue
         meanprpi=sumprpi/nc
         do 40 j=1,nc
            sqprpdif(j)=(prpnstyc(j)-meanprpi)*
     & (prpnstyc(j)-meanprpi)
40       continue
         sumsqpdf=0d0
         do 42 j=1,nc
            sumsqpdf=sumsqpdf+sqprpdif(j)
42       continue
         stdevdfi=sqrt(sumsqpdf/(nc-1))
         hi=1.06*stdevdfi*(nc**(-0.2))
```

```
      hi=booth
      print *, 'hi:',hi
   iter=1
   print *, 'iteration:', iter
   write (7,300) iter
   write (7,250) hi
   call mksynth(newrwc,newewc,ebybc,ebylc,phbybc
  & ,ehbybc,ehbylc,synewrwc,synewewc,synebybc,synebylc,syphbybc
  & ,syehbybc,syehbylc,synprpyc,prpnstyc,prpnstyt,hi,nt,nc)
   call passhi(prpnstyt,synprpyc,nt,nc,hi)
   write (7,250) hi
   iter=2
   hsize=0.005
   hidiff=1
   do while ((hidiff.gt.hsize).and.(iter.lt.101))
      print *, 'iteration:', iter
      write (7,300) iter
      write (7,250) hi
      oldhi=hi
      call mksynthx(newrwc,newewc,ebybc,ebylc,phbybc
  & ,ehbybc,ehbylc,synewrwc,synewewc,synebybc,synebylc,syphbybc
  & ,syehbybc,syehbylc,synprpyc,prpnstyc,prpnstyt,hi,nt,nc)
     call passhi(prpnstyt,synprpyc,nt,nc,hi)
     hidiff=sqrt((oldhi-hi)*(oldhi-hi))
     iter=iter+1
     end do
   totdiff=0d0
   meanact=0d0
   meansynth=0d0
   meandiff=0d0
   sumdifsq=0d0
   sdp=0d0
   tplus=0d0
   booth=hi
   write (3,*) booth
   type *; '      ***************************************************
  &*********************'
   type *, '      |   Local Linear Regression: hi =',hi,'
  &|'
   type *, '      ***************************************************
  &*********************'
   type *, '      |     treatment              |        Synthetic         |
  &     t-stat        |'
   type *, '      ***************************************************
  &*********************'
   type *, '      |                            |                          |
  &                   |'
   call ttesting(newewt,synewewc,nt,totdiff,meandiff,sumdifsq
```

```
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relnewew=meanact
 synnewew=meansynth
 newewtpl=tplus
 sdrnewew=stderact
 sdsnewew=stdersyn
 call ttesting(newrwt,synewrwc,nt,totdiff,meandiff,sumdifsq
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relnewrw=meanact
 synnewrw=meansynth
 newrwtpl=tplus
 sdrnewrw=stderact
 sdsnewrw=stdersyn
 call ttesting(ebybt,synebybc,nt,totdiff,meandiff,sumdifsq
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relebyb=meanact
 synebyb=meansynth
 ebybtpl=tplus
 sdrebyb=stderact
 sdsebyb=stdersyn
 call ttesting(ebylt,synebylc,nt,totdiff,meandiff,sumdifsq
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relebyl=meanact
 synebyl=meansynth
 ebyltpl=tplus
 sdrebyl=stderact
 sdsebyl=stdersyn
 call ttesting(phbybt,syphbybc,nt,totdiff,meandiff,sumdifsq
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relphbyb=meanact
 synphbyb=meansynth
 phbybtpl=tplus
 sdrphbyb=stderact
 sdsphbyb=stdersyn
 call ttesting(ehbybt,syehbybc,nt,totdiff,meandiff,sumdifsq
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relehbyb=meanact
 synehbyb=meansynth
 ehbybtpl=tplus
 sdrehbyb=stderact
 sdsehbyb=stdersyn
 call ttesting(ehbylt,syehbylc,nt,totdiff,meandiff,sumdifsq
& ,sdp,tplus,meanact,meansynth,stderact,stdersyn)
 relehbyl=meanact
 synehbyl=meansynth
 ehbyltpl=tplus
 sdrehbyl=stderact
 sdsehbyl=stdersyn
```

```
      type *, '      ****************************************************
     &*******************'
      return
 200  format(i10,i4,3d25.15)
 250  format(1d25.10)
 300  format(i3)
      end



c     The subroutines mksynth and mksynthx generate the synthetic wages
c     and elasticites.

      subroutine mksynth(newrwc,newewc,ebybc,ebylc,phbybc
     & ,ehbybc,ehbylc,synewrwc,synewewc,synebybc,synebylc,syphbybc
     & ,syehbybc,syehbylc,synprpyc,prpnstyc,prpnstyt,hi,nt,nc)
      double precision diff(4500),actual(4500),synth(4500)
     & ,ssknl,ssknl2,sumnum,sumkern,s,kernel,hi,numerat(4500),weighsum
     & ,synewrwc(4500),synewewc(4500),synebybc(4500),synebylc(4500)
     & ,syphbybc(4500),syehbybc(4500),syehbylc(4500),newrwc(4500)
     & ,newewc(4500),ebybc(4500),ebylc(4500),phbybc(4500)
     & ,ehbybc(4500),ehbylc(4500),prpnstyc(4500),prpnstyt(4500)
     & ,weight(400,3500),synprpyc(4500),sumkern2,spls
      integer i,j,k,nt,nc
      do 150 i=1,nt
         ssknl=0d0
         ssknl2=0d0
         sumnum=0d0
         do 160 j=1,nc
            sumkern=0d0
            sumkern2=0d0
            do 154 k=1,nc
               s=(prpnstyc(i)-prpnstyt(k))/hi
               spls=prpnstyc(k)-prpnstyt(i)
               kernel=15*(s*s-1)*(s*s-1)/16
               sumkern=sumkern+kernel*spls
               sumkern2=sumkern2+kernel*spls*spls
 154         continue
            s=(prpnstyc(i)-prpnstyt(j))/hi
            spls=prpnstyc(j)-prpnstyt(i)
            kernel=15*(s*s-1)*(s*s-1)/16
            numerat(j)=sumkern2*kernel-sumkern*kernel*spls
            sumnum=sumnum+numerat(j)
 160     continue
         weighsum=0d0
         synprpyc(i)=0d0
         synewrwc(i)=0d0
```

```
         synewewc(i)=0d0
         synebybc(i)=0d0
         synebylc(i)=0d0
         syphbybc(i)=0d0
         syehbybc(i)=0d0
         syehbylc(i)=0d0
         do 155 j=1,nc
            weight(i,j)=numerat(j)/sumnum
            weighsum=weighsum+weight(i,j)
            synprpyc(i)=synprpyc(i)+weight(i,j)*prpnstyc(j)
            synewrwc(i)=synewrwc(i)+weight(i,j)*newrwc(j)
            synewewc(i)=synewewc(i)+weight(i,j)*newewc(j)
            synebybc(i)=synebybc(i)+weight(i,j)*ebybc(j)
            synebylc(i)=synebylc(i)+weight(i,j)*ebylc(j)
            syphbybc(i)=syphbybc(i)+weight(i,j)*phbybc(j)
            syehbybc(i)=syehbybc(i)+weight(i,j)*ehbybc(j)
            syehbylc(i)=syehbylc(i)+weight(i,j)*ehbylc(j)
 155      continue
150   continue
      return
 300  format(3d25.10)
      end


      subroutine mksynthx(newrwc,newewc,ebybc,ebylc,phbybc
     & ,ehbybc,ehbylc,synewrwc,synewewc,synebybc,synebylc,syphbybc
     & ,syehbybc,syehbylc,synprpyc,prpnstyc,prpnstyt,hi,nt,nc)
      double precision diff(4500),actual(4500),synth(4500)
     & ,ssknl,ssknl2,sumnum,sumkern,s,kernel,hi,numerat(4500),weighsum
     & ,synewrwc(4500),synewewc(4500),synebybc(4500),synebylc(4500)
     & ,syphbybc(4500),syehbybc(4500),syehbylc(4500),newrwc(4500)
     & ,newewc(4500),ebybc(4500),ebylc(4500),phbybc(4500)
     & ,ehbybc(4500),ehbylc(4500),prpnstyc(4500),prpnstyt(4500)
     & ,weight(400,3500),synprpyc(4500),sumkern2,spls
      integer i,j,k,nt,nc
      do 150 i=1,nt
         ssknl=0d0
         ssknl2=0d0
         sumnum=0d0
         do 160 j=1,nc
            sumkern=0d0
            sumkern2=0d0
            do 154 k=1,nc
               s=(synprpyc(i)-prpnstyt(k))/hi
               spls=synprpyc(k)-prpnstyt(i)
               kernel=15*(s*s-1)*(s*s-1)/16
               sumkern=sumkern+kernel*spls
               sumkern2=sumkern2+kernel*spls*spls
```

```
154          continue
           s=(synprpyc(i)-prpnstyt(j))/hi
           spls=synprpyc(j)-prpnstyt(i)
           kernel=15*(s*s-1)*(s*s-1)/16
           numerat(j)=sumkern2*kernel-sumkern*kernel*spls
           sumnum=sumnum+numerat(j)
160       continue
         weighsum=0d0
         synprpyc(i)=0d0
         synewrwc(i)=0d0
         synewewc(i)=0d0
         synebybc(i)=0d0
         synebylc(i)=0d0
         syphbybc(i)=0d0
         syehbybc(i)=0d0
         syehbylc(i)=0d0
         do 155 j=1,nc
            weight(i,j)=numerat(j)/sumnum
            weighsum=weighsum+weight(i,j)
            synprpyc(i)=synprpyc(i)+weight(i,j)*prpnstyc(j)
            synewrwc(i)=synewrwc(i)+weight(i,j)*newrwc(j)
            synewewc(i)=synewewc(i)+weight(i,j)*newewc(j)
            synebybc(i)=synebybc(i)+weight(i,j)*ebybc(j)
            synebylc(i)=synebylc(i)+weight(i,j)*ebylc(j)
            syphbybc(i)=syphbybc(i)+weight(i,j)*phbybc(j)
            syehbybc(i)=syehbybc(i)+weight(i,j)*ehbybc(j)
            syehbylc(i)=syehbylc(i)+weight(i,j)*ehbylc(j)
155       continue
150    continue
       return
300    format(3d25.10)
       end




c      The subroutine passhi creates the new value for h (the bandwidth
c      parameter of the kernel) from the arguements passed to it.

       subroutine passhi(prpnstyt,synprpyc,nt,nc,hi)
       double precision prpnstyt(4500),synprpyc(4500),hi
     & ,fprpdif(1400,3500),fsumprpi,fmeanpri,fsqprdif(4500),fsumsqdf
     & ,fstdevdf
       integer i,j,nc,nt
       nc=nt
       do 400 i=1,nt
         do 410 j=1,nc
            fprpdif(i,j)=prpnstyt(i)-synprpyc(j)
```

```
 410     continue
 400   continue
         fsumprpi=0d0
         do 420 j=1,nc
            fsumprpi=fsumprpi+synprpyc(j)
 420     continue
         fmeanpri=fsumprpi/nc
         do 430 j=1,nc
            fsqprdif(j)=(synprpyc(j)-fmeanpri)*
     & (synprpyc(j)-fmeanpri)
 430      continue
         fsumsqdf=0d0
         do 440 j=1,nc
            fsumsqdf=fsumsqdf+fsqprdif(j)
 440     continue
         fstdevdf=sqrt(fsumsqdf/(nc-1))
         hi=1.06d0*fstdevdf*(nc**(-0.2d0))
       print *, 'hi:',hi
       return
       end



c      The subroutine ttesting performs the t-tests for significant
c      differences between the actual and synthetic wages and
c      elasticites.

       subroutine ttesting(actual,synth,nt,totdiff,meandiff
     & ,sumdifsq,sdp,tplus,meanact,meansynth,stderact,stdersyn)
       double precision diff(5000),actual(5000),synth(5000)
     & ,totdiff,meandiff,diffsq(5000),sumdifsq,sdp,ntplus
     & ,tplus,sumact,meanact,sumsynth,meansynth,difactsq,sumdfasq
     & ,difsynsq,sumdfssq,varact,varsyn,stderact,stdersyn
       integer i,nt
       sumact=0d0
       sumsynth=0d0
       meanact=0d0
       meansynth=0d0
       difactsq=0d0
       sumdfasq=0d0
       difsynsq=0d0
       sumdfssq=0d0
       varact=0d0
       varsyn=0d0
       do 10 i=1,nt
          sumact=sumact+actual(i)
          sumsynth=sumsynth+synth(i)
```

```
10   continue
     meanact=sumact/nt
     meansynth=sumsynth/nt
     do 20 i=1,nt
        difactsq=(actual(i)-meanact)*(actual(i)-meanact)
        sumdfasq=sumdfasq+difactsq
        difsynsq=(synth(i)-meansynth)*(synth(i)-meansynth)
        sumdfssq=sumdfssq+difsynsq
20   continue
     varact=sumdfasq/(nt-1)
     varsyn=sumdfssq/(nt-1)
     ntplus=nt
     stderact=sqrt(varact)/sqrt(ntplus)
     stdersyn=sqrt(varsyn)/sqrt(ntplus)
     sdp=sqrt((varact/nt)+(varsyn/nt))
     tplus=(meanact-meansynth)/sdp
     write (8,250) meanact,meansynth,tplus
     write (8,250) stderact,stdersyn
     type *, '     |                    ', ' |                 '
    & , '       | ', '                  |'
     return
250  format(3d25.10)
     end
```

Appendix C

C.1 Matching Algorithm: Based on Lechner (1999)

1. Estimation of Propensity Score via a Probit Model.
2. Compute $v\hat{\beta}$ and its conditional variance $\text{var}(V\hat{\beta}|V = v)$ for each observation.
3. Split people into a treatment vector (people with YTS experience) and a control vector.
4. Randomise the order of observations in both the T and C vectors.
5. Draw the first person in the randomised T vector.
6. Calculate caliper of propensity score for observation in terms of its predicted probit index, $v_{n_t}\hat{\beta}$, and its conditional variance $\text{var}(V\hat{\beta}|V = v_{n_t})$.
7. Find observations from C vector (denote these as j) which lie within the caliper. I.e:

   $v_j\hat{\beta} \in [v_{n_t}\hat{\beta} \pm C\sqrt{var(v_{n_t}\hat{\beta})}]$ Where $C = 1.65$, so that each treatment persons caliper forms a 90% confidence interval for their propensity score, $v_{n_t}\hat{\beta}$
8. If $j \geq 1$ then calculate the above only for those j people within the caliper.
9. Remove the matched individual, j, from C.
10. Repeat until all individuals from treatment vector are matched with a control.

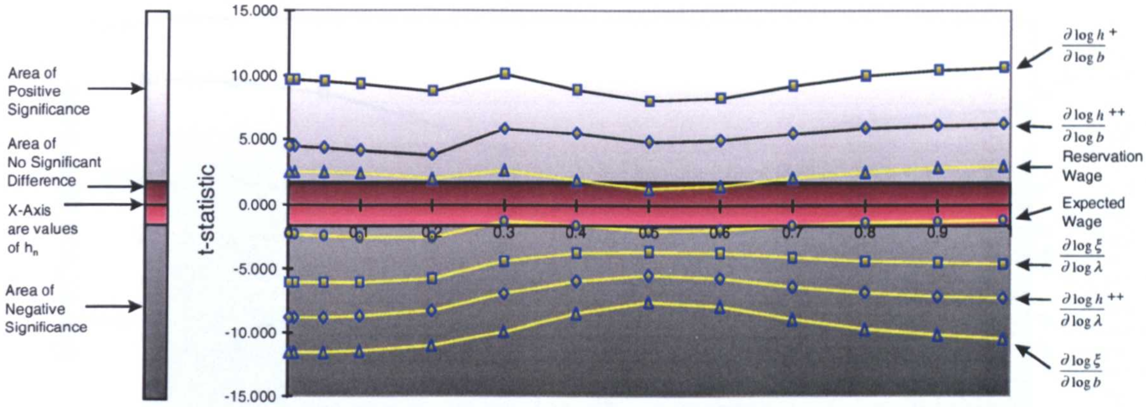

Figure C.1: Kernel Regression Matching Performance as Measured by t-statistic Fluctuations for Differing h_n (Males)

C.2 Interpretation of Charts within Chapter 4

Chapter 4 contains a number of plots which are explained here to avoid repetition.

C.2.1 Charts Representing t-statistic Fluctuations

Figure C.1 compares kernel regression matched males with and without any form of YTS experience. In it we present a plot of these t-statistics as h_n was varied between 0 and 1. The x-axis represents the variation in h_n from 0 to 1. The y-axis depicts the t-statistic values for the differences between treatment and synthetic controls of wages/elasticities produced using the values of h_n. Therefore, the yellow plot lines represent the movement of these t-statistics for each wage/elasticity pair as h_n increases. The wages/elasticities to which the yellow lines belong are indicated to the right of the figure. The red area represents the area of no significant difference between a synthetic control person and their treatment counterpart. Hence yellow plots, which reside beyond this region, indicate the presence of a treatment effect.

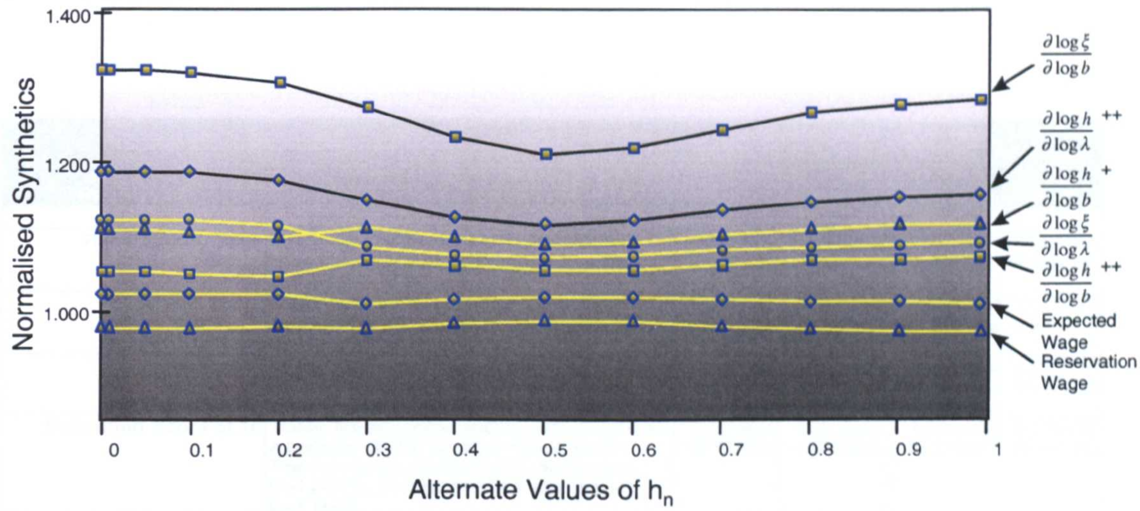

Figure C.2: Kernel Regression Matching Performance as Measured by Fluctuations in the Percentage Differences Between Treatment and Synthetic Elasticities for Various h_n (Males)

C.2.2 Charts Representing Percentage Difference Fluctuations

Figure C.2 presents the way in which the magnitudes of the differences between treatment and synthetic wages/elasticities vary over h_n for the same male sub-sample as Figure C.1 using kernel regression matching. Here we have standardised the treatment wages/elasticities to 1. Hence, a synthetic control wage/elasticity with a value greater than 1 indicates that the wage/elasticity increased after treatment.

C.3 Data

Regressors	Description	Un-Mtchd Males Mean	Un-Mtchd Females Mean	Matched Males Mean	Matched Females Mean
Reservation Wage	Responses to the question: '*What is the lowest weekly take home pay you would consider for a full time job?*'	96.16908 (0.49596)	88.6961 (0.42353)	98.75319 (0.65870)	88.61614 (0.52228)
Expected Wage	Responses to the question: '*How much weekly take home pay do you expect to earn in your next job?*'	126.1087 (0.66737)	115.8429 (0.54831)	126.8414 (0.90866)	114.2719 (0.69372)
Ethnic Origin	White: Respondent is white, Other: Respondent belongs to another ethnic group.	0.081128 (0.27306)	0.073842 (0.26154)	0.066863 (0.24985)	0.069749 (0.25477)
Education Score at 16	These are education scores calculated from a persons GCSE (or equivalent) results. The scores are computed as A=5 points, B=4 points, C=3 points, D=2 points and E=1 point.	8.084804 (9.72350)	8.861285 (9.66383)	4.265487 (5.51636)	5.339342 (6.40732)
Maths GCSE	1 Respondent held a GCSE (or equivalent) in Maths at grade C or above, 0 Otherwise.	0.215931 (0.41152)	0.188986 (0.39154)	0.086529 (0.28121)	0.107367 (0.30964)
English GCSE	1 Respondent held a GCSE (or equivalent) in English at grade C or above, 0 Otherwise.	0.256618 (0.43682)	0.325824 (0.46873)	0.139626 (0.34668)	0.206505 (0.40488)
No. of Siblings	The total number of siblings (not a dummy).	1.807843 (1.55990)	1.858156 (1.62853)	2.073255 (1.69402)	2.023119 (1.72236)
Career Service Interview	1 Respondent attended a career service interview, 0 Otherwise.	0.141667 (0.34875)	0.111181 (0.31439)	0.174041 (0.37924)	0.142633 (0.34977
Live With Parents	Respondent lived with parent(s).	0.977696 (0.14769)	0.952232 (0.21330)	0.97296 (0.16224)	0.939655 (0.23817)
YTSI	Respondent had a spell on the first incarnation of YTS.	0.083333 (0.27642)	0.075511 (0.26424)	0.167158 (0.37321)	0.14185 (0.34896)
YTSII	Respondent had a spell on the second incarnation of YTS.	0.116667 (0.32106)	0.131206 (0.33766)	0.234022 (0.42349)	0.246473 (0.43104)
YT	Respondent had a spell on YT	0.053922 (0.22589)	0.06237 (0.24185)	0.09882 (0.29849)	0.111677 (0.31503)
Pred. Training allowance	An estimate of the respondents training allowance.	- (-)	- (-)	3.998719 (0.16567)	3.952304 (0.14224)
Regions	North, Yorkshire & Humberside, East Midlands, East Anglia, Greater London, South East, South West, West Midlands, North West and Wales. North of England is used as a reference group.	- (-)	- (-)	- (-)	- (-)
Cohorts	Dummies for Cohorts 1 to 6. Cohort 1 is used as a reference group.	- (-)	- (-)	- (-)	- (-)
	Exogenous Variables				
Youth Unemployment	These variables were collected by region and come from Regional Trends. Wages are adult figures and are dis-aggregated by gender, so females in the database are assigned mean female wages for their home region in the relevant time period. Regional training place figures are normalised by using the population of 16-19 year olds in the region for each year. Population figures provided by the Office for National Statistics (ONS).	0.064805 (0.03561)	0.064578 (0.03562)	0.076662 (0.04027)	0.075117 (0.03982)
Regional YT Places		0.083314 (0.02429)	0.083677 (0.02434)	0.090544 (0.02228)	0.090273 (0.02213)
Average Wages		233.1733 (32.5365)	159.0188 (24.2911)	220.9955 (24.4659)	150.5093 (19.6396)
LEA Unemployment	Using the National On-line Manpower Information System (NOMIS), unemployment figures were collected for each month for each Local Education Authority (LEA). Rates were then calculated by dividing through by Census population figures of the number of individuals in the labour market for each LEA. The 1981 Census is used for the period to 1991 and the 1991 Census is used thereafter.	10.01333 (4.13270)	9.922007 (4.14132)	11.11412 (4.36587)	10.96832 (4.31522)
Sample Size		**4115**	**4770**	**2034**	**2552**

Table C.1: Summary Statistics of the Four YCS Subsets Used in this Study

C.4 Figures

Expected Wage	Reservation Wage: Under 27.5	100	125	150	175	200	225	250	Over 250	Total
Under 27.5	3	10	2	1	0	0	0	0	0	16
100	21	3305	84	35	11	3	4	1	2	3466
125	2	2539	797	16	9	0	1	2	0	3366
150	0	583	832	210	2	4	1	0	1	1633
175	0	133	390	216	85	2	2	0	1	829
200	0	75	135	101	49	29	0	0	0	389
225	0	15	34	49	27	18	11	0	0	154
250	0	18	26	20	40	22	9	10	0	145
Over 250	0	7	16	9	13	12	8	17	18	100
Total	26	6685	2316	657	236	90	36	30	22	10098

Table C.2: The Joint Distribution of Reservation and Expected Wages

Cohort	Of School Leaving Age In:	Sample Size	Year in Which Sweep Was Conducted: 1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	Final Response Rate
1	1983-84	8064	1 (69%)	2 (75%)	3 (84%)								43%
2	1984-85	14430		1 (74%)	2 (80%)	3 (83%)							49%
3	1985-86	16208			1 (77%)	2 (76%)	3 (76%)						44%
4	1987-88	14116					1 (71%)	2 (74%)	3 (78%)				41%
5	1989-90	14511							1 (72%)	2 (75%)	3 (77%)		42%
6	1990-91	29922								1 (69%)	2 (74%)	3 (75%)	40%

Table C.3: The Cohort Design Structure of the YCS Survey

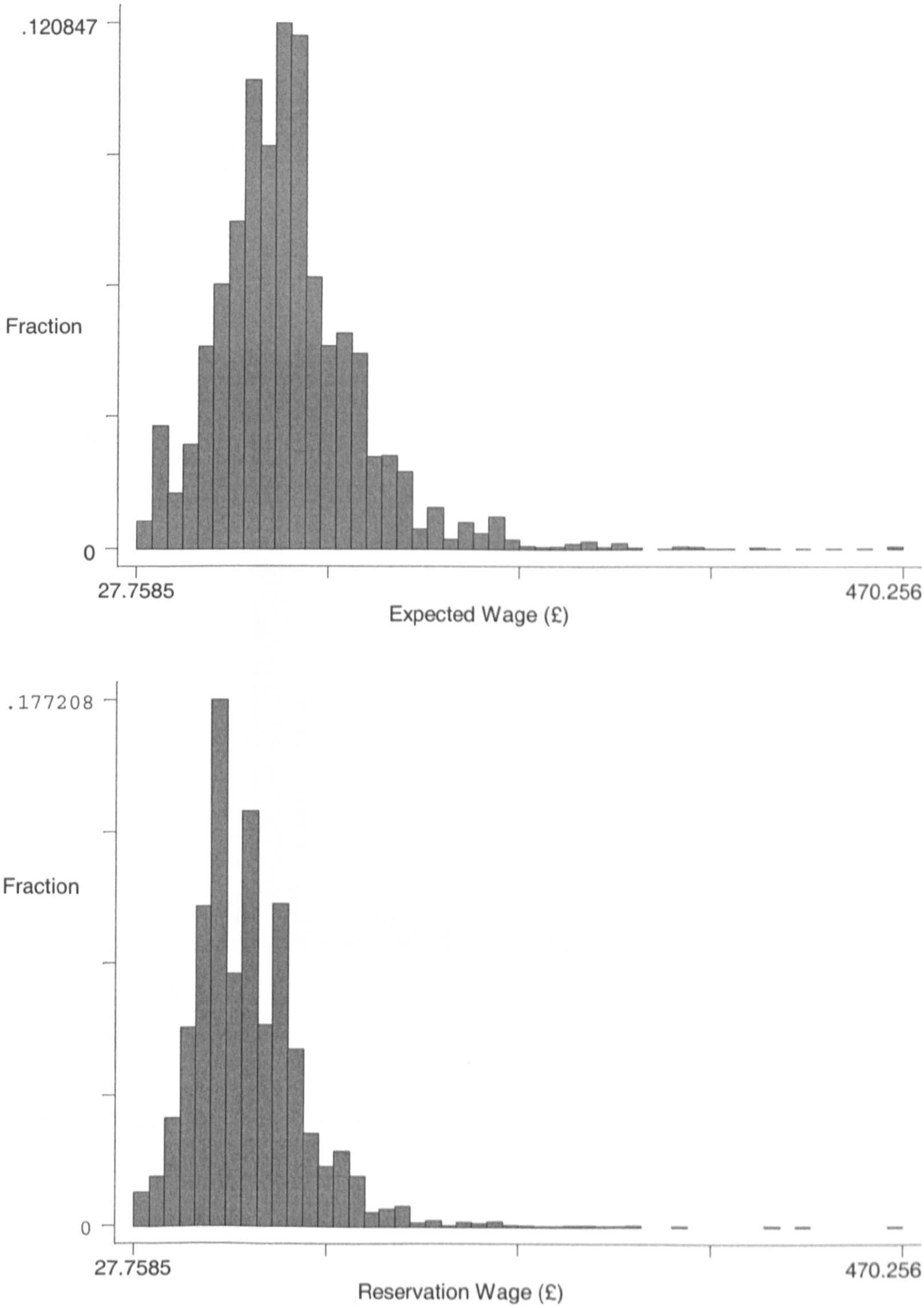

Figure C.3: Distributions of both Reservation and Expected Wages for All Cohorts

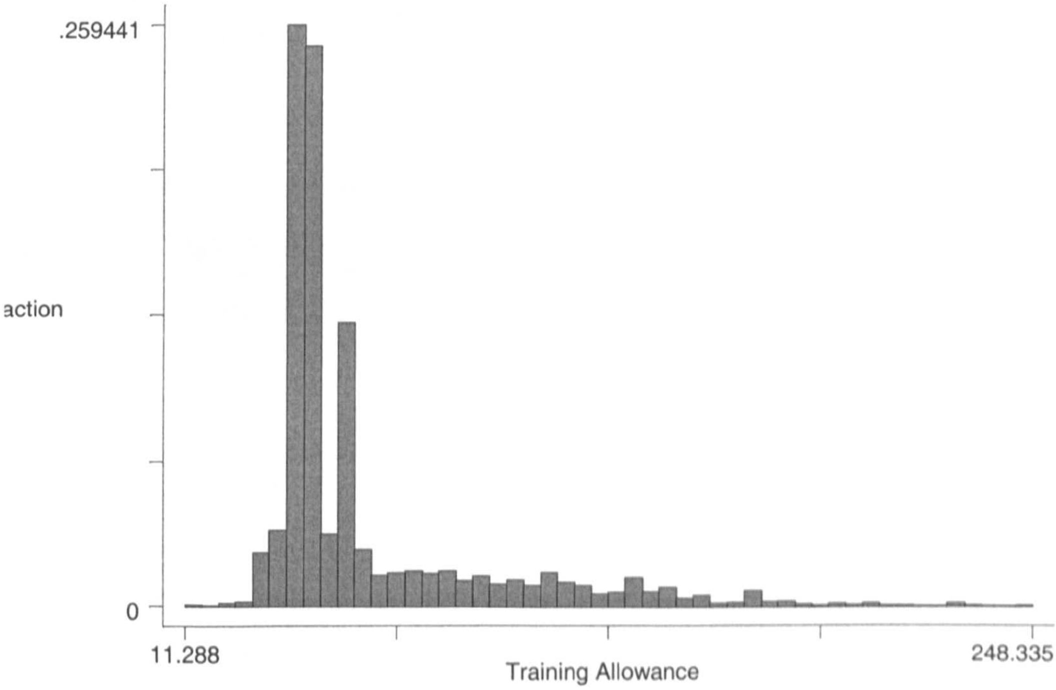

Figure C.4: Distributions of both Reservation and Expected Wages for All Cohorts

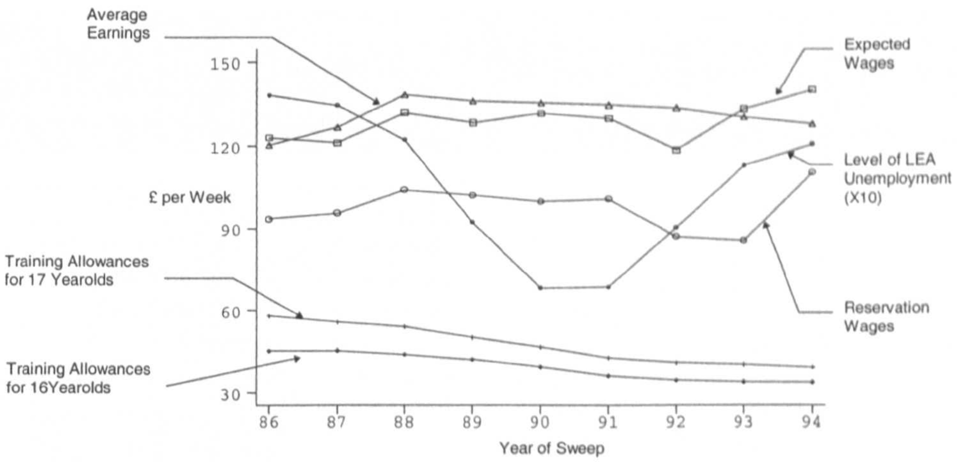

Figure C.5: Graph Showing the Paths of Various Economic Indicators for the Period of Study. Using the Male Matched Sample. Average Earnings (Source: Regional Trends), Reservation and Expected Wages are Gender Specific. LEA unemployment is multiplied by 10.

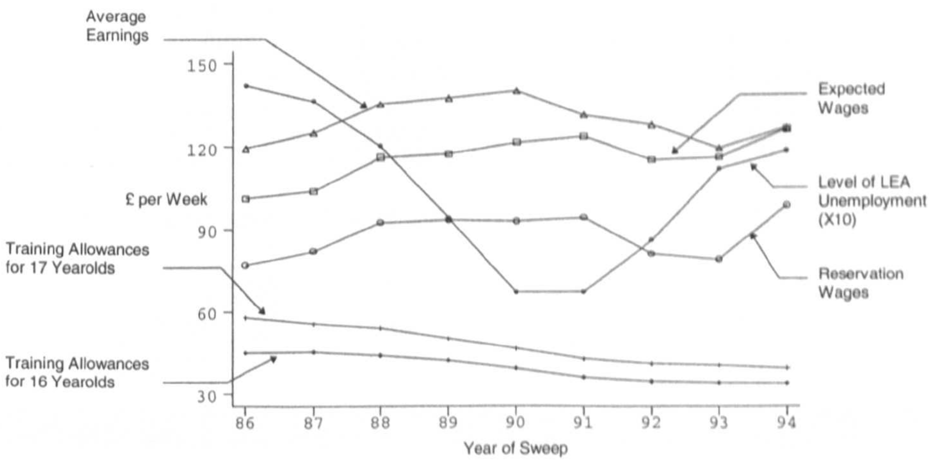

Figure C.6: Graphs Showing the Paths of Various Economic Indicators for the Period of Study. Using the Female Matched Sample

Estimated Probit Models for Calculation of Propensity Scores				
	Men		Women	
Regressors	Coefficients	Standard Error	Coefficients	Standard Error
Ethnic Origin	**-0.27955**	0.090444	-0.12212	0.0853204
Education Score at 16	**-0.03625**	0.004897	**-0.04007**	0.0041856
Maths GCSE	**-0.29432**	0.084958	-0.00719	0.0744607
English GCSE	0.046779	0.080678	0.008856	0.0637166
No. of Siblings	0.016166	0.015491	**-0.02473**	0.013766
Career Service Interview	**0.139481**	0.063511	0.089003	0.0636125
Youth Unemployment	-0.89669	1.68043	-1.54819	1.519593
Regional YT Places	4.712332	3.767507	**7.657992**	3.422754
Average Wages	**0.034929**	0.003623	**0.042935**	0.0039148
Live With Parents	-0.04585	0.147203	0.076554	0.0936863
LEA Unemployment	0.005294	0.008308	**0.015029**	0.0074929
Yorkshire & Humberside	0.048007	0.109685	**-0.21767**	0.0937065
East Midlands	**-0.47221**	0.10109	**-0.60949**	0.0843748
East Anglia	-0.06355	0.161182	-0.14231	0.1467006
Greater London	-0.10677	0.158371	0.028652	0.1406996
South East	**-0.40535**	0.179728	**-0.35222**	0.1591976
South West	**-3.52576**	0.374787	**-3.12641**	0.3006614
West Midlands	**-2.38212**	0.261426	**-1.98503**	0.2099086
North West	**-0.31698**	0.165377	**-0.29694**	0.1475537
Wales	0.054306	0.117229	**-0.24516**	0.1005625
Cohort 2	-0.0661	0.099427	-0.13534	0.0927769
Cohort 3	**-0.77264**	0.13317	**-0.68744**	0.1213671
Cohort 4	**-0.96933**	0.16877	**-0.95999**	0.149922
Cohort 5	**-1.50203**	0.175574	**-1.43872**	0.1605171
Cohort 6	**-1.9456**	0.1705892	**-2.01512**	0.167976
Constant	**-7.15183**	0.8751402	**-6.12655**	0.6709916
$\chi^2(25)$	702.74		717.54	
Prob. > χ^2	0.0000		0.0000	
Pseudo R^2	0.1520		0.1285	
Sample Size	4080		4794	

Table C.4: Male and Female Probit Model Estimations of YTS Participation

Regressors	Men: Reservation Wages Coef.	Std. Error	Men: Expected Wages Coef.	Std. Error	Women: Reservation Wages Coef.	Std. Error	Women: Expected Wages Coef.	Std. Error
Ethnic Origin	10.89702	3.464913	15.45376	4.837764	-0.7331	2.108997	2.938049	2.804193
Education Score at 16	-0.4328	0.219833	-0.48605	0.306934	0.061281	0.124128	0.365425	0.165045
Maths GCSE	-0.17245	2.87702	-4.4208	4.016939	2.713062	2.128552	-0.91137	2.830194
English GCSE	2.849834	2.333186	9.246686	3.257629	2.778337	1.661183	-1.79623	2.208765
No. of Siblings	0.062976	0.501925	-0.40546	0.700795	-0.27913	0.330924	-1.02411	0.440008
Career Service Interview	4.612076	1.792269	6.053792	2.502393	-1.34014	1.429507	-0.57241	1.900721
Youth Unemployment	-189.823	45.1496	-213.118	63.03854	-123.306	36.04723	-79.8102	47.92961
Regional YT Places	9.139832	122.273	31.0028	170.7193	310.8223	117.4818	51.41146	156.2078
YTSI	33.63028	13.58253	32.92925	18.96413	8.759707	6.803698	-8.38859	9.046426
YTSII	17.21374	6.120887	17.90767	8.546074	4.548775	4.144305	-2.17982	5.510407
YT	67.00084	19.53328	72.06455	27.27267	15.60367	9.103098	-3.72546	12.10379
Average Wages	0.593443	0.139266	0.760058	0.194445	1.300284	0.121493	1.432638	0.161541
Pred. Training Allowance	124.6503	42.01457	136.9275	58.66138	50.7928	25.52927	-6.55406	33.94458
Live With Parents	5.848049	4.619193	8.146328	6.449386	-12.5849	3.478181	-8.12764	4.624706
LEA Unemployment	0.74072	0.30864	0.733179	0.430928	-0.53105	0.193458	-0.85525	0.257228
Yorkshire & Humberside	0.042642	2.994493	2.386416	4.180956	1.958798	2.574407	-1.72809	3.423018
East Midlands	-0.72384	3.285392	-1.53823	4.587114	-8.81026	2.711551	-4.40834	3.605369
East Anglia	14.95051	4.953959	17.77626	6.916791	8.421936	4.315854	1.690732	5.738505
Greater London	2.185735	4.383977	7.481988	6.120975	11.2164	4.23051	6.220016	5.625029
South East	4.912493	5.527611	6.95248	7.717733	0.811709	4.053202	-1.66377	5.389274
South West	-49.4123	10.9313	-55.2298	15.26245	-46.3038	7.834674	-49.474	10.41725
West Midlands	-26.946	7.719467	-27.5619	10.77804	-34.2984	5.433954	-47.9169	7.225168
North West	-0.55041	4.601312	3.317585	6.424421	5.315443	5.193808	-4.91661	6.905862
Wales	-2.69004	3.4771	-0.77979	4.854779	-0.06344	2.72888	1.485483	3.628411
Cohort 2	0.039398	2.642059	-2.5973	3.688883	0.489976	2.251418	-4.55783	2.993561
Cohort 3	8.467371	11.3175	-1.38315	15.80167	-10.436	4.52875	-28.9266	6.021578
Cohort 4	3.692701	12.41048	-3.26482	17.3277	-15.523	5.893457	-33.902	7.836138
Cohort 5	-28.5584	10.42568	-32.7629	14.55648	-32.9492	6.141893	-43.1951	8.166468
Cohort 6	-46.5554	5.561294	-56.113	7.764762	-46.6177	5.127124	-50.3621	6.817196
Constant	-541.831	162.2768	-598.091	226.5733	-295.898	119.9743	-22.587	159.5219
F(29,2004) or F(29,2522)	9.93		8.05		13.00		12.80	
Prob. > F	0.00		0.00		0.00		0.00	
R^2	0.1256		0.1043		0.1301		0.1283	
Sample Size	2034		2034		2552		2552	

Table C.5: Male and Female Regression Model Estimations for Reservation and Expected Wages

Regressors	log of Training Allowance	Standard Error	log of Wages	Standard Error	log of Reservation Wage	Standard Error
	Dependent Variables					
Black	**-0.0488**	**0.0310**	**0.0115**	**0.0243**	**0.0549**	**0.0264**
Asian	-0.0296	0.0232	0.0501	0.0170	0.0503	0.0179
Other Ethnic Origin	**-0.0117**	**0.0312**	**0.0095**	**0.0178**	**0.0024**	**0.0236**
Gender	0.1780	0.0364	-0.0803	0.0228	0.4054	0.0324
Education Score at 16	**0.0055**	**0.0006**	**0.0052**	**0.0004**	**0.0037**	**0.0006**
Maths GCSE	0.0236	0.0099	0.0024	0.0063	0.0067	0.0109
English GCSE	**0.0006**	**0.0090**	**0.0267**	**0.0059**	**0.0050**	**0.0102**
YTS Type I	**-0.3218**	**0.0335**	**-0.0700**	**0.0113**	**-0.0665**	**0.0149**
YTS Type II	-0.2317	0.0202	-0.0242	0.0066	0.0089	0.0120
YT	**-0.4517**	**0.0094**	**-0.0287**	**0.0086**	**0.0989**	**0.0149**
Employment Training	-	-	-0.0518	0.0195	0.0632	0.0415
Government Scheme	**-**	**-**	**-0.0392**	**0.0110**	**0.0292**	**0.0187**
Siblings	**-0.0016**	**0.0025**	**-0.0024**	**0.0017**	**0.0020**	**0.0023**
School Job Interview	-0.0045	0.0084	-	-	-0.0058	0.0102
Ever Turned Job Down	**-**	**-**	**-**	**-**	**0.0331**	**0.0107**
Ever Applied for Job	-	-	-	-	-0.0423	0.0072
Regional YT Places	**-1.6785**	**0.6313**	**0.4779**	**0.3890**	**-0.7483**	**0.4623**
Regional Wages	0.0028	0.0005	-0.0003	0.0003	0.0065	0.0004
Lives with Parents	**0.0473**	**0.0216**	**-**	**-**	**-0.0142**	**0.0181**
LEA Unemployment	-0.0070	0.0013	-0.0040	0.0009	-0.0034	0.0011
Yorkshire & Humberside	**0.0498**	**0.0128**	**-0.0085**	**0.0108**	**0.0154**	**0.0159**
East Midlands	0.0695	0.0130	-0.0043	0.0105	-0.0007	0.0144
East Anglia	**-0.0140**	**0.0274**	**0.0123**	**0.0180**	**0.0214**	**0.0237**
Greater London	0.0115	0.0265	0.0222	0.0176	0.0044	0.0225
South East	**-0.0067**	**0.0294**	**0.0626**	**0.0194**	**0.0099**	**0.0256**
South West	-0.1003	0.0601	0.2914	0.0362	-0.2837	0.0454
West Midlands	**-0.1177**	**0.0473**	**0.1519**	**0.0279**	**-0.2146**	**0.0333**
North West	0.0059	0.0291	0.0499	0.0189	0.0067	0.0240
Wales	**-0.0220**	**0.0160**	**0.0057**	**0.0128**	**0.0419**	**0.0167**
Cohort 2	**0.0472**	**0.0237**	**-0.0545**	**0.0207**	**0.0345**	**0.0158**
Cohort 3	-0.1854	0.0407	0.0337	0.0164	-0.1349	0.0183
Cohort 4	**-0.2234**	**0.0415**	**0.0316**	**0.0174**	**-0.1345**	**0.0216**
Cohort 5	-0.3234	0.0397	-0.0458	0.0170	-0.4140	0.0224
Cohort 6	**0.2098**	**0.0395**	**-0.0681**	**0.0170**	**-0.3239**	**0.0214**
Constant	**3.8049**	**0.1250**	**4.7167**	**0.0739**	**3.3255**	**0.1062**
SE	**0.3752**	**0.0012**	**0.3486**	**0.0009**	**0.3025**	**0.0023**

Table C.6: Tobit Regression Estimates of Training Allowances, Wages and Reservation Wages

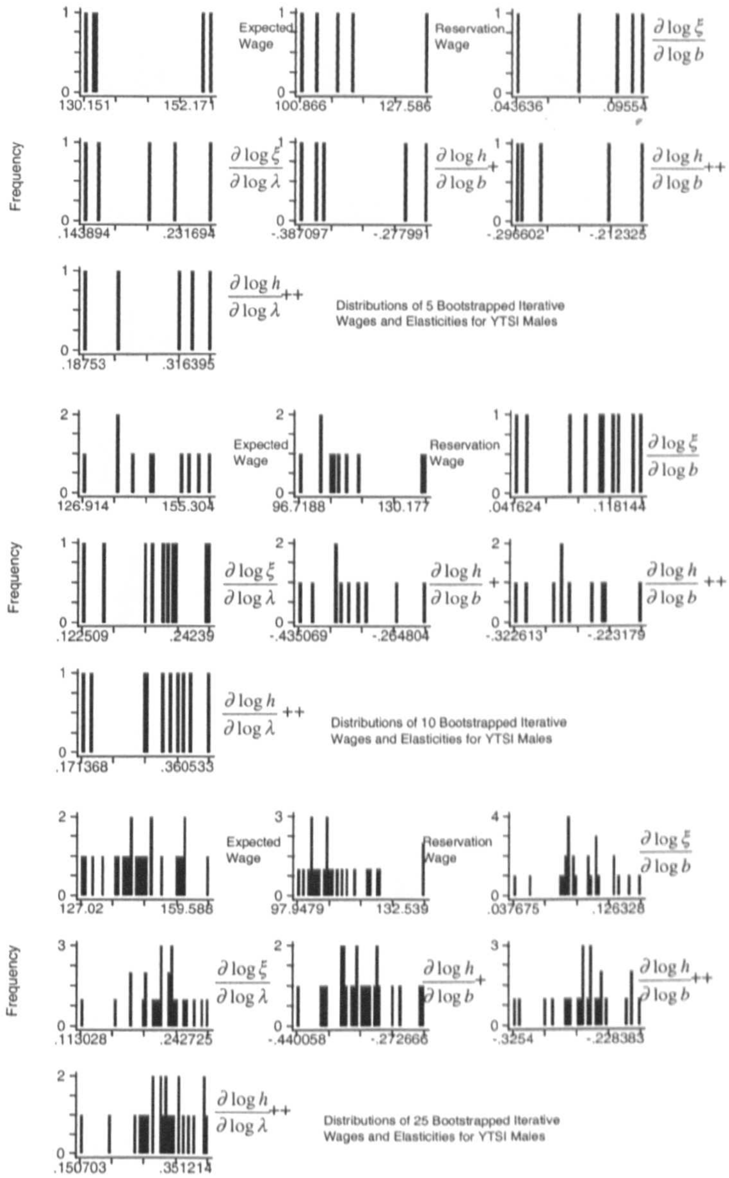

Figure C.7: Distributions of Estimated Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching for Various Sample Sizes (Males)

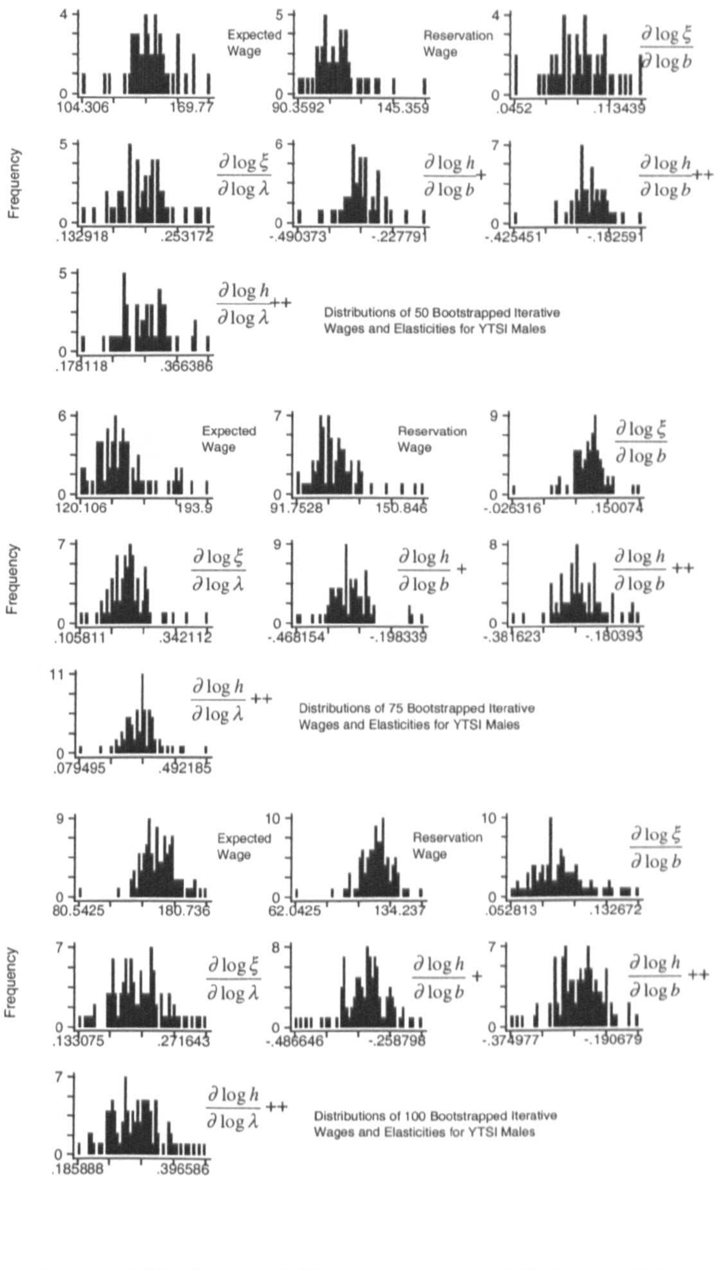

Figure C.8: Distributions of Estimated Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching for Various Sample Sizes (Males)

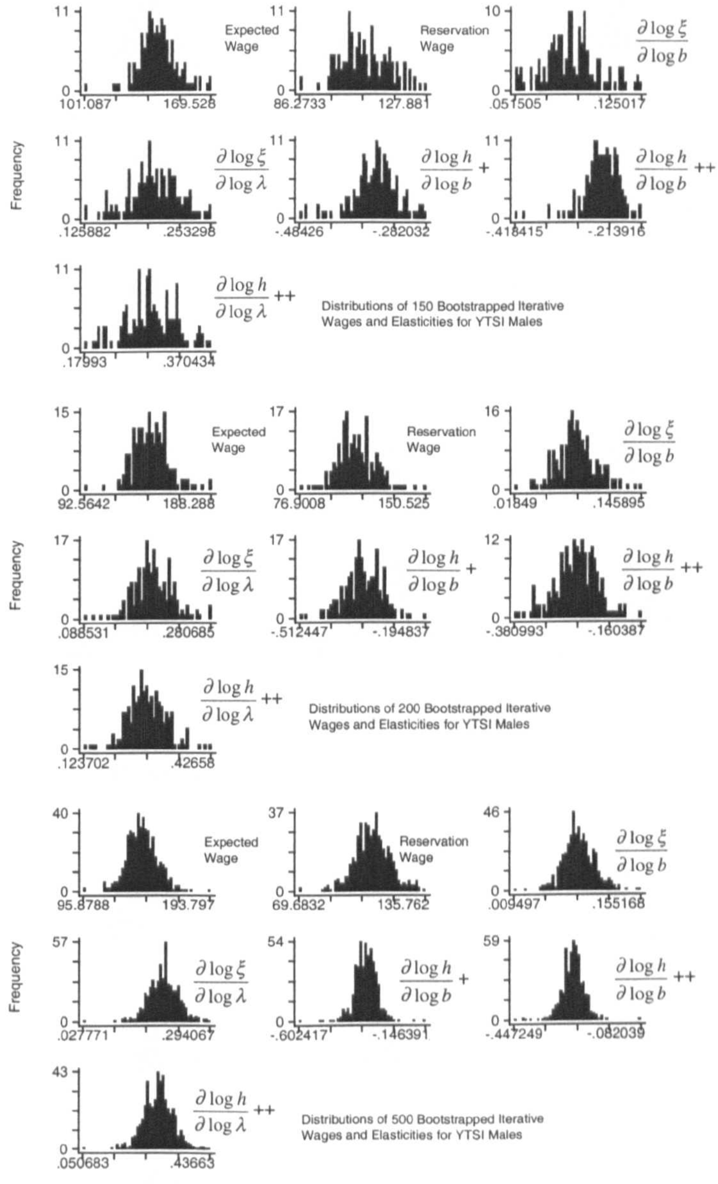

Figure C.9: Distributions of Estimated Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching for Various Sample Sizes (Males)

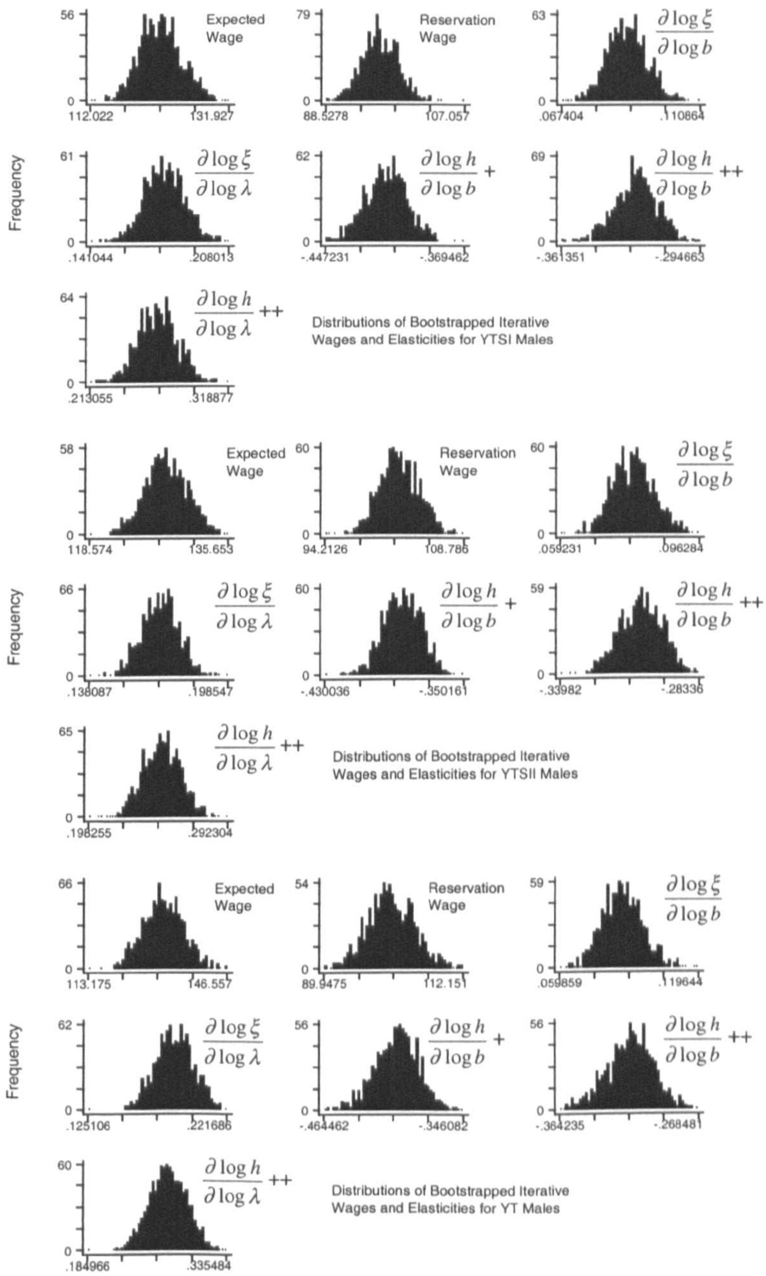

Figure C.10: Distributions of Estimated Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching (Males)

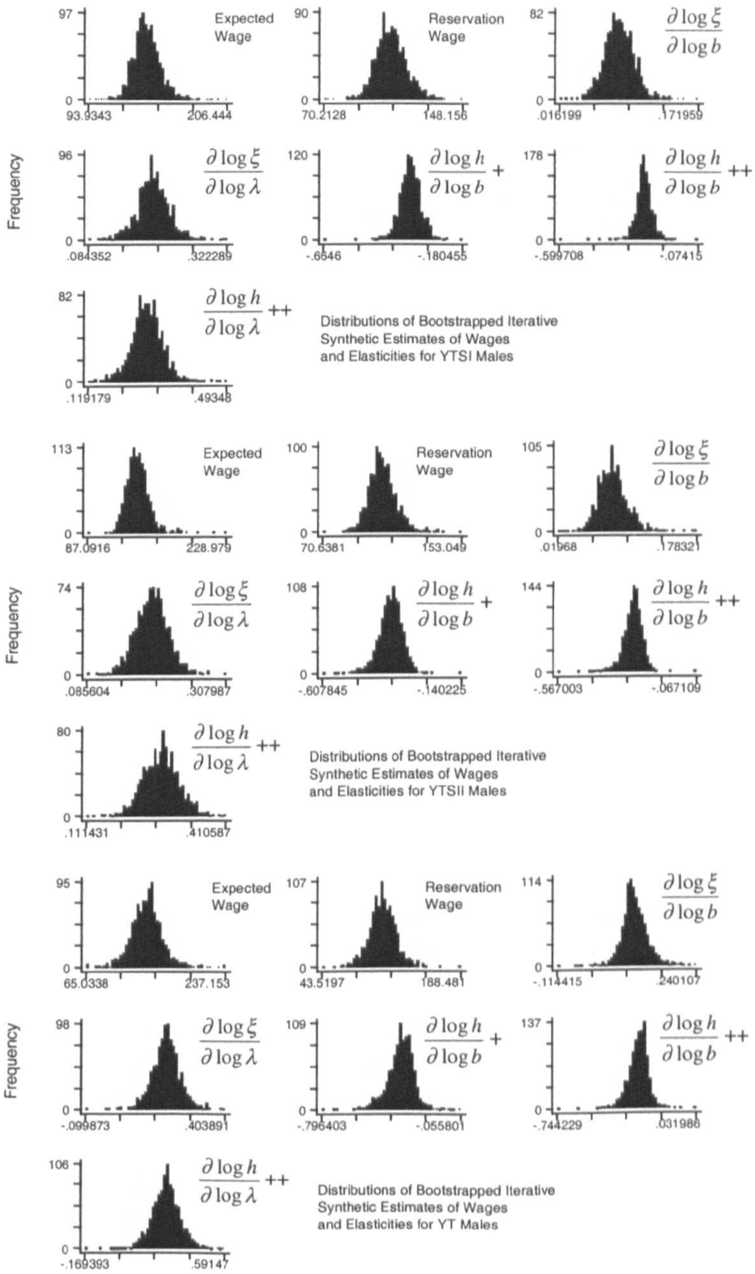

Figure C.11: Distributions of Estimated Synthetic Control Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching (Males)

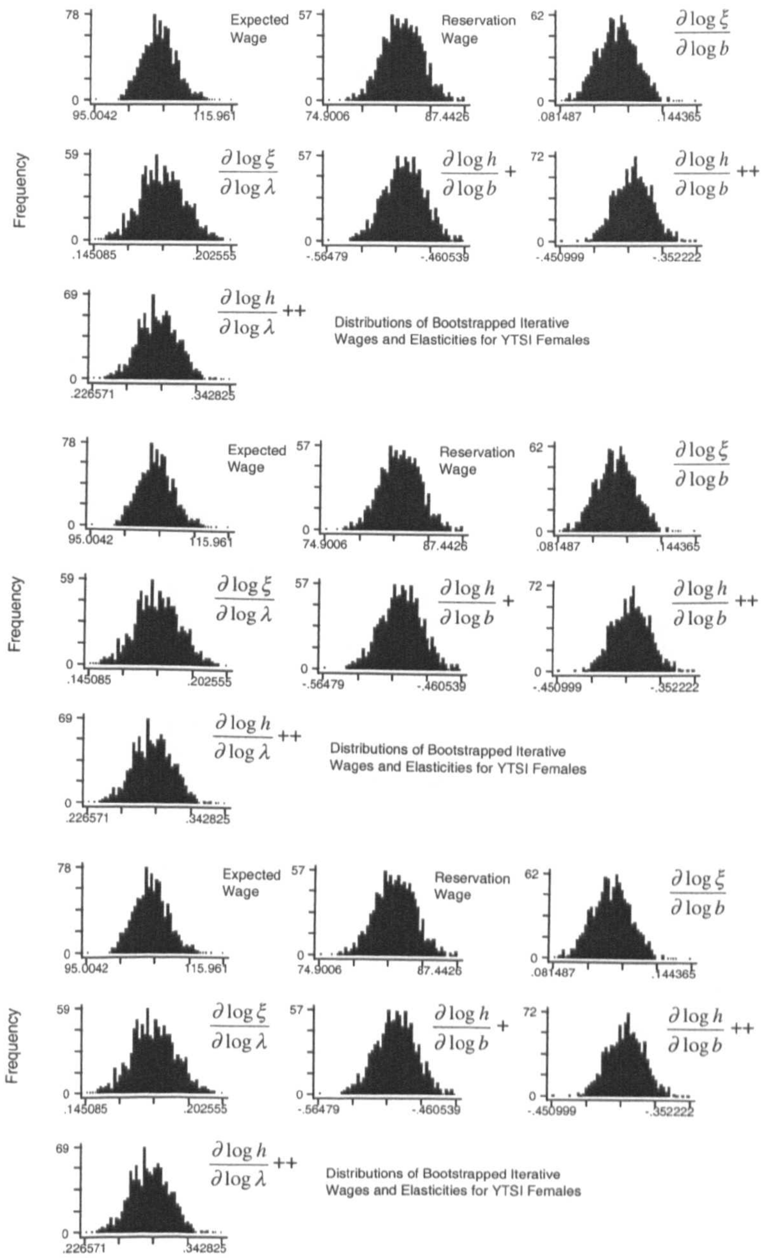

Figure C.12: Distributions of Estimated Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching (Females)

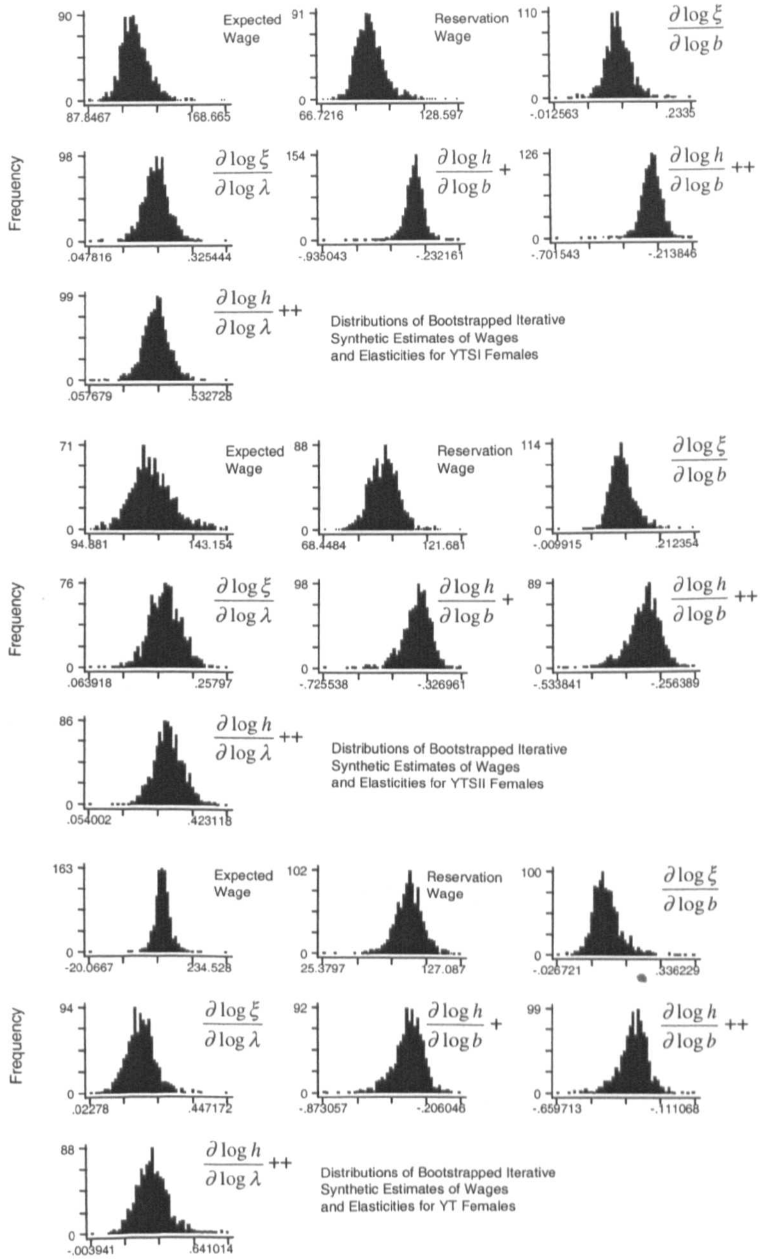

Figure C.13: Distributions of Estimated Synthetic Control Elasticities and Wages Using Bootstrapped Iterative Local Linear Regression Matching (Females)