Evaluating the labour market impact of Working Families' Tax Credit using difference-in-differences

Richard Blundell, Mike Brewer and Andrew Shephard*

Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE

Final version

Abstract

A difference-in-differences methodology cannot identify the labour market impact of WFTC alone because other taxes and benefits changed at the same time as its introduction. However, a comparison of the change in employment rates for parents against adults without children should underestimate any positive labour supply impact of WFTC for lone parents. Using two different household surveys, we find WFTC and associated reforms increased lone parents' employment by around 3.6 percentage points (ppt). For couples with children, we find that WFTC and associated reforms had no significant effect on mothers' employment, and was associated with a -0.5ppt change in fathers' employment, with the reforms encouraging households to have one earner rather than two. Overall, these changes correspond to between 25,000 and 59,000 extra workers depending upon the data source used. Robustness analysis of our identifying assumptions is generally favourable to our conclusions for lone parents.

 Published by HM Revenue and Customs 2005. ISBN 1-904983-12-X. JEL classification: H24, H31, I38. Author affiliations: Richard Blundell (UCL and Institute for Fiscal Studies), Mike Brewer and Andrew Shephard (Institute for Fiscal Studies). Support for this paper came from the ESRC-funded Centre for Public Policy at the IFS, from the Inland Revenue, as part of the project called "Econometric Research: Impact of Working Families' Tax Credit", and from the Bank of England. Data from the Labour Force Survey is made available through the UK Data Archive (UKDA). Data from the Family Resources Survey is made available by the Department for Work and Pensions, and is also available at the UKDA. The authors are very grateful for helpful discussions and comments on earlier drafts from Stuart Adam, Mike Bielby, Ian Bond, Howard Reed, Mike Joyce and Jonathan Thomas. All errors and omissions remain those of the authors alone, and not of the institutions mentioned here.

1 Introduction

Working Families' Tax Credit (WFTC) replaced Family Credit (FC) in October 1999 as the main component of state support for low income working families with children. WFTC was central to the government's 'Making Work Pay' agenda, with the dual aims of improving incentives to work, and encouraging low-paid individuals to progress in the earnings distribution.[1]

In this paper we attempt to estimate the labour market impact of Working Families' Tax Credits using a difference-in-differences methodology. Other studies, such as Brewer and Gregg (2001), Francesconi and Van der Klaauw (2004), Gregg and Harkness (2003), and Leigh (2004), have estimated similar models for lone mothers and have arrived at similar conclusions. With the exception of Leigh (2004), all these studies restrict their analysis to estimating the impact of WFTC on lone mothers. We attempt to estimate the labour market impact on couples also, although unfortunately the results turn out to be not very robust. In a related paper, and using some of the same data as here, Brewer *et al.* (2005) pursue a different estimation strategy using a structural model of labour supply. For lone mothers in particular, Brewer *et al.* arrive at conclusions that are similar to those reached here, which increases our confidence in the results.

We begin by presenting describing the reform in Section 2.1 and in Section 2.2 we provide some basic facts concerning recent changes in employment rates. Section 2.3 then discusses the difference-in-differences methodology used in our analysis, together with the specification adopted. Our main results using Labour Force Survey data are presented in Section 3.1, where we also conduct some robustness analysis and experiment with time effects and heterogeneous responses. We then present some results using the Family Resources Survey in Section 3.4, and summarise our main results in Section 4.1. In Section 5 we discuss results obtained in other studies while Section 6 concludes.

[1] See Brewer and Shephard (2004).

2 Background

2.1 Description of reforms

Working Families' Tax Credit (WFTC) was introduced in the UK in October 1999 as a replacement to Family Credit (FC), and was fully phased in by April 2000. Eligibility for the programme depended on hours of paid employment, the number of children, income, capital and formal childcare costs. Couples were assessed jointly. Unlike the Earned Income Tax Credit in the US, there was no "phase-in": families fulfilling the work condition (an adult in the family unit must work 16 or more hours a week) were immediately eligible for the maximum credit, but earnings above a threshold - £90 a week in October 1999 - reduced the credit at a rate of 55% of net income (so each pound of earnings after income tax and national insurance reduced awards of WFTC by 55p; the combined WFTC-income tax-national insurance effective marginal tax rate for someone paying basic-rate income tax was 69%: see Brewer (2001)). Financial assets over £3,000 reduced the award; savings over £8,000 removed eligibility completely. There was a small extra credit for families where someone worked more than 30 hours a week, and support for childcare was paid in addition to this. Spending on Family Credit in 1998/9 was £2.4 billion (bn), and this rose in cash terms to £4.6bn by 2000/1 and £6.3bn by 2002/3 (implying real rises of 85% and 140%), and there was no attempt to present the reform as revenue neutral.

Although it owed much to its predecessor, two key differences between WFTC and FC were the generosity of WFTC and the payment mechanism.[2] WFTC was more generous than FC in three ways: it had higher credits, particularly those for young children, families could earn more before the credit was phased out, and it had a lower withdrawal rate. The change in the payment mechanism was that, while FC was paid direct as a cash benefit, WFTC was paid by employers through the wage packet (who are themselves reimbursed by the Inland Revenue) unless a couple collectively decided that the non-working adult should apply for and therefore be paid WFTC. WFTC also significantly changed the system of

[2] A detailed history of in-work benefits in the UK, and a comparison of WFTC and FC can be found in Blundell and Hoynes (2003), with shorter accounts in Blundell *et al.* (1999 and 2000) and Dilnot and McCrae (1999).

support for formal childcare costs. Under FC, childcare costs up to £60 (£100) a week for families with 1 (2) children could be disregarded before the credit was phased out, which only benefited families earning more than the earnings threshold. Under WFTC, there was a payable childcare tax credit. It was potentially much more generous than the FC childcare disregard, providing a 70% subsidy to the parent on costs up to £150 a week for families with two or more children up to age 15 (16 for disabled children), and was paid in addition to WFTC, rather than an income disregard (for couples, the eligibility condition was that both must be working 16 or more hours). One final change is that Family Credit treated child support (or maintenance) above £15 a week as income, but WFTC disregarded all child maintenance when calculating awards.

2.2 Trends in employment rates of parents

Since the introduction of WFTC, there have been changes in the employment rates of parents. Figure 2.1 shows that the proportion of working-age lone parents in employment has increased from around 47.0% in Summer 1996, to around 54.9% by Spring 2002 - an increase of almost 8 percentage points (ppt), with a 3ppt increase since WFTC was introduced. The trend for single women without children has been much more static over the period, averaging 77%. The aim of the following sections will be to estimate the extent to which the recent changes in the employment rate of lone parents is due to WFTC or other reforms. In fact, as is evident in Figure 2.1, the employment rate of lone parents began to converge with that of single women without children before WFTC was introduced. Differences in the way this is accounted for explains some of the differences in the estimates of the impact of WFTC obtained by this and other studies: we discuss this further in Section 2.4 and Section 5. The trend for single men (not shown) is similar to that for single women without children.

Figure 2.2 shows trends in employment rates for men in couples (both with and without children), mothers in couples, and childless women in couples. The figure shows that the three groups examined have very different rates of employment throughout the time period considered, but exhibit similar *changes* in these employment rates over this entire period.

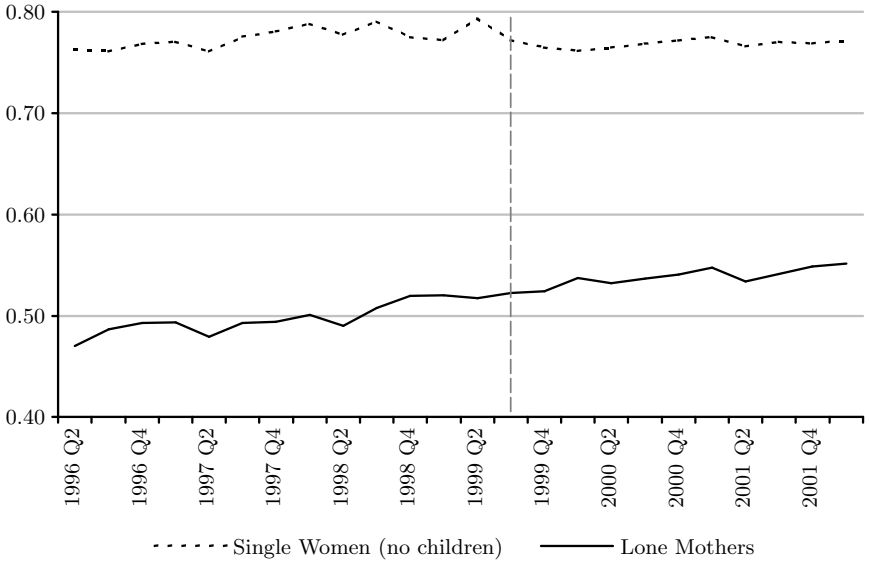

Figure 2.1: Employment rate for single women. Calculated using Labour Force Survey for Great Britain, various years.

For men in couples, the employment rate rises from 87.1% in Summer 1996 to 89.4% in Spring 2002. For women in couples without children the corresponding increase is from 75.3% to 77.5%, and for mothers in couples employment rises from 66.4% to 70.5% over this period. In particular, since WFTC was introduced in October 1999, the employment rate of mothers in couples has increased by a smaller amount than has that of women in couples without children (an increase of 0.5ppt compared to 1.2ppt). Again, our analysis seeks to estimate the extent to which these differential employment trends are due to the introduction of WFTC.

2.3 The Difference-in-Differences Methodology

Our goal is to identify the effect that Working Families' Tax Credit (WFTC) had upon the labour market outcomes of those who are potentially eligible to receive it (in other words, families with children). The evaluation problem is that we do not get to observe the outcomes that would have arisen had WFTC not been in existence. A difference-in-differences estimator exploits the existence of a comparison group in an attempt to estimate the impact of the treatment (in this case, WFTC) on the eligible group (in this case,

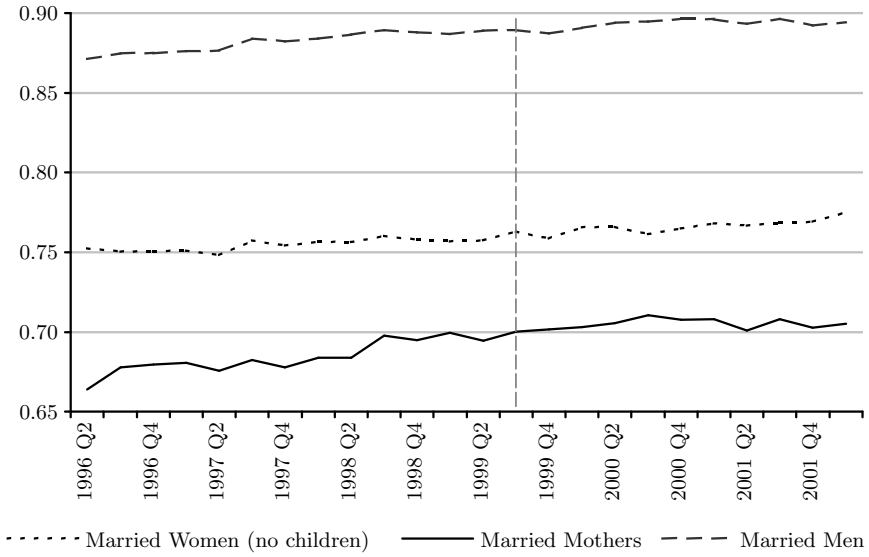

Figure 2.2: Employment rate for men and women in couples. Calculated using Labour Force Survey for Great Britain, various years.

parents). This relies on the assumption that the average change in employment status for the non-eligible group since the introduction of WFTC will reflect some underlying, non-WFTC related, employment trend. By subtracting this change from that of the WFTC eligible group, we are able to purge the latter change of any non-WFTC related components. This is the difference-in-differences estimator. The approach used follows that of Eissa and Liebman (1996), who used the difference-in-differences methodology to evaluate the impact of Earned Income Tax Credit (EITC) upon labour force participation in the United States. A more formal presentation of this technique is given in Appendix A.

The choice of comparison group is crucial in such analysis. A good comparison (or control) group should be as similar to the treatment group as possible in all dimensions other than eligibility for WFTC. Given the eligibility criteria for WFTC, a natural candidate for the comparison group are similar individuals without children. There are, of course, other possible candidates for comparison groups. For example, we could compare the changing employment outcomes of low-education lone mothers with those of high-education lone-mothers. The latter group may largely be ineligible for WFTC if education attainment and earnings potential are positively correlated (although we will, inevitably, misclassify some individuals to the extent that it is not a perfect predictor). For such a construction

of treatment and comparison groups, the division of workers into high and low education groups is, to some extent, arbitrary. Furthermore, by restricting the sizes of these treatment and control groups, it becomes more difficult to obtain precise estimates of the effect of the reforms.[3]

For our estimator to identify the impact of Working Families' Tax Credit, it must be true that there are no other contemporaneous shocks affecting the relative labour market outcomes of treatment and comparison groups. This assumption is almost certainly not satisfied with our choice of comparison group because the government was making other reforms at the same time as WFTC which may be expected to have had a differential effect on parents relative to non-parents. The most important of these is the increase in the generosity of child allowances in Income Support and jobseeker's allowance for parents (discussed in Brewer et al (2005)). The difference-in-differences method cannot separate the impact of this reform from that of WFTC: we are only able to identify the effect of the combined impact of all the government's reforms affecting parents. For lone parents, however, we may expect that these increases in Income Support generosity would have undermined work incentives for this group. With this in mind, it is likely that we are underestimating the 'true' effect of WFTC for this group. For couples with children, however, it is much less clear what effect other contemporaneous reforms have upon their work incentives, and so it is equally unclear whether our estimates will be systematically under- or over-estimating the effect of WFTC.[4]

Finally, while the assumption of common shocks is very strong (as the characteristics of treatment and comparison groups may differ on average), it does become more reasonable once we condition upon observable characteristics. This is discussed in more detail in the following section, and in Section 3.2 we will undertake some robustness analysis.

[3]Leigh (2004), which we discuss later, uses a number of control groups, and finds little variation in the estimated impact of WFTC.

[4]The way in which other estimates of the impact of WFTC treat the contemporaneous reforms is discussed later.

2.4 Empirical Specification

We estimate a model of employment, using a probit model.[5] The probit model will give estimates of how various explanatory factors affect the probability of an individual working. In particular, it will give us an estimate of the effect that the reforms affecting parents had upon labour market outcomes.

The outcome variable that we are interested in is an individual's employment status, equal to one if the individual works any number of hours, and zero otherwise. Given that we expect children to be an important determinant of an individual's decision to work, we include indicator variables for the number of dependent children (one, two, and three or more). Similarly, we include indicators for the age of the youngest child (under four, and between five and ten years). Indicator variables are also included for region of residence, housing tenure type, ethnic group (equal to one for non-white), and, where relevant, the employment status of the individual's partner. General economic conditions are controlled for by inclusion of a real deseasonalised GDP series, which may also be interpreted as acting like a general time trend; we also include seasonal controls.[6] Age enters our specification through a third-order polynomial, to allow for sufficient flexibility in fit; age may be considered as acting as a proxy for factors highly correlated with it, such as accumulated experience, which is not measured by the LFS. Educational attainment is controlled for by including a set of variables that indicate the age at which full time education was completed. The omitted category is individuals with less than sixteen years education, and indicator variables are included for education being completed at ages 16, 17 or 18, and 19 and above.[7] We construct a variable equal to one if the individual was observed after October 1999, the date when WFTC was introduced. Interacting this post-WFTC period variable with an indicator for the presence of dependent children gives the treatment variable. All other factors that affect participation are designated to the disturbance term

[5]The probit model is explained in Appendix B.

[6]We also experimented by including a polynomial time trend rather than using the real GDP series. The regression results turn out not to be sensitive to this particular assumption.

[7]Later in this section we experiment using more detailed educational controls. An advantage of this specification, however, is that it allows us to estimate a comparable regression using the FRS (see Section 3.4).

with the usual exogeneity assumptions imposed. A test of the null hypothesis that WFTC and contemporaneous reforms had no effect on labour force participation decisions therefore amounts to testing whether the estimated treatment effect is significantly different from zero.

Rather than considering the relevant WFTC period as being after it was introduced, a possible alternative is to define the period of interest as being after it was actually announced. Between these two events, there was an increase in the employment rate of lone mothers (see Figure 2.1) which Francesconi and Van der Klaauw (2004) interpret as a behavioural response in anticipation of the actual reform. However, Gregg and Harkness (2003) argue that this increase was due to other policy changes, rather than an announcement effect. We do not consider this alternative in this paper.

Our main regression results are obtained using data from the Labour Force Survey between Spring 1996 and Summer 2002.[8] To allow both for the phase-in period of WFTC, and for individuals to have sufficient time to re-optimise their behaviour following the introduction of WFTC, we drop observations from Summer 1999 to Spring 2000 inclusive.[9] We used two datasets for our analysis: the LFS and the FRS. Both of these are random surveys of households in Great Britain. Given that the LFS has a larger sample than the FRS, we would usually prefer to use the LFS for this sort of analysis; we present results from the FRS partly as another robustness check, and partly because the FRS is the data-set used in estimates based on a structural model of labour supply (see Brewer *et al.*, 2005). In both cases we omit observations for those individuals who are above working age,[10] and those who remain in full-time education. Those individuals who have found work but are waiting to start are classified as working for the purposes of our analysis. This latter group of individuals is small in size and our results are not sensitive to this particular classification.

[8]The Labour Force Survey is a representative survey of households in the UK, with sample sizes of around 60,000 in each quarter.

[9]The phase-in period exists because, until the end of March 2000, some individuals were still receiving Family Credit.

[10]Results are robust to restricting the sample to those individuals aged under 55.

3 Results

We first present the baseline results from the LFS, and for brevity, some of the regression output has been omitted from the presentation.[11]. Section 3.2 tests whether our identification assumptions are reasonable by using the same methods to test for a (hopefully non-existent) treatment effect in time periods when there were no major policy changes. Section 3.3 explores whether the treatment effect is the same for all individuals, by testing for time effects and variation by age of and number of children. Section 3.4 presents estimates based on the FRS.

3.1 Results from the Labour Force Survey

Results are shown below separately for single women, women in couples, single men and men in couples.

Single Women

The regression is first performed for single women, with the output displayed in Table 3.1.[12] This table, and all others, report the marginal effect for continuous variables, with all variables set to their mean values; this means that the numbers reported are the responses of an individual with average characteristics. For discrete variables, the figure reported is the change in probability as the variables changes in value from zero to one.[13]

The results show that children are an extremely important determinant of a single woman's decision to work: a single woman with one dependent child has a 10 percentage point (ppt) lower probability of working than if she had no children, whilst with two dependent children the probability is reduced by 15ppt, and with three children by almost 28ppt. If the youngest child is under five then there is a 32ppt lower probability of them working, while if they are between five and ten years old this probability is reduced by

[11] Full regression output is available upon request.

[12] All standard errors in this section correspond to the marginal effects rather than the coefficient estimates.

[13] All other variables continue to be evaluated at their mean value. See Appendix B for more details on the calculation of these marginal effects.

13ppt. The probability of work is monotonically increasing in the age at which education was completed, and a non-white individual in this group has a 5ppt lower probability of working.

Table 3.1: Regression Results, Single Women

	Marginal Effect	Standard Error	z	$P > \|z\|$
Age	0.445	0.031	14.33	0.000
Age Squared	−0.083	0.009	−9.48	0.000
Age Cubed	0.029	0.008	3.72	0.000
Children, 1	−0.099	0.005	−20.09	0.000
Children, 2	−0.147	0.006	−25.89	0.000
Children, 3+	−0.278	0.007	−38.05	0.000
Youngest Child, 0-4	−0.319	0.005	−58.54	0.000
Youngest Child, 5-10	−0.125	0.005	−24.19	0.000
Edage, 16	0.088	0.003	27.98	0.000
Edage, 17-18	0.159	0.003	46.31	0.000
Edage, 19+	0.172	0.003	45.84	0.000
Non-white	−0.054	0.003	−16.87	0.000
Post April 2000	−0.016	0.006	−2.71	0.007
Treatment Effect	0.036	0.005	7.63	0.000
Maximised Log Likelihood				-111843.75
Observations				233208

Note: The age variables are scaled as follows: Age is divided by 10, Age Squared by 100 and Age Cubed by 10000. Estimates for region of residence, housing tenure type, quarter, and real GDP have been omitted.

The estimate of the treatment effect for single women is 3.6ppt, a result which is significant at the 1% level. This suggests that WFTC, together with the contemporaneous reforms affecting parents, raised the employment rate of lone mothers by 3.6ppt. Moreover, as was noted in Section 2.3, this estimate is likely to underestimate the true effect of WFTC reform in isolation on lone mothers' employment.

Women in Couples

The specification of our model of employment for women in couples is identical to that for single women, with two exceptions: we include an indicator variable for whether the

woman's partner is working, and we interact this variable with the treatment effect, because economic theory suggests that recipients' responses will depend on the employment status of their partner. The net employment effect for women in couples will therefore depend upon these two treatment effects and the relative sizes of the two groups. Results are presented in Table 3.2.

Table 3.2: Regression Results, Women in Couples

	Marginal Effect	Standard Error	z	$P > \|z\|$
Age	0.055	0.003	18.52	0.000
Age Squared	−0.001	0.000	−14.27	0.000
Age Cubed	0.000	0.000	7.85	0.000
Children, 1	−0.052	0.003	−18.89	0.000
Children, 2	−0.097	0.003	−33.14	0.000
Children, 3+	−0.203	0.004	−54.95	0.000
Youngest Child, 0-4	−0.287	0.003	−97.28	0.000
Youngest Child, 5-10	−0.083	0.003	−29.67	0.000
Edage, 16	0.050	0.002	27.55	0.000
Edage, 17-18	0.088	0.002	44.17	0.000
Edage, 19+	0.104	0.002	49.10	0.000
Non-white	−0.063	0.002	−31.42	0.000
Post April 2000	0.025	0.004	6.74	0.000
Partner Working	0.235	0.002	113.09	0.000
Treatment Effect				
Partner Working	−0.001	0.003	0.42	0.672
Partner not Working	0.026	0.005	5.20	0.000
Maximised Log Likelihood				-254258.03
Observations				510542

Note: The age variables are scaled as follows: Age is divided by 10, Age Squared by 100 and Age Cubed by 10000. Estimates for region of residence, housing tenure type, quarter, and real GDP have been omitted.

The results give a qualitatively similar picture to those for single women. We find that the women's partners' employment status is highly correlated with their own employment status: other things equal, having a working partner raises the probability of work by almost 24ppt.

The treatment effect for women whose partners are not working is estimated to be

2.6ppt (significant at the 1% level), and, for women with working partners, the marginal effect is not significantly different from zero. Overall, given that the former group of individuals is much smaller than the latter, our analysis implies an approximately neutral, and statistically insignificant, treatment effect of WFTC and contemporaneous reforms of +0.4ppt for all women in couples.[14]

Single Men

The regression results for single men are presented in Table 3.3. Even though the proportion of lone fathers is very small, the sample size of the LFS is sufficiently large to detect a positive and significant effect of WFTC and contemporaneous reforms upon this group's employment status of 4.6ppt.

Table 3.3: Regression Results, Single Men

	Marginal Effect	Standard Error	z	$P > \|z\|$
Age	0.020	0.002	9.05	0.000
Age Squared	0.000	0.000	−4.49	0.000
Age Cubed	0.000	0.000	−1.73	0.084
Children, 1	−0.083	0.012	−7.48	0.000
Children, 2	−0.077	0.015	−5.37	0.000
Children, 3+	−0.136	0.023	−6.22	0.000
Youngest Child, 0-4	−0.203	0.021	−10.09	0.000
Youngest Child, 5-10	−0.137	0.015	−9.54	0.000
Edage, 16	0.080	0.003	25.44	0.000
Edage, 17-18	0.110	0.003	31.21	0.000
Edage, 19+	0.138	0.003	38.71	0.000
Non-white	−0.061	0.003	−19.57	0.000
Post April 2000	0.006	0.005	1.15	0.250
Treatment Effect	0.046	0.011	3.83	0.000
Maximised Log Likelihood				-109434.17
Observations				238615

Note: The age variables are scaled as follows: Age is divided by 10, Age Squared by 100 and Age Cubed by 10000. Estimates for region of residence, housing tenure type, quarter, and real GDP have been omitted.

[14] This calculation is based on the relative sizes of the two groups in the LFS data set.

Men in Couples

The results for men in couples are displayed in Table 3.4. The analysis here suggests that although the presence of children does have a significant impact on the outcomes of men in couples, they are relatively small in magnitude. We also find that the age at which education was completed has a much smaller impact upon the probability of work than for women, perhaps a consequence of the generally higher participation rates of male workers.

Table 3.4: Regression Results, Men in Couples

	Marginal Effect	Standard Error	z	$P > \|z\|$
Age	0.005	0.002	2.77	0.006
Age Squared	0.000	0.000	1.82	0.068
Age Cubed	0.000	0.000	−9.22	0.000
Children, 1	−0.005	0.002	−2.51	0.012
Children, 2	−0.001	0.002	−0.33	0.740
Children, 3+	−0.037	0.003	−14.34	0.000
Youngest Child, 0-4	0.022	0.002	11.36	0.000
Youngest Child, 5-10	0.006	0.002	2.99	0.003
Edage, 16	0.015	0.001	13.04	0.000
Edage, 17-18	0.026	0.001	17.78	0.000
Edage, 19+	0.042	0.001	30.20	0.000
Non-white	−0.033	0.001	−23.41	0.000
Post April 2000	0.012	0.002	4.94	0.000
Partner Working	0.144	0.001	121.23	0.000
Treatment Effect				
Partner Working	−0.010	0.002	−4.18	0.000
Partner not Working	0.005	0.002	2.19	0.029
Maximised Log Likelihood				-168345.78
Observations				531838

Note: The age variables are scaled as follows: Age is divided by 10, Age Squared by 100 and Age Cubed by 10000. Estimates for region of residence, housing tenure type, quarter, and real GDP have been omitted.

The estimated treatment effects imply that WFTC and related reforms increased employment rates of men whose partners were not working by +0.5ppt, and reduced employment rates of men whose partners were working by -1.0ppt. The average effect is a

statistically significant -0.5ppt. However, our robustness analysis undertaken in Section 3.2 reveals some concerns for this group.

3.2 Robustness Analysis

The validity of the results presented thus far depend upon the identifying assumptions made. If the assumptions are inappropriate then the strategy may fail to identify the treatment effect. It is therefore desirable to undertake some form of robustness analysis.

To do this we now repeat the original analysis for a hypothetical change in policy affecting working families with children. Specifically, using data from Spring 1996 to Spring 1998, we shall suppose that there was some policy change in Spring 1997. Given that such a change did not take place, if our estimation technique reports highly significant treatment effects then it would be suggestive that there are differential trends between the treatment and comparison groups that have not been appropriately controlled for. Moreover, it would cast doubt upon whether the marginal effects presented so far actually correspond to treatment effects. The results of our robustness analysis are presented in the lower panel of Table 3.5 (this table also summarises results from Tables 3.1 to 3.4). It would be desirable to perform this test for many other time periods, but there are few time periods where one can confidently state that there were no major policy changes with differential effects on parents and non-parents.

The marginal effects reported in Table 3.5 for both single women and single men are insignificant. This supports the idea that, conditional upon our vector of controls, the assumption of no differential trends may be satisfied. However, for couples we do find some positive and significant effects (the marginal effect for men in couples with a non-working partner is not strongly significant, however). Detecting a significant effect here is suggestive that there may be trends in this group that have not been adequately controlled for, and so our earlier result for women in couples in Table 3.2 and men in couples in Table 3.4 must be treated with caution. These results could be used to construct a trend-adjusted difference-in-differences estimator. Under the assumption that we have a time-invariant trend differential between treatment and comparison groups (as given in Table 3.5), we

Table 3.5: Treatment Effect using LFS Data: Original and Robustness Specifications

		Marginal Effect	Standard Error	z	$P > \|z\|$	Observations	Maximised Log Likelihood
Original Specification	Single Women	0.036	0.005	7.63	0.000	233208	-111843.75
	Single Men	0.046	0.011	3.83	0.000	238615	-109434.17
	Women in Couples						
	Partner Working	0.001	0.003	0.42	0.672	510542	-254258.03
	Partner not Working	0.026	0.005	5.20	0.000		
	Men in Couples						
	Partner Working	−0.010	0.002	−4.18	0.000	531838	-168345.78
	Partner not Working	0.005	0.002	2.19	0.029		
Robustness Specification	Single Women	−0.018	0.011	−1.62	0.105	74754	-36192.20
	Single Men	0.020	0.027	0.73	0.465	77662	-36074.17
	Women in Couples						
	Partner Working	0.013	0.006	2.08	0.038	166807	-83619.45
	Partner not Working	0.022	0.008	2.72	0.006		
	Men in Couples						
	Partner Working	0.006	0.005	1.16	0.246	173298	-56014.68
	Partner not Working	0.008	0.005	1.83	0.067		

Note: For details of the empirical specification, see Section 2.4.

can arrive at trend-adjusted estimate of the impact of WFTC and related reforms by subtracting the marginal effects in Table 3.5 from our original estimates. For example, for men in couples this would imply that the treatment impact of WFTC and related reforms was -1.6ppt if their partner was working and -0.3ppt if their partner was not working. The assumption of time-invariant trend differentials is strong, however.[15]

As a different kind of robustness check, we re-estimated the treatment effects using indicators for the highest qualification obtained, rather than the age at which full-time education was completed. The full set of treatment effects (original and robustness specifications) are in Table 3.6, and our results prove to be robust across these different educational classifications.

3.3 Time effects and Heterogeneous Responses

Our six reported treatment effects presented so far correspond to the change in employment since WFTC and contemporaneous reforms were introduced, averaged across all parents in each of the six groups. However, it is possible that employment responses to WFTC and contemporaneous reforms have changed over time, and that any such responses vary by observable characteristics of individuals within these six groups. In this section, we test for both these effects.

Time effects

Time effects can be estimated by including an additional treatment variable, interacting the presence of children with an indicator for the year 2001/02. The coefficient on this variable therefore captures an additional effect of WFTC and the other reforms affecting parents beyond that in 2000/01.[16] If, for example, the first coefficient were positive, and the second statistically indistinguishable from zero, this would suggest that WFTC and other reforms had a one-off positive effect upon employment. If, though, the second coefficient

[15] Gregg and Harkness (2003) allow for differential trends between lone parents and single people without children, and find it makes little difference to the estimated impact of WFTC and related reforms on lone parents.

[16] All results in this section use a regression specification that is comparable to that earlier in Section 3.1, when the age at which full-time education was completed, was used.

Table 3.6: Treatment Effect using LFS Data: Alternative Educational Controls

		Marginal Effect	Standard Error	z	$P > \|z\|$	Observations	Maximised Log Likelihood
Original Specification	Single Women	0.039	0.005	8.22	0.000	233208	-109908.26
	Single Men	0.049	0.011	4.24	0.000	238615	-109175.87
	Women in Couples						
	Partner Working	0.003	0.003	1.00	0.315	510542	-252219.78
	Partner not Working	0.030	0.005	6.14	0.000		
	Men in Couples						
	Partner Working	−0.013	0.002	−5.35	0.000	531838	-171054.67
	Partner not Working	0.004	0.002	1.63	0.103		
Robustness Specification	Single Women	−0.007	0.011	−0.67	0.501	74754	-35528.39
	Single Men	0.030	0.025	1.13	0.258	77662	-35976.48
	Women in Couples						
	Partner Working	0.015	0.006	2.40	0.016	166807	-83072.70
	Partner not Working	0.029	0.008	3.58	0.000		
	Men in Couples						
	Partner Working	0.002	0.005	0.44	0.660	173298	-56780.47
	Partner not Working	0.010	0.004	2.14	0.033		

Note: For details of the empirical specification, see Section 2.4.

were negative, then it would suggest that the initial positive response diminished over time. Results are presented in Table 3.7.

For single women, we find a one-off level effect in 2000/01, with no additional effect in 2001/02. Results for couples are mixed. We find that the estimated treatment effects are lower (more negative) for all four groups (i.e. men and women with and without working partners) in 2001 than in 2000. However, the fall in the treatment effect is only significantly negative for women in couples whose partners are working, and men in couples whose partners are not working. But the general picture remains that WFTC and related reforms had little distinguishable impact on the employment rates of individuals in couples.

Table 3.7: Multi-Period Treatment Effects

	WFTC effect		Additional April 2001 effect	
	Marginal Effect	$P > \|z\|$	Marginal Effect	$P > \|z\|$
Single Women	0.035	0.000	0.002	0.813
Single Men	0.049	0.001	−0.006	0.761
Women in Couples				
Partner Working	0.005	0.126	−0.011	0.019
Partner not Working	0.029	0.000	−0.011	0.235
Men in Couples				
Partner Working	−0.009	0.004	−0.005	0.233
Partner not Working	0.008	0.006	−0.009	0.040

Note: For details of the empirical specification, see Section 2.4.

Heterogeneous responses to WFTC

We allow for heterogeneous responses by interacting the treatment variable with other explanatory variables. In particular, we investigate how responses vary with the number of dependent children and the age of the youngest child; results are summarised in Table 3.8 and Table 3.9.

For lone mothers, we find that the positive treatment effect (of +4.5ppt) is limited to women whose youngest child is under eleven: these is no significant effect on women whose children are all aged over 11. There is little variation, though, by number of dependent children. For lone fathers, there is also a larger (more positive) treatment impact for those with very young children, and there is also evidence of variation by number of children, with the largest effect for those with three or more children. Note that it is not possible in the present difference-in-differences framework to determine whether these heterogeneous responses by the age of children are due to different elasticities or because of the different treatment of them in the tax and benefit system.

Results for couples are more mixed. For both men and women in couples whose partners are not working, average treatment effect are positive (see Tables 3.4 and 3.2). We find that these positive effects are limited to parents with children under 5: the treatment effects for those whose youngest child is of school-age are generally lower (less positive) and insignificant (with the exception of men whose youngest child is over 11, where we find a significantly negative treatment effect which is hard to rationalise).

For men and women whose partners are working, average treatment effects are negative (again, see Tables 3.4 and 3.2). We find that the treatment effect becomes more positive as the age of youngest child increases, so that the only groups with significantly negative treatment effects are those whose youngest child is under 5.

There is also some variation by number of children for individuals in couples. The negative impact of WFTC and related reforms on women with working partners appears to be concentrated amongst families with three or more children, but none of the treatment effects are significantly different from zero. However, the reverse is the case for men with working partners: we find a negative treatment effect if they have fewer than three children, and no significant effect otherwise. For men without working partners, the significant positive effects are limited to those with two or more dependent children.

Table 3.8: Treatment Effects by Age of Youngest Child

	0 - 4		5 - 10		11+	
	Marginal Effect	$P > \|z\|$	Marginal Effect	$P > \|z\|$	Marginal Effect	$P > \|z\|$
Single Women	0.044	0.000	0.045	0.000	0.010	0.201
Single Men	0.085	0.006	0.030	0.136	0.045	0.006
Women in Couples						
Partner Working	−0.006	0.117	0.008	0.057	0.007	0.132
Partner not Working	0.079	0.000	0.003	0.750	−0.015	0.077
Men in Couples						
Partner Working	−0.027	0.000	0.000	0.954	−0.007	0.052
Partner not Working	0.019	0.000	−0.005	0.239	−0.015	0.001

Note: For details of the empirical specification, see Section 2.4.

Table 3.9: Treatment Effects by Number of Children

	One Child		Two Children		Three+ Children	
	Marginal Effect	$P > \|z\|$	Marginal Effect	$P > \|z\|$	Marginal Effect	$P > \|z\|$
Single Women	0.032	0.000	0.041	0.000	0.038	0.000
Single Men	0.037	0.017	0.038	0.077	0.108	0.001
Women in Couples						
Partner Working	0.003	0.403	0.005	0.127	−0.011	0.018
Partner not Working	0.021	0.005	0.037	0.000	0.018	0.043
Men in Couples						
Partner Working	−0.014	0.000	−0.013	0.000	0.002	0.691
Partner not Working	0.000	0.923	0.008	0.011	0.007	0.057

3.4 Results from the Family Resources Survey

In this section we report results from performing the same analysis as was undertaken in Section 3.1, but using data from the Family Resources Survey (FRS). This therefore offers an additional robustness check on the results obtained to date. The specification has been chosen to match that of our earlier analysis as closely as possible, although we now have a less fine regional dissagregation, and have to measure education by the age at which the individual left full-time education.[17] Furthermore, the sample size is not sufficiently large to allow for time effects and heterogeneous responses as in Section 3.3. Rather than presenting the full set of results here, we instead present the treatment effects of both the main model and the robustness check in Table 3.10.

Results from the original specification are broadly consistent with those using the LFS. The results for single women are reassuringly similar - the estimation here suggests an impact of 3.7ppt compared with the 3.6ppt obtained when using the LFS. For single men we still find a positive effect (2.6ppt) although it is no longer significant (using the LFS we found a significant effect of 4.6ppt).

For women in couples, we fail to find a statistically significant treatment effect if their partner is working. The implied treatment effect when their partner is not working is positive and significant, but at 4.3ppt it is larger than was obtained using the LFS (2.6ppt). Similar effects are also obtained for men in couples, though the magnitudes are slightly larger than in the comparable LFS-based regressions. A man whose partner is working has an estimated treatment effect of -1.8ppt, whereas if they are not in work the treatment effect is about 1.5ppt (using the LFS data we obtained -1.0ppt and 0.5ppt respectively).

We performed the same robustness check as for the LFS, testing for a hopefully non-existent treatment effect in Spring 1997 (see Section 3.2). This fails to find an effect for single men and women, finds a significant effects for men in couples (a positive effect of around 2ppt regardless of their partners employment status), but not for women in couples. Again, this suggests that there may be some underlying trend differentials that have not

[17] We use data from Spring 1996 to Spring 2002. Observations from Summer 1999 to Spring 2000 are again omitted from our analysis.

Table 3.10: Treatment Effect using FRS Data: Original and Robustness Specifications

		Marginal Effect	Standard Error	z	$P > \|z\|$	Observations	Maximised Log Likelihood
Original Specification	Single Women	0.037	0.014	2.69	0.007	25163	-13077.50
	Single Men	0.021	0.031	0.81	0.418	23296	-11774.88
	Women in Couples						
	Partner Working	−0.006	0.009	−0.72	0.470	53416	-27140.79
	Partner not Working	0.043	0.015	2.74	0.006		
	Men in Couples						
	Partner Working	−0.018	0.007	−0.25	0.013	54399	-17488.91
	Partner not Working	0.015	0.006	2.39	0.017		
Robustness Specification	Single Women	0.008	0.021	0.38	0.705	11930	-6256.85
	Single Men	−0.070	0.051	−1.43	0.152	11321	-6050.44
	Women in Couples						
	Partner Working	−0.000	0.012	−0.04	0.970	26082	-13519.81
	Partner not Working	−0.008	0.022	−0.36	0.717		
	Men in Couples						
	Partner Working	0.019	0.009	2.09	0.036	26497	-8926.91
	Partner not Working	0.021	0.008	2.37	0.018		

Note: For details of the empirical specification, see Section 2.4.

been adequately controlled for, particularly for men in couples.

4 Summary of Results

In this section, we provide a summary of our results obtained from our difference-in-differences analysis using both LFS and FRS data. We also use grossing weights to give estimated population sizes of the different groups, and this allows us to estimate the number of individuals moving in or out of employment due to WFTC and contemporaneous reforms that affect parents.[18] In Table 4.1 the point estimates are presented together with the 95% confidence interval around them: in general, confidence intervals are larger with the FRS because it is a smaller sample.

The estimated impact on lone mothers based on the LFS is that WFTC increased the number of workers by 60,000, with a lower confidence bound of 45,000 and an upper confidence bound of 76,000. A very similar central estimate is obtained from the FRS (61,000), though the confidence interval around this estimate is larger (18,000 - 106,000).

The overall impact across all parents is +59,000 using the LFS, more than twice as large as the +25,000 obtained from the FRS. The main source of this discrepancy is the difference in the estimated effect on women in couples whose partner is working: the LFS suggests a small positive impact of 6,000, but the FRS suggests a fall of -27,000, but neither of these effects is significantly different from zero at the 5% level.

5 Other Difference-in-Differences Evaluations of WFTC

Other studies have also evaluated the labour market of WFTC. In this section, we summarise the findings from these studies. With the exception of Brewer et al (2005), all use the difference-in-differences methodology, and so, just as with our results, none can claim to estimate the impact of WFTC alone. All except Leigh (2004) and Brewer et al (2005) look only at lone parents (or lone mothers).

An early *ex-post* evaluation of WFTC on lone parents is reported in Brewer and Gregg

[18] This is based on an annual average over the treatment period.

Table 4.1: Summary of our Results

			Number of individuals		
		Marginal Effect	Central Estimate	Lower Bound	Upper Bound
Summary using LFS Data	Single Women	0.036	60,000	45,000	76,000
	Single Men	0.046	7,000	4,000	11,000
	Women in Couples				
	Partner Working	0.001	6,000	−21,000	32,000
	Partner not Working	0.026	17,000	11,000	23,000
	Men in Couples				
	Partner Working	−0.010	−41,000	−61,000	−21,000
	Partner not Working	0.005	10,000	1,000	18,000
Summary using FRS Data	Single Women	0.037	61,000	18,000	106,000
	Single Men	0.026	4,000	−5,000	12,000
	Women in Couples				
	Partner Working	−0.006	−27,000	−109,000	50,000
	Partner not Working	0.043	23,000	8,000	39,000
	Men in Couples				
	Partner Working	−0.018	−61,000	−109,000	−10,000
	Partner not Working	0.015	25,000	5,000	46,000

(2001). This used LFS data from 1997 to Summer 2000. Their results suggest that the initial effect of WFTC and related reforms was to increase the employment rate of lone parents by around 1.4 ppts, compared with single adults without children with similar characteristics. However, because a relatively small amount of data after WFTC was used, the confidence interval around this result was large.

A more recent study by Gregg and Harkness (2003) uses difference-in-differences techniques together with propensity score matching. Using data from the LFS and the General Household Survey, the paper estimates that the combined effects of the government's reforms affecting lone parents between 1998 and to 2002 was to raise the employment rate by around 5ppt, and increase the hours worked by those in employment.

Meanwhile, Leigh (2004) estimates that WFTC and related reforms raised the employment of both lone parents and mothers in couples by 1ppt, increased average hours by a single hour, and increased earnings by 4%. This study also uses LFS data, but makes use of the longitudinal aspect of this data-set, estimating a model allowing for individual fixed-effects: this effectively means that the treatment effects of WFTC and related reforms are identified from changes in individuals' moves into and out of work over a 15 month period. The paper focuses on a very short time period, comparing individuals observed just before and just after WFTC, and this may explain why the estimated impacts for lone parents are smaller than the other studies.

Using British Household Panel Survey data from 1991 to 2001, Francesconi and Van der Klaauw (2004) estimate that WFTC increased lone parent employment rates by 7ppt, driven both by both slower rates of exit from the labour market and by higher rates of entry into it.

6 Conclusion

Working Families' Tax Credit (WFTC) replaced Family Credit (FC) in October 1999 as the main component of state support for low-income working families with children. Using data from both the Labour Force Survey and the Family Resources Survey, this paper

uses difference-in-differences techniques to estimate the combined labour market impact of WFTC and contemporaneous reforms affecting parents.

We estimate that these reforms increased employment rates for lone mothers by 3.6 percentage points (ppts). The result is within the wide range of estimates from other studies of the impact on lone parents, and corresponds to around 60,000 more lone mothers in work. The response was larger for those with younger children and those with two or more children.

Most studies of WFTC have focused exclusively upon lone parents, but we also estimated the impact of WFTC and contemporaneous reforms on people in couples. We found no significant impact on the employment of mothers in couples, but a significant negative effect of -0.5ppt (-1.2ppt when using the FRS) for fathers in couples, driven by the estimated responses of men with working partners. There is evidence of heterogenous responses for people in couples, with the reforms having more impact (whether positive or negative) on families with young children. There is also limited evidence that the effect became more negative over time, although most of our estimated impacts are small, and few are statistically significant. However, our robustness check suggests that there may be differential trends between adults in couples with and without children which have not been adequately controlled for, and this would imply that the true impact of the WFTC and combined reforms on employment was more negative than estimated here.

Appendices

A Difference-in-Differences

We let y_i denote the employment outcome of individual i, with the convention that λ_1 corresponds to the treatment group and λ_0 to the control group. Similarly, we indicate the receipt of WFTC by θ_1 and non-receipt by θ_0. Because each individual's employment status is a binary variable equal to one if they are employed and zero otherwise, we are only interested in averages across individuals (hence the subsequent use of the expectations operator $\mathrm{E}[\cdot]$). Finally, we shall denote the period following the introduction of WFTC as t_1 and the period prior to this as t_0.

With this notation the treatment effect of WFTC is given by the quantity $\mathrm{E}[y_{t_1}|\lambda_1, \theta_1] - \mathrm{E}[y_{t_1}|\lambda_1, \theta_0]$. However, since we do not actually observe the counterfactual state of the world we cannot evaluate $\mathrm{E}[y_{t_1}|\lambda_1, \theta_0]$ and so by implication the treatment effect.

One potential way to estimate the effect that WFTC had upon the employment status of eligible individuals is to simply calculate the average difference between the outcomes of these individuals between time t_1 and t_0, i.e. evaluate the quantity $\mathrm{E}[y_{t_1}|\lambda_1, \theta_1] - \mathrm{E}[y_{t_0}|\lambda_1, \theta_0]$. The problem with this however, is that it assumes that the counterfactual is no change, so that had the WFTC not been introduced, employment for this group would have remained unchanged.

Difference-in-differences analysis addresses this issue by exploiting the presence of a comparison (non-eligible) group. If we assume that there are no differential trends in the labour market employment status of the treatment and comparison groups, then it follows that the difference-in-differences estimator of the treatment effect is given by,

$$(\mathrm{E}[y_{t_1}|\lambda_1, \theta_1] - \mathrm{E}[y_{t_0}|\lambda_1, \theta_0]) - (\mathrm{E}[y_{t_1}|\lambda_0, \theta_0] - \mathrm{E}[y_{t_0}|\lambda_0, \theta_0])$$

The term in the first set of parentheses gives the change in employment for the treatment group following the introduction of WFTC, whereas the latter term gives the change in employment of the comparison group over the same time period. Subtracting this second

term purges the first of any non-WFTC related employment changes, so that all that remains is the treatment effect.

Given that the characteristics of the treatment group may differ on average from those of the comparison group, for the empirical implementation it is necessary use regression analysis to augment our specification through inclusion of a vector $\boldsymbol{x}$ of control variables. Treatment is then given by a dummy variable D equal to one if the individual is in both the treatment group and in the treatment period. More formally, $D = 1(\lambda = \lambda_1)1(t = t_1)$ where $1(\cdot)$ is the indicator function. For more details, see Blundell and Dias (2002).

B The Probit Model

For the Probit model we assume that an individual's propensity to work y_i^* is linearly related to some set of observed characteristics, or: $y_i^* = \boldsymbol{x_i}'\boldsymbol{\beta} + \delta D + \epsilon_i$ where $\epsilon_i \sim \mathcal{N}(0, \sigma)$.

In practice we observe only the binary variable y_i, equal to one if an individual works (any hours), and zero otherwise.

$$y_i = \begin{cases} 1 & \text{if} \quad y_i^* > 0 \\ 0 & \text{if} \quad y_i^* \leq 0 \end{cases}$$

The probability than an individual works is therefore given by,

$$\begin{aligned} \text{Prob}\,(y_i = 1) &= \text{Prob}\,(y_i^* > 0) \\ &= \text{Prob}\left(\epsilon_i > -\boldsymbol{x}_i'\boldsymbol{\beta} - \delta D\right) \\ &= \text{Prob}\left(\frac{\epsilon_i}{\sigma} > \frac{-\boldsymbol{x}_i'\boldsymbol{\beta} - \delta D}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{-\boldsymbol{x}_i'\boldsymbol{\beta} - \delta D}{\sigma}\right) \\ &= \Phi\left(\frac{\boldsymbol{x}_i'\boldsymbol{\beta} + \delta D}{\sigma}\right) \end{aligned}$$

The likelihood function is then given by product of the probabilities across all individuals. Noting that the probability of not working is given by $\text{Prob}(y_i = 0) = 1 - \Phi(\boldsymbol{x}_i'\boldsymbol{\beta} + \delta D_i)$ it follows that,

$$\mathcal{L} = \textstyle\prod_{i=1}^{n} \Phi(\boldsymbol{x}_i'\boldsymbol{\beta} + \delta D_i)^{y_i}[1 - \Phi(\boldsymbol{x}_i'\boldsymbol{\beta} + \delta D_i)]^{1-y_i}.$$

The maximum likelihood estimates are the parameters which maximise the likelihood function $\mathcal{L}$. Note that only the ratios $\frac{\beta_k}{\sigma}$ and $\frac{D}{\sigma}$ are identified. For continuous variables the marginal effect of an explanatory factor upon the probability of work is given by the derivative

$$\tfrac{\partial \text{Prob}(y=1)}{\partial x_k} = \phi\Big(\tfrac{\boldsymbol{x}'\boldsymbol{\beta} + \delta D}{\sigma}\Big)\tfrac{\beta_k}{\sigma}$$

where $\phi(\cdot)$ is the standard normal density. Note that the marginal effect is a function of the explanatory variables. It is conventional to evaluate this at the mean value of the explanatory variables, so the marginal effects have the interpretation of corresponding to a representative, or average, individual.

By definition there can be no continuous change in the value of a discrete random variable. For dummy variables the marginal effect is therefore defined as the change in probability as the dummy changes in value from zero to one. For example, the marginal effect of treatment D is given by

$$\text{Prob}(y=1|D=1) - \text{Prob}(y=1|D=0) = \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}+\delta}{\sigma}\right) - \Phi\left(\frac{\boldsymbol{x}'\boldsymbol{\beta}}{\sigma}\right)$$

where again we evaluate all other variables (including those which are discrete) at their mean value.

References

[1] Blundell, R. & Dias, M (2000), 'Evaluation Methods for Non-Experimental Data', *Fiscal Studies*, vol. 21, No.4 427-468.

[2] Blundell, R., A. Duncan, J. McCrae & C. Meghir (2000), 'The Labour Market Impact of the Working Families' Tax Credit', *Fiscal Studies* vol. 21, No.1 75-104.

[3] Brewer, M., A. Duncan, A. Shephard, & M.J. Suárez (2005), 'Did Working Families' Tax Credit work? The final evaluation of the impact of in-work support on parents' labour supply and take-up behaviour in the UK', Inland Revenue Working Paper 2 (available from `www.hmrc.gov.uk/research/`).

[4] Brewer, M. & P. Gregg (2001), 'Lone Parents, the WFTC and Employment in Households with Children', in The State of Working Britain: Update 2001, (eds) R. Dickens, J. Wadsworth & P. Gregg, LSE Centre for Economic Peformance.

[5] Brewer, M. & Shephard, A. (2004), 'Has Labour Made Work Pay?', York: Joseph Rowntree Foundation.

[6] Eissa, N. & J.B. Liebman (1996), 'Labor Supply Responses to the Earned Income Tax Credit,' *The Quarterly Journal of Economics* 111: 605-637.

[7] Francesconi, M. & W. Van der Klaauw (2004), 'The Consequence of 'In-Work' Benefit Reform in Britain: New Evidence from Panel Data', ISER WP (available from `www.iser.essex.ac.uk/pubs/workpaps/pdf/2004-13.pdf`).

[8] Gregg, P. & S. Harkness (2003), 'Welfare Reform and Lone Parents in the UK', CMPO Working Paper Series No. 03/072.

[9] Leigh, A., (2004), 'Optimal Design of Earned Income Tax Credits: Evidence from a British Natural Experiment', mimeo, Australian National University, (available from `econrsss.anu.edu.au/ aleigh/pdf/Optimal Design of Tax Credits.pdf`).