The Economic Journal **107** (*September*), 1375–1390. © Royal Economic Society 1997. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden MA 02148, USA.

THE LABOUR SUPPLY, UNEMPLOYMENT AND PARTICIPATION OF LONE MOTHERS IN IN-WORK TRANSFER PROGRAMMES*

Paul Bingley and Ian Walker

In-work transfer schemes have recently been suggested as a device for encouraging labour force participation and reducing the severity of the disincentives associated with out-of-work income support schemes. Here we estimate a discrete choice labour supply model which allows for endogenous in-work welfare programme participation. We also allow for involuntary unemployment so as to distinguish between programme non-participation and an inability to obtain work and thereby generate an eligibility to the programme.

In-work transfer schemes have recently been suggested as a device for encouraging labour force participation and reducing the severity of the disincentives associated with out-of-work transfer programmes (such as Income Support (IS) in the United Kingdom, and Aid for Families with Dependent Children (AFDC) in the United States). Indeed the United Kingdom and the United States have in-work transfer schemes called Family Credit (FC) and the Earned Income Tax Credit (EITC) respectively and the expansion of EITC has been central to the direction of welfare reform in the United States (see Scholz (1994)). In-work transfer programmes have been suggested as a more appropriate policy for relieving poverty than alternatives such as a minimum wage because of their potentially beneficial effects on work incentives and a similar expansion to that for EITC has recently been proposed for FC to cover individuals without children.[1]

One difficulty with such schemes is that their design may be unavoidably complicated or personally intrusive to ensure that they are targeted sufficiently finely. The result may be that individuals may feel that participation in such programmes is stigmatised or not understand that they have some eligibility. The implication of such stigma or ignorance is that individuals may not participate in such programmes despite their potential entitlement. We use stigma as a generic term because we are empirically unable to distinguish between stigma, other costs or ignorance. The problem is known as 'programme non-participation'. FC is intended to counter the adverse work incentive effects of IS which itself is intended to guarantee a basic living standard – if it is to do this then it is important that it does not suffer from low participation rates. Yet this is precisely what we seem to observe.

* We are grateful to the European Union *Human Capital and Mobility* programme, the UK Economic and Social Research Council, and the European Union *Stimulation Programme for Economic Science* for support under grants ECBCHRXCT930225, R000232160, and E90100010 respectively. Financial support from the Economic and Social Research Council for the *Centre for the Microeconomic Analysis of Fiscal Policy* at the Institute for Fiscal Studies is also gratefully acknowledged. The data have been made available to us with the permission of the Controller of Her Majesty's Stationary Office by the ESRC Data Archive. We are grateful to seminar participants at Arhus, Keele, the IFS, and Bristol. The usual disclaimer applies.

[1] See DSS (1995*a*).

The difficulties in estimating the impact of such transfer programmes on labour supply behaviour are that: labour supply and programme participation decisions may be determined simultaneously, for example, if labour supply depends on the marginal wage and the marginal wage depends on whether one participates in an income transfer programme; and, individuals may choose not to participate in the labour market if some in-work benefit is stigmatised so it is necessary to distinguish between individuals who choose this position from those individuals who would want to work (and, perhaps, claim the in-work benefit) but simply cannot find a job – i.e. labour market rationing.

Here we estimate a labour supply model which allows for endogenous in-work welfare programme participation. We also allow for involuntary unemployment so as to distinguish between programme non-participation and an inability to obtain work. This is applied to the United Kingdom Family Credit programme for a sample of UK lone mothers. The estimates suggest that FC plays an important role in overcoming the adverse incentive effects associated with Income Support and that it does this without having an adverse effect on the probability of working full-time. However, the probability of working would rise further were it not for the apparent stigma associated with FC participation.

Our approach is distinct from previous US work as we are primarily concerned with an in-work transfer programme and many potentially eligible individuals may choose to remain out of work, because of FC stigma, and therefore not appear to be entitled. Thus, we also need to model the probability of wanting to work for individuals observed not to work. Reducing stigma would promote programme participation for those currently eligible and would promote out of work individuals to want to work and become eligible. Thus the methodology we employ extends recent US studies by modelling labour supply and programme participation using discrete choice methods to allow for one of the 'choices' to be involuntary unemployment.[2]

The main welfare programmes in the United Kingdom are Income Support (IS), Housing Benefit (HB) and Family Credit (FC). IS is the UK equivalent to the US Aid for Families with Dependant Children (AFDC) and is intended to ensure that household incomes[3] do not fall below some minimum. HB covers a proportion of the rent[4] and rates (a local property tax) of households not on IS. FC is payable to low income families but only if hours of work exceed some level – that is it is an in-work benefit.

In the United States AFDC has an unambiguous negative impact on work incentives (i.e. on compensated labour supply) and so there is then a conflict between reducing stigma and improving work incentives. Lowering the stigma associated with the programme induces extra participation in it and therefore

[2] Moreover, we allow for wages to be determined differently across employment states because of the large differential between part-time and full-time wages rates that is a feature of the UK labour market. The raw differential is of the order of 20% and Ermisch and Wright (1993) estimate differential returns to education and experience which imply a full-time/part-time differential of 8·5%.

[3] Net of housing costs including mortgage interest payments in many cases.

[4] But not loan payments for mortgagees.

reduces labour force participation and hours of work. In contrast in the United Kingdom, while IS has an unambiguously negative effect on work incentives, FC exhibits a notch in the budget constraint which increases the probability of working (although, because FC is means tested, it presumably acts as a disincentive to working long hours). IS has a programme participation rate that is close to 100% for lone mothers but improving participation in FC would have a beneficial effect on the incentive to participate in the labour force. Since programme participation is often found to be an increasing function of the level of entitlement this offers the virtuous prospect that an increase in FC entitlement could increase both programme participation and labour force participation.

Moffitt (1992) gives an extensive overview of the US literature but the only UK literature (see Blundell *et al.* (1987) for example) assumes that labour supply is exogenous. Moffitt (1983) was the first to analyse the decision not to participate in a welfare programme despite a positive potential benefit as a utility maximising decision subject to programme induced stigma[5]. The econometric framework is an hours Tobit with endogenous AFDC participation applied to a small sample of US female household heads. Disutility of welfare participation is assumed to be separable from the utility function, which implies this disutility affects programme participation, but not labour supply conditional upon benefit receipt. This assumption is maintained in most subsequent work applied to single decision maker households.

Keane and Moffitt (1991) extends earlier work by Fraker and Moffitt (1988). AFDC, food stamps and subsidised housing are modelled as a trivariate probit simultaneously with a discrete choice ordered probit for labour supply, as well as a wage equation. The notable feature of the methodlogy is that, because of the high dimensionality of the integration required, estimation involved the use of simulation methods to compute the likelihood.[6] They found that there are economies to joint programme participation. This is interpreted as stigma (since they assume that stigma arises from participation in a single programme and does not rise additively with additional programmes) or money-and-time costs (it is administratively easier to participate in subsequent programmes after single programme participation is established). The statistical significance of the elements of their error covariance matrix suggests that their approach would dominate a simple multinomial logit. The drawback of this approach is that it assumes that if full-time work is preferred to non-participation, then so is part-time work, so that utility comparisons are not made between all alternatives. The validity of this depends upon the budget constraint being convex.

In Hoynes (1996) AFDC-UP (Unemployed Parents) participation and simultaneous husband and wife discrete choice labour supply is modelled as a

[5] See Besley and Coate (1995*a*, *b*) for relevant theoretical arguments.

[6] Kiefer *et al.* (1995) notes that such simulation methods are computationally intensive, may suffer from numerical instabilities, poor convergence properties and are likely to be sensitive to starting values. They propose an alternative method based on a Monte Carlo EM algorithm which they demonstrate to have superior properties especially for large dimensional problems.

multinomial logit. It finds significant labour supply and programme participation effects of different AFDC withdrawal rates. The approach relaxes the assumption of separability between labour supply and programme participation which is a feature of previous work. This allows for a unordered choice-based framework which is appropriate when there are non-convex budget constraints. However, it does so at the cost of imposing the unpalatable assumption associated with a multinomial logit that the odds of choosing one position over another is independent of any other possibilities available.

In the case of two-parent households, modelling joint household labour supply and multiple programme participation decisions would be a major undertaking if we are to allow a full correlation structure between unobservables. However, lone mothers are an important client group for the welfare system, a group which has grown considerably in size in recent years. Moreover much of the US work relates to lone mothers. Thus we restrict our analysis to this group. Since IS and HB are characterised by participation close to 100% for lone parents we consider only the modelling of FC participation. Thus we focus on lone mothers and consider only participation in FC.

Policy makers are interested in the implications of changes in the provisions of transfer programmes for government revenue, the extent of eligibility, and the distribution of net incomes. Estimates of the number of programme participants to expect following some reform are frequently made from data on income distributions alone. Such estimates will be biased downwards in the case of in-work income support programmes which create positive work incentive effects since they encourage programme participation through increasing hours of work and gross earnings. The opposite holds for programmes that have adverse work incentive effects such as AFDC in the United States or IS in the United Kingdom.

The possibility of programme non-participation introduces additional complexity to any analysis of the effects of reforms. For example, estimates of likely programme participation from an increase in programme eligibility will be biased upwards if some newly eligible individuals do not participate in the programme. Thus, a comprehensive analysis of welfare programme effects requires consideration of both labour supply effects and programme participation effects.

I. THE FAMILY CREDIT PROGRAMME AND THE FES DATA

FC superseded Family Income Supplement (FIS) in 1988. In 1977/8 FIS expenditure was £63 million (at 1992 prices) and there were approximately 70,000 recipients. In 1991/92 FC expenditure had risen to £864 million (at 1992 prices) and the number of recipients to 420,000. Official UK estimates (see DSS (1995*b*)) of the extent of non-participation varies across programmes and by type of individual. The participation rates (measured as total expenditure divided by estimated total entitlements) for IS and HB are quoted as 84–93 and 91–5% respectively where lone parents are thought to be at the high end of the ranges quoted. However for FC, participation is quoted as 73%

overall. Fry and Stark (1993) find that lone mothers have lower than average FC participation rates. For 1989 they compute that the participating eligible lone mothers were 53% of the estimated eligibles, and the amount received constituted just 44% of the estimated total amount of eligibility.

The nature of the UK welfare system relevant to the labour supply of lone mothers is well documented elsewhere. In our analysis we include all welfare programme entitlements into the budget constraints that individuals are assumed to face. However, only Family Credit is described in detail here as it is the welfare programme of interest in the present application. After 1988 entitlement has been driven by the formula:

$$FC = \max\{0, MFC - \max[0, 0{\cdot}7(Y_1 - AA)]\} \quad \text{if} \quad h \geqslant 24,$$

where Y_1 is weekly after tax income averaged over the previous five weeks or two months. AA is called an 'applicable amount' and MFC is the maximum FC payable which depends on the number and ages of children.[7] In terms of the budget constraint there is a discontinuity at 24 hours of work per week up to MFC, followed by a section with slope equal to the gross wage rate w (ignoring income taxation and social security contributions) until income reaches AA which is then followed by a section with a 70% clawback rate until the entitlement is exhausted at the value of Y such that $MFC = 0{\cdot}7(Y - AA)$. There was a marked increase in the MFC in 1988 which was somewhat larger for families with two children. h is the number of hours worked per week?

Our data consist of 15 pooled cross-sections of Family Expenditure Surveys[8] from 1978 to 1992 which yield a sample of 4,248 lone mothers who are householders. We compute eligibility and the level of entitlement from a very detailed routine[9] that acknowledges all of the features of the tax, welfare and social security contribution systems (except for in-kind transfers) using estimated wage equations. A particular feature to be aware of is that HB entitlement depends upon the level of FC receipt, so that FC non-participation may not have such a dramatic effect on net income as one would otherwise imagine.

[7] Prior to 1988 the formula was $FC = \min[0, MFC - 0{\cdot}5(Y_0 - PA)]$ where Y_0 is pre-tax income and PA is the 'prescribed amount' which depends on the number and ages of children. Since the marginal tax rate from 1988 has been 25%, the FIS taper of 50% of pre-tax income closely approximates the FC 70% taper on post-tax income. Prior to the 1988 reforms FIS carried with it a passport to free school meals. From 1988 this was dropped and the amounts paid for children of all ages increased in (partial) compensation. FC has been subject to a wealth test since 1988 although this is set quite high and the effect is probably quite weak. We approximate this test by applying an assumed interest rate to the recorded investment income data. FC is payable for 6 months from establishing entitlement regardless of any change in circumstances. This feature of FC begs the question of what is the appropriate way to model its effects. A strict interpretation of the incentive mechanism is that during the assessment period over which eligibility is established the FC carries a 70% taper but, once eligibility is established, FC becomes a lump-sum transfer and should have only income effects on labour supply behaviour. We intend to pursue this in future work.

[8] Fry and Stark (1993) consider the deficiencies of the FES data in detail. It is difficult to use data prior to 1978 because of the absence of schooling data, and data beyond 1992 does not contain appropriate information about housing costs to deal with changes in the local tax system that occurred at this time.

[9] The routine is based on the Institute for Fiscal Studies' TAXBEN programme but is less complicated in that it considers only lone mothers, although more complicated in that it considers not just tax/benefits at the observed level of hours but also at other levels, and it deals with all of the changes that have taken place between 1978 and 1992. See Johnson *et al.* (1990) for details of TAXBEN.

A peculiar feature of FC is that once entitlement has been established it continues for 6 months with no obligation for recipients to report changes in circumstances. Programme participation is summarised in Table 1 where the

Table 1
Labour Force Status and Programme Participation

Labour market status	Ineligible non-participants	Ineligible participants	Eligible non-participants	Eligible participants	Total number
Not working	2,939	53	0	0	2,992
Part-time	292	23	142	125	582
Full-time	495	22	90	67	674
All	3,726	98	232	192	4,248

data is divided by labour force status. The table uses actual observed (usual) income and hours data to compute entitlement.

There are relatively few ineligible participants suggesting that we are quite successful at predicting FC entitlements despite the nature of the programme which allows continued receipt after a change in circumstances. There are two problem areas in the data. First, most of the ineligible participants are ineligible by virtue of their hours being below 24 and this is probably a feature of being able to continue to receive FC even after a change in circumstances. Secondly, there are significant numbers of eligible participants and non-participants at full-time hours but these have quite a low level of entitlement on average (£5.65 and £4.73 respectively) so that their eligibility is going to be very sensitive to their predicted wages.

II. ECONOMETRIC FRAMEWORK

Budget constraints faced by UK lone mothers are likely to take a complicated piecewise-linear form with two important non-convexities due to FC – one due to the notch at 24 hours interacting with the 100% tax faced by those on IS, and one arising from eligibility ceasing as earnings rise. Moreover, we find strong evidence that there is non-convexity in the gross budget constraint induced by a marked differential between the wages of full-time workers and part-time workers. There are severe difficulties associated with estimating a model that allows for the possibility of continuous hours substitution[10] in the face of such complex budget constraints. Our approach follows much of the literature on simultaneous labour supply and programme participation modelling and approximates the continuous choice with a choice among discrete alternatives. Each regime is characterised by some specific hours level. Individuals choose between these alternatives and no hours substitution is

[10] See MaCurdy *et al.* (1990) for an approach which is based on estimating an hours equation where the budget constraint is approximated by a quadratic function. However, the possibility of finding a suitable reasonable continuous approximation to the constraint faced by UK lone mothers is remote. A second approach can be found in Blundell *et al.* (1992) where they sidestep the problem by taking a 'deep selection' of high earning women only and correcting for the resultant bias using some reduced form. The weakness of this approach is that the absence of valid exclusion restrictions to achieve identification.

allowed within regimes. Though obviously an approximation, one could argue that choice between discrete hours offers is an empirical regularity, and modelling hours as a continuous choice may, in fact, be a mis-specification.

McFadden (1984) surveys the discrete response literature and motivates a choice of modelling framework appropriate to the present context following Hausman and Wise (1978). That is, we estimate an unordered Probit random utility model over four states (full-time work; part-time with FC participation; part-time without FC participation; and non-participation) and we control for the fact that some of the labour market non-participants would rather be working. Since these choices are determined by the income levels associated with each state, and we only observe one of these, we need to predict incomes for each state from the income in the observed state. Since it would be computationally burdensome to estimate the wages associated with each state jointly with the choices of state, and since we only require consistent wage predictions in order to estimate the determinants of each state, we adopt a two-step procedure. In the first step we estimate full-time and part-time wage equations which uses a reduced form for labour market status to control for the endogeneity of labour market status and use these estimates to predict incomes in the part-time and full-time positions. Incomes for non-participants are computed from a knowledge of the welfare system and observed unearned (non-transfer) income. In the second step, we estimate the random utility model using the predicted income in each state.

We approximate the budget constraint by just four discrete points which we think of as labour market non-participation (NP), part-time work (PT) and full-time (FT) work[11] where part-time workers may choose to participate in FC (PT(1)) or not (PT(0)). Since FC entitlement at zero hours is zero and is typically small at full-time hours, and since the participation rates for IS and HB are close to 100% for lone parents we assume that FC participation is the only welfare participation issue, and even then only at the part-time position.

The choice between the four alternatives is assumed to be driven by differences in the utility attached to them. Let the utility associated with choosing state k be $U_i(y_k, h_k; \mathbf{X}_i)$ where y_k is the income associated with this choice (including any FC received) and $\mathbf{X}_i$ is a vector of individual characteristics.

Now consider a statistical specification[12] which allows for random variation in behaviour due to an additive disturbance and variation in tastes,[13]

[11] We define these as: usual weekly hours less than 15, between 15 and 34, and 35 or greater, respectively. We then compute their incomes at zero hours, 24 hours and 40 hours. An important criticism of discrete choice modelling is the arbitrary nature of the definition of the alternatives. Sensitivity of the estimates to different definitions of what constitutes part-time and full-time was tested. The parameters were not significantly affected by the definition of part-time but the full-time criteria of 35 hours is obviously more crucial since increasing this brings the 'full-time' hours peak into the part-time definition.

[12] Fischer and Nagin (1981) denote this a random coefficients, covarying disturbances model. Their notation is used above.

[13] ϵ_{hi} could represent unobserved attributes of alternatives or individuals which affect choice but are uncorrelated with V_{hi}. This may also be a pure element of random choice. Note that, by itself, an additive disturbance term allows no variation in tastes. Individuals with identical observables may have different tastes, and this variance must be incorporated explicitly in the taste parameter. It can easily be shown that failure to incorporate taste variation induces a downward bias to estimates of taste parameters.

$U_i^*(y_k, h_k; \mathbf{X}_i, \epsilon_{ik})$, where U_{ik}^* is latent unobservable utility of state k for individual i. Thus, the utility gain from moving from alternative k to j is

$$U_i^*(y_j, h_j; \mathbf{X}_i, \epsilon_{ij}) - U_i^*(y_k, h_k; \mathbf{X}_i, \epsilon_{ik}) \equiv U_{ij} - U_{ik}. \tag{1}$$

In a discrete choice model the choices j and k are assumed to be common across i so this utility difference can be expressed as

$$U_{ij}^* - U_{ik}^* = \mathbf{g}(y_{ij}, y_{ik})(\bar{\boldsymbol{\psi}} + \tilde{\boldsymbol{\psi}}_i) + \mathbf{X}_i \boldsymbol{\beta}_{jk} + (\epsilon_{ij} - \epsilon_{jk}), \tag{2}$$

where $\mathbf{g}(y_{ij}, y_{ik})$ is a row vector of differences of functions of the net incomes, and $\bar{\boldsymbol{\psi}}$ reflects the mean tastes of the sample while $\tilde{\boldsymbol{\psi}}_i$ is a coefficient vector which shows how i differs from the mean, and $(\epsilon_{ij} - \epsilon_{ik})$ is an additive disturbance assumed to be iid across i but not necessarily across j.[14] The choice of $\mathbf{g}(\,)$ is arbitrary although here we assume linearity for simplicity.

Keeping a linear form in the $\boldsymbol{\psi}$ parameters enables us to introduce random parameters as described in Hausman and Wise (1978). In this case assuming that the error terms are jointly normally distributed and that the parameters are independently normally distributed is a natural choice and leads to tractable expressions for the probabilities. The model thus obtained belongs to the class of Multinomial Probit Random Utility Models.

As usual in this class of model only the utility differences between the number of alternatives minus one can be identified (see Pudney (1991) for a discussion). This results in the need for identifying restrictions on the variance covariance of the error terms ϵ_{ij}. Therefore we parameterise the variance covariance matrix of the terms $(\epsilon_{ij} - \epsilon_{ik})$ such that the variance of $(\epsilon_{i1} - \epsilon_{i2})$ equals one.

It is possible to interpret the parameters $\boldsymbol{\beta}_{jk}$ as a gain (or a loss) in utility from having the characteristics $\mathbf{X}_i$ when one compares the alternative j to the alternative k, where this latter choice is the reference.

From (2) the probability of observing i in state j is given by

$$P_{ij} = \Pr(U_{ij}^* > U_{ik}^*) = \Pr[\mathbf{g}(y_{ij}, y_{ik})(\bar{\boldsymbol{\psi}} + \tilde{\boldsymbol{\psi}}_i) + \mathbf{X}_i \boldsymbol{\beta}_{jk} > (\epsilon_{ij} - \epsilon_{ik})] \quad \forall j \neq k. \tag{3}$$

We define the choices k by discretising the actual hours distribution whereby we allocate an hours level of h_k to all those individuals whose hours fall within a range around h_k. On the one hand this induces an errors in variables problem but, on the other hand, it allows for a non-structural form approach to be taken. The larger the number of alternatives chosen the less the measurement error problem induced but the greater the computational complexity.

We assume that an individual may chose among four alternatives – three labour market states, one with or without programme participation. However, those ineligible for the programme face only three alternatives, since programme participation is not an option. This restricted choice set is accounted for in estimation (see likelihood appendix for details).

The FES asks those with zero hours in the labour market whether they are actively looking for a job and we use this information to discriminate between

[14] Hausman and Wise (1978) assume ϵ_{hi} to be iid across alternatives too.

voluntary non-participation and involuntary unemployment. This is important because individuals who are involuntary unemployed are not observed to be in their most preferred state, and must be classified appropriately in a choice model. This group is assumed to reveal that some positive hours state is preferred to zero and they are distinguished by the following latent and observed unemployment rationing equations

$$R_i = \mathbf{1}[(R_i^* = \mathbf{Z}_i \boldsymbol{\tau} + v_i) > 0], \tag{4}$$

where R_i^* is the latent variable corresponding to the observed rationing, R_i, which we define to be unity if i is observed to be not working and searching for work and zero otherwise, $\mathbf{Z}_i$ is a matrix of demand side variables; $\boldsymbol{\tau}$ is a corresponding vector of parameters; and v_i is a random error.

Individuals observed in any positive hours labour market state are assumed to prefer their observed state to all alternatives and not to be rationed in exercising this preference.

While this is an extension that has not previously been considered in the labour supply and programme participation literature, we consider it important because we would otherwise understate the extent of programme non-participation. That is, we would assume that non-workers prefer that state to part-time work with programme participation when in fact they simply cannot find a job.[15]

Details concerning stochastic specification, identification and likelihood contributions are relegated to the Appendix.

III. ESTIMATES

The results are given in Table 2.[16] The direct interpretation of comparison-specific parameters in an unordered model is problematic. A positive sign implies that a variable is associated with increasing the probability of moving to the destination state. $PT(0)$ is part-time with FC non-participation, $PT(1)$ is part-time work with FC participation. For example, a positive coefficient on *Widow* in the $PT(0) \rightarrow NP$ equation means that being a widow makes one more likely to prefer NP than $PT(0)$. They are presented here to give an idea of sign, magnitude and significance.

The interpretation of the parameters is less than transparent since they tell us about the impact of characteristics on the probability of choosing one state rather than the default (of voluntary non-participation). Thus the constant terms tell us what the increase in utility would be of a move to voluntary non-participation for the default individual, while the coefficients on the

[15] The specification for the determination of wages is $\log W_i^h = \mathbf{Z}_i \gamma^t + e_i^h$, for $h = PT, FT$ where PT and FT indicate part-time and full-time labour force status. We are simply interested in obtaining consistent wage predictions so, although employment status is likely to be endogenous to the wage, we estimate the wage equations by including the Mills Ratio from a bivariate Probit model of participation *vs.* non-participation and full-time *vs.* part-time work conditional on participation. We include the level of unearned income in the reduced form labour force status equations but not in the wage equations to achieve identification. Details are available on request.

[16] To keep the dimensionality down we appeal to the results in Blundell *et al.* (1987) which tests for the exogeneity of rationing in their double hurdle model of labour supply and finds that the data support this.

Table 2

Labour Supply, FC Participation and Rationing Estimates

	Choice			
Variable	$PT(0) \rightarrow NP$	$PT(1) \rightarrow NP$	$FT \rightarrow NP$	Ration
Constant	0·4576 (0·2175)	0·7322 (0·0621)	0·0411 (0·3447)	−0·8179 (0·3002)
Renter	0·1965 (0·0348)	0·0524 (0·0847)	0·5320 (0·0929)	0·2891 (0·0466)
Age	−0·5239 (0·9851)	1·1299 (2·2829)	0·1514 (1·6461)	−2·5613 (1·4839)
Age^2	1·3517 (1·2582)	0·5377 (2·8317)	2·3044 (2·1740)	0·7686 (1·9806)
Child 0–4	0·9344 (0·0717)	1·5570 (0·2225)	1·8654 (0·2960)	0·4272 (0·0572)
Child 5–10	0·2961 (0·0482)	0·4958 (0·1076)	1·0276 (0·1645)	0·1990 (0·0377)
Widow	0·3331 (0·0420)	1·2195 (0·2491)	0·3753 (0·0892)	0·0252 (0·0638)
Unemployment				0·2234 (0·1156)
North	0·2403 (0·0627)	−0·1380 (0·1332)	0·6811 (0·1368)	−0·0158 (0·1026)
Yorkshire	0·1710 (0·0437)	0·0896 (0·0954)	0·4084 (0·0915)	−0·0486 (0·0664)
N. West	0·2541 (0·0457)	0·4100 (0·1109)	0·4702 (0·1005)	−0·1814 (0·0806)
E. Midlands	−0·0227 (0·0431)	−0·3838 (0·1361)	0·5086 (0·1053)	−0·0010 (0·0569)
W. Midlands	0·0983 (0·0462)	0·0139 (0·1100)	0·4850 (0·1074)	0·1527 (0·0733)
E. Anglia	−0·1615 (0·0489)	−0·5140 (0·1472)	0·4329 (0·1088)	−0·0391 (0·0696)
S. West	−0·0846 (0·0330)	−0·0177 (0·0798)	0·3290 (0·0747)	−0·2090 (0·0651)
Wales, Scot, NI	0·1894 (0·0412)	−0·0740 (0·1007)	0·2338 (0·0703)	−0·2178 (0·0821)
1979–80	0·0724 (0·0480)	0·0343 (0·1401)	−0·4267 (0·1064)	−0·1392 (0·0852)
1981–82	0·0018 (0·0481)	−0·4675 (0·1617)	−0·2329 (0·0943)	0·2217 (0·0941)
1983–84	0·1236 (0·0491)	−0·4639 (0·1734)	0·1552 (0·0866)	0·2073 (0·1105)
1985–86	0·0654 (0·0478)	−0·3951 (0·1572)	0·0507 (0·0867)	0·2970 (0·1150)
1987–88	0·2051 (0·0525)	−0·2513 (0·1510)	−0·1089 (0·0961)	0·4738 (0·0932)
1989–90	0·2687 (0·0547)	−0·3065 (0·1551)	−0·1218 (0·0993)	0·1631 (0·0845)
1991–92	0·3283 (0·0615)	−0·1687 (0·1424)	0·0679 (0·1016)	0·1022 (0·0884)
$\bar{Y}$	— —	7·6871 (0·9436)	— —	— —
$\tilde{Y}$	— —	1·8117 (0·2069)	— —	— —
σ	3·4274 (1·0608)	1	3·5166 (1·0658)	1
$\rho_{PT(0) \rightarrow NP, PT(1) \rightarrow NP}$	—	−0·4781 (0·2597)	—	—
$\rho_{PT(0) \rightarrow NP, FT \rightarrow NP}$	—	0·7416 (0·2954)	—	—
$\rho_{PT(1) \rightarrow NP, FT \rightarrow NP}$	—	−0·4469 (0·5682)	—	—
Observations	—	4,248	—	—
Mean log L	—	−0·95853	—	—

characteristics tell us how different individuals depart from this default.[17] Thus, denoting utility as $U(y, h, PP)$ where y is income, h is hours of work, and PP is an indicator of FC programme participation, it is interesting to derive the utility difference associated with FC programme participation as

$$U(y, PT, 1) - U(y, PT, 0) = U(y, PT, 1) - U(y, NP, 0) - [U(y, PT, 0) - U(y, NP, 0)],$$

which for the default individual is 0·2746 (= 0·7322 − 0·4576). In order to put some perspective on this figure we need to compare it with the utility gain associated with an additional pound on income (the coefficient on Y in Table 2 divided by 100) which is estimated to be 0·0769. Thus, we can infer that the utility loss associated with participating in FC is, for given income, the same as the utility loss associated with a reduction in income of £3.57 (i.e.

[17] More than the usual degree of heroism is required to make welfare inferences from the estimates. In particular, in addition to cardinal comparability we maintain that utility is linear in income.

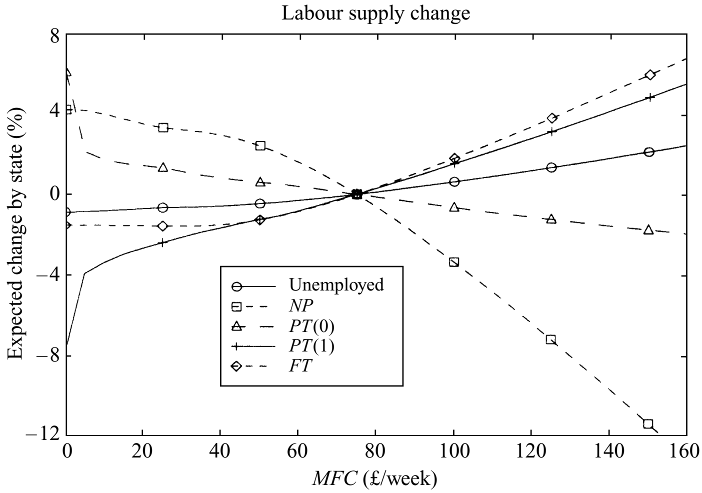

Fig. 1. Simulated labour market and programme participation effects of variations in *MFC*.

0·2746/0·0769). We can compute this figure for all individuals in the data and we find that the average utility loss from FC participation is £5·91 (S.D. = 1·47).

We can also use the estimates to infer the utility loss associated with working. That is, we can compute the loss associated with part-time work as $U(y, PT, 0) - U(y, NP, 0)$ which, again by comparing it with the utility gain associated with an additional pound of income, implies that part-time work reduces utility by an average of £16·34 (S.D. = 2·77). Similarly, full-time work reduces utility by £26·93 (S.D. = 4·86), and the average utility loss from rationing is £10·67 (S.D. = 3·50). Note that although the standard deviation around these estimated means are high the distributions are highly skewed and there are, in fact, no instances where the value falls to zero.

Thus the estimates suggest that the economic framework appears to be broadly supported by the data in that there is a utility loss associated with FC programme participation, there is a gain associated with additional income, and there is a loss associated with additional work.

Moreover, significant correlations between unobservables in the choice equations suggests that a multinomial probit would dominate a multinomial logit. A significant random parameter on income supports the random utility model approach which allows for taste heterogeneity.

To get a broad feel for the effect of the constraint on behaviour we simulated the effect of increasing the incomes of part-timers by £10. The experiment that increases income at $PT(0)$ and not $PT(1)$ is particularly successful at reducing non-participation. The experiment that increases FC by £10 increases $PT(1)$ by 1·3% mainly at the expense of $PT(0)$ and *NP*, but with essentially no adverse impact at all on the full-time probability. Although the traditional take-up rate measure remains largely unchanged at 51·5%, because the rise in the probability of being in $PT(1)$ is matched by the fall in $PT(0)$, it would be difficult to argue this was a sensible measure of the obvious improvement in programme participation. Further simulations suggests that labour supply is somewhat less elastic at high hours than at low hours which is consistent with the results in Blundell *et al.* (1992) and with the view that work incentive effects may be predominantly participation effects (see Heckman (1993)).

Our results suggest that FC possesses the two most desirable features of an in-work transfer programme – it strongly encourages individuals to work, and it has little adverse effect on the incentives of those already in work. Thus it is interesting to see if the extent of FC support is set at the most appropriate level. Thus we simulate the impact of changes in the level of the MFC (just for this sample of lone mothers) – when the MFC is set to zero all FC entitlements disappear, and as FC increases more and more individuals become entitled and the levels of entitlement rise. In Fig. 1 we show the effect of variations in the MFC on the proportion of individuals in each category.

Doubling the MFC effectively: reduces voluntary non-participation by about 10% points from a base of 66·1%, increases the involuntary unemployed (because more individuals wish to work the higher is FC), reduces PT workers who do not participate in FC ($PT(0)$) because participation is more attractive the greater is the level of entitlement, increases $PT(1)$ for the same reason, and FT also rises (even more so that $PT(1)$) because of greater income at FT hours (as well as at PT hours).[18] Eliminating FC by setting *MFC* to zero has the opposite effects although not symmetrically so.

One might hope that, by reducing expenditure on IS, increasing the *MFC* would be relatively inexpensive. Fig. 2 shows that FC expenditure rises (falls)

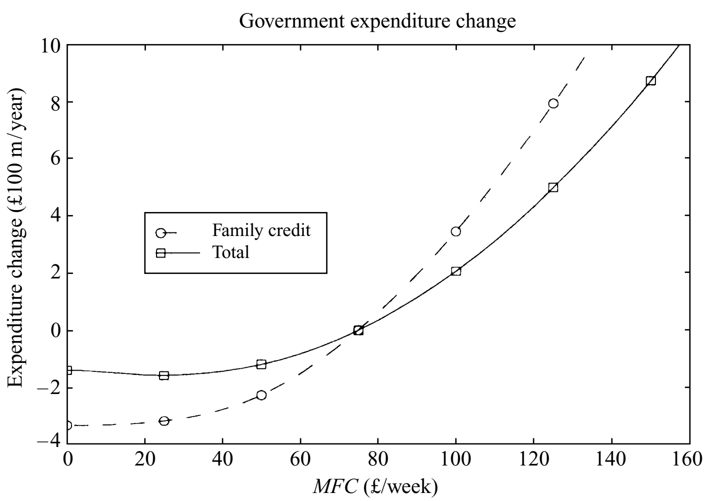

Fig. 2. Simulated effects of variations in *MFC* on government expenditure.

steeply as *MFC* rises from its existing level and that this offsets the additional cost by around 40% through rising income tax and NI and falling IS. Note

[18] In simulations we assume FC participation at full time. This is consistent with the estimation since full-time FC entitlements are found to be relatively small in the data.

that abolishing FC altogether reduces FC expenditure to zero from its current (1992 on lone parents) cost of approximately £350m. However, the impact of reducing the *MFC* on total government expenditure (tax and NI less benefits payments) leads to a reduction in total expenditure as *MFC* falls to around £35 per week but that government expenditure rises as *MFC* falls below this point. The reason for this is that a small reduction in *MFC* causes a large rise in *NP* and hence additional IS expenditure, while a larger drop causes FT work to begin to rise and FC participation falls steeply as entitlements get very small.

IV. CONCLUSION

This paper has been concerned with estimating a model of labour supply and participation in Family Credit (as in-work transfer programme which is known to suffer from poor participation rates) for lone mothers in the United Kingdom. A feature of the modelling is the treatment of labour force non-participants where we discriminate between voluntary non-participants and the involuntary unemployed. This is important in this context, since to treat all non-workers as being on their labour supply curves would exaggerate the impact of the stigma associated with Family Credit. That is, we would be attributing some labour force non-participation to ignorance and/or stigma associated with FC rather than to an inability to obtain a job.

The weakness of the analysis is that it cannot tell us why programme non-participation occurs. Our results simply suggest that it does occur and we can infer from them the loss in income that is equivalent to FC participation *per se*. However, we find that, on average, this 'stigma' is in the order of £6 per week (compared with average receipts of £25).

The results are extremely important for public policy. There are two important findings. First, we show that an increase in FC has a large impact on the probability of taking up part-time work and some impact on wanting (but not being able) to participate but essentially no adverse effect on the probability of working full-time. Thus, FC seems to contribute to overcoming the unemployment trap without inducing a serious poverty trap problem for those already in work. Secondly, however, we find evidence of some considerable 'stigma' which implies that FC is not as effective at countering the disincentive effect of the Income Support programme or at countering poverty amongst the working poor as it might otherwise. If it were possible to simply eliminate the stigma associated with FC this would have an important impact on the labour force non-participation rate for lone mothers and would imply large savings in government expenditure on Income Support for non-working lone mothers.

The first result is encouraging for recent proposals to widen the scope of FC. Moreover, the implications of extending FC coverage to a wider clientele, such as non-parents, may imply that its potential reputational externality might be diminished. Thus ET may not suffer from such stigma and it may even reduce the stigma currently associated with FC itself.

Our simulations which vary the extent of FC provision reinforce the assertions above. Indeed, we find that extending FC provision higher up the income distribution makes both part-time and full-time participation rise.

Centre for Labour Markets & Social Research, Århus Keele University

Date of receipt of final typescript: October 1996

References

Besley, Timothy and Coate, Stephen (1992*a*). 'Understanding welfare stigma: tax-payer resentment and statistical discrimination.' *Journal of Public Economics*, vol. 48, pp. 165–83.

Besley, Timothy and Coate, Stephen (1992*b*). 'Workfare versus welfare: incentive arguments for work requirements in poverty-alleviation programs.' *American Economic Review*, vol. 82, pp. 249–61.

Blundell, Richard, Duncan, Alan and Meghir, Costas (1992). 'Taxation in empirical labour supply models: lone mothers in the UK.' Economic Journal, vol. 102, pp. 265–78.

Blundell, Richard, Fry, Vanessa and Walker, Ian (1988). 'Modelling the take-up of means-tested benefits: the case of housing benefits in the United Kingdom.' Economic Journal, vol. 98, pp. 58–74.

Blundell, Richard, Ham, John and Meghir, Costas (1987). 'Unemployment and female labour supply.' Economic Journal, vol. 97, pp. 44–64.

Department of Social Security (1995*a*). *Piloting Change in Social Security: Helping People into Work*. London: HMSO.

Department of Social Security (1995*b*). *Social Security Statistics 1995*. Government Statistical Service. London: HMSO.

Ermisch, John F. and Wright, Robert E. (1993). 'Wage offers and full-time and part-time employment by British women.' *Journal of Human Resources*, vol. 28, pp. 111–33.

Fischer, Gregory W. and Nagin, Daniel (1981). 'Random versus fixed coefficient quantal choice models.' In *Structural Analysis of Discrete Data*. (ed. Charles Manski and Daniel McFadden), pp. 273–304. Cambridge, MA: MIT Press.

Fraker, Thomas and Moffitt, Robert (1988). 'The effect of food stamps on labor supply: a bivariate selection model.' *Journal of Public Economics*, vol. 53, pp. 25–56.

Fry, Vanessa and Stark, Graham (1993). *The Take-Up of Means-Tested Benefits 1984–90*. London: Institute for Fiscal Studies.

Hausman, Jerry and Wise, David A. (1978). 'A conditional probit model for qualitative choice: discrete decisions recognizing interdependence and heterogeneous preferences.' *Econometrica*, vol. 46, pp. 403–26.

Heckman, James (1993). 'What has been learned about labor supply in the last 20 years?' *American Economic Review*, vol. 82, no. 2, pp. 116–21.

Hoynes, Hilary Williamson (1996). 'Welfare transfers in two-parent families: the case of AFDC-UP.' *Econometrica*, vol. 64, no. 2, pp. 295–332.

Johnson, Paul, Stark, Graham and Webb, Steven. (1990). 'TAXBEN2: the new IFS tax-benefit model.' IFS Working Paper 90/5.

Keane, Michael and Moffitt, Robert (1991). 'A structural model of multiple welfare program participation and labor supply.' (mimeo).

Kiefer, Nicholas M., Natarajan, Ranjini and McCulloch, Charles E. (1995). 'Maximum likelihood for the multinomial probit model.' CLS Working Paper, Århus.

MaCurdy, Thomas, Green, D. and Paarsch, H. (1990). 'Assessing empirical approaches for analysing taxes and labour supply.' *Journal of Human Resources*, vol. 25, pp. 415–90.

McFadden, Daniel (1984). 'Econometric analysis of qualitative response models.' In *Handbook of Econometrics*, vol. 2 (ed. Zvi Griliches and Michael D. Intriligator). New York: Elsevier, pp. 1396–457.

Moffitt, Robert (1983). 'An economic model of welfare stigma.' *American Economic Review*, vol. 73, pp. 1023–35.

Moffitt, Robert (1992). 'The incentive effects of the US welfare system: a survey.' *Journal of Economic Literature*, vol. 92, pp. 1–61.

Pudney, Stephen (1991). *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*. London: Basil Blackwell.

Scholz, John K. (1994). 'Earned income tax credit: participation, compliance and anti-poverty effectiveness. *National Tax Journal*, vol. 47, pp. 63–87.

Appendix

The Likelihood

The general case of four alternatives which is discussed in the text is only applicable for those lone mothers (79%) who have positive predicted entitlement to FC at part-time work. Due to high predicted wages or high unearned income for the remainder of the individuals the $PT(0)$ and $PT(1)$ regimes collapse. Since $PT(1)$ corresponds to participating in FC, this is no longer an option, and $PT(1)$ becomes degenerate for this group – who are now considered to face only three alternatives.

In what follows, for the purposes of exposition only, we ignore the random parameterisation and the existence of state-specific characteristics. Essentially this amounts to considering simpler variance-covariance matrices[19] and omitting income functions and associated mean parameters from the likelihood.[20]

Stochastic assumptions for the error difference terms are captured by the following variance-covariance matrices of a multivariate normal distribution:

$$V\begin{bmatrix} \epsilon_1-\epsilon_2 \\ \epsilon_1-\epsilon_3 \\ \epsilon_1-\epsilon_4 \\ v \end{bmatrix} = \begin{bmatrix} \alpha & \beta & \gamma & 0 \\ \beta & \delta & \psi & 0 \\ \gamma & \psi & \pi & 0 \\ 0 & 0 & 0 & \kappa \end{bmatrix} = \begin{bmatrix} 1 & \rho_{23}\sigma_{13} & \rho_{24}\sigma_{14} & 0 \\ \rho_{23}\sigma_{13} & \sigma_{13}^2 & \rho_{34}\sigma_{13}\sigma_{14} & 0 \\ \rho_{24}\sigma_{14} & \rho_{34}\sigma_{13}\sigma_{14} & \sigma_{14}^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The variance of the first error difference α has been normalised to one, so δ and π may be estimated up to scale. Also the variance of the unemployment ration error κ has been normalised to one. Essentially we have a switching regression with an exogeneous switch. According to rationing the switch is between being observed in most preferred state and in the least preferred. The switch is exogenous because unemployment ration is assumed independent of choice.[21] This requires only three dimensional integration, which is numerically tractable. The model is parametrically identified, but exclusion restrictions are required for non-parametric identification. These arise naturally in the present context, if we are allowed a supply and demand interpretation, where income is assumed to affect only choice and aggregate unemployment only affects rationing.

The first likelihood presented is that of a tetranomial probit for the full choice set (for the potentially FC eligibles) with full correlation structure but without random coefficients.

$$L_4 = \begin{cases} \prod_{i\in 1} \Phi_3\left(\frac{X\omega_{12}}{\sqrt{\alpha}}, \frac{X\omega_{13}}{\sqrt{\delta}}, \frac{X\omega_{14}}{\sqrt{\pi}}, \beta, \gamma, \psi\right) \\ \prod_{i\in 2} \Phi_3\left(\frac{-X\omega_{12}}{\sqrt{\alpha}}, \frac{X(\omega_{13}-\omega_{12})}{\sqrt{(\alpha+\delta-2\beta)}}, \frac{X(\omega_{14}-\omega_{12})}{\sqrt{(\alpha+\pi-2\gamma)}}, \alpha-\beta, \alpha-\gamma, \alpha+\psi-\gamma-\beta\right)\Phi_1(Z\tau) \\ \prod_{i\in 3} \Phi_3\left(\frac{X(\omega_{12}-\omega_{13})}{\sqrt{(\alpha+\delta-2\beta)}}, \frac{-X\omega_{13}}{\sqrt{\delta}}, \frac{X(\omega_{14}-\omega_{13})}{\sqrt{(\delta+\pi-2\psi)}}, \delta-\beta, \gamma+\delta-\beta-\psi, \delta-\psi\right)\Phi_1(Z\tau) \\ \prod_{i\in 4} \Phi_3\left(\frac{X(\omega_{12}-\omega_{14})}{\sqrt{(\alpha+\pi-2\gamma)}}, \frac{X(\omega_{13}-\omega_{14})}{\sqrt{(\delta+\pi-2\psi)}}, \frac{-X\omega_{14}}{\sqrt{\pi}}, \beta+\pi-\gamma-\psi, \pi-\gamma, \pi-\psi\right)\Phi_1(Z\tau) \\ \prod_{i\in 5} \Phi_3\left(\frac{-X\omega_{12}}{\sqrt{\alpha}}, \frac{-X\omega_{13}}{\sqrt{\delta}}, \frac{-X\omega_{14}}{\sqrt{\pi}}, \beta, \gamma, \psi\right)\Phi_1(-Z\tau), \end{cases}$$

where states 1 to 4 are respectively unrationed (i.e. voluntary) non-participation in the labour market, part-time and FC non-participant, full-time, and part-time and FC

[19] Without additive terms for individual income covariance.

[20] The simplest linear functions of income specification is presented. Quadratic and log income failed to improve the likelihood and cubic income failed to converge.

[21] Blundell *et al.* (1987) estimate a double hurdle model (using the same searching question) and find a significant correlation between the error terms. However they go on to show that this makes little difference to the parameter estimates.

participant. State 5 is rationed unemployed. Φ_1 and Φ_3 are univariate and trivariate normal cumulative distribution functions.

Now consider the trinomial Probit for the reduced (to three alternatives) choice set. The parameter values are restricted to be the same, correcting for a degenerate likelihood contribution. Hence the likelihood contributions are as follows:

$$L_3 = \begin{cases} \prod_{i \in 1} \Phi_2\left(\dfrac{X\omega_{12}}{\sqrt{\alpha}}, \dfrac{X\omega_{13}}{\sqrt{\delta}}, \beta\right) \\ \prod_{i \in 2} \Phi_2\left(\dfrac{-X\omega_{12}}{\sqrt{\alpha}}, \dfrac{X(\omega_{13}-\omega_{12})}{\sqrt{(\alpha+\delta-2\beta)}}, \alpha-\beta\right)\Phi_1(Z\tau) \\ \prod_{i \in 3} \Phi_2\left(\dfrac{X(\omega_{12}-\omega_{13})}{\sqrt{(\alpha+\delta-2\beta)}}, \dfrac{-X\omega_{13}}{\sqrt{\delta}}, \delta-\beta\right)\Phi_1(Z\tau) \\ \prod_{i \in 5} \Phi_2\left(\dfrac{-X\omega_{12}}{\sqrt{\alpha}}, \dfrac{-X\omega_{13}}{\sqrt{\delta}}, \beta\right)\Phi_1(-Z\tau), \end{cases}$$

where Φ_2 is a bivariate normal cumulative distribution function.