Review of Economic Studies (1992) 59, 537-557 0034-6527/92/00270537$02.00

Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets

MANUEL ARELLANO
London School of Economics and CEMFI, Madrid

and

COSTAS MEGHIR
University College London and Institute for Fiscal Studies

First version received March 1989; *final version accepted August* 1991 (*Eds.*)

We develop an empirical model of labour supply that is consistent with on-the-job search and which is identified and estimated by combining two data sets: the U.K. Family Expenditure Survey which contains information on income and expenditure and the U.K. Labour Force Survey, which has data on hours and job search behaviour. We provide statistical evidence on the compatibility of the two samples for the purposes of estimating our model. We find that search has a direct negative effect on hours of work and we establish a strong positive effect of wages on hours.

1. INTRODUCTION

The purpose of this paper is to develop an empirical model of female labour supply which is consistent with intertemporal optimization under uncertainty in the presence of job search activity and to estimate it using two complementary household-level data sets.

The empirical models of labour supply developed in the literature, whether intertemporal (e.g. Heckman and MaCurdy (1980), Blundell and Walker (1986) and Altonji (1986)), or not (e.g. Heckman (1974), Hausman (1980) Cogan (1981) amongst others) ignore the dynamics implied by search theory.[1] More recently, Blundell, Ham and Meghir (1988) have presented a labour supply model which allows for the presence of job search among the unemployed. Yet an empirically tractable model of labour supply in the presence of on-the-job search has not been fully developed before.

The way one specifies a labour supply model in the presence of on-the-job search depends on the way job offers are made and accepted. As an initial description one could make two hypotheses. First, offers arrive as fixed wage–hours packages (see for example Altonji and Paxson (1987)). The individual accepts or rejects an offer and if accepted may continue to search on-the-job. In this case observed hours are not desired hours in the traditional sense. As an alternative to the above, one could assume, as we do, that offers are indexed by the wage only which is probably a good description of the operation

1. For an evaluation of static labour supply models see Mroz (1987).

of the U.K. female labour market.[2] The implications of such a model as characterized by Burdett and Mortensen (1978) is that all working individuals (whether seeking or not) are on their labour supply curve. The latter will differ in a specific way between seekers and non-seekers and we discuss this in detail in the paper.

We estimate the model by using two independent sources: the U.K. Labour Force Survey (LFS) and the U.K. Family Expenditure Survey (FES), both for 1983. Both data sets contain a comparable set of conditioning variables but while the LFS contains search variables, it lacks information on earnings and other financial information. The situation is reversed in FES with the result that our model would not be estimable using either survey separately. We are thus led to discuss issues in identifying and estimating structural parameters from complementary data sources.

Ideas relating to combining data sets have been in existence for some time. For example, Hartley (1958) and Hartley and Hocking (1971) studied maximum likelihood estimation from incomplete data. Maddala (1971) considered estimates that are obtained by pooling time series and cross-section observations. Jorgenson, Lau and Stoker (1982) combined time series with cross-section information in the estimation of a demand system. Bound, Griliches and Hall (1986) analysed the problem of pooling estimates across covariance matrices of different size. In addition, a by-product of our approach is similar in spirit to the methodology proposed by Ham and Hsiao (1984).

The combination of the two data sets enables us to estimate a female labour supply model on the LFS for the first time. The latter is a much bigger sample than the FES. We estimate a strong and positive wage elasticity, in contrast to previous U.K. studies, which is robust to a number of alternative over-identifying assumptions. The other substantive finding in this paper is that job seekers work significantly fewer hours than non-seekers *given* wages, other income and demographic characteristics.

The paper proceeds as follows. In Section 2 we present the theoretical model of labour supply in the presence of on-the-job search. Section 3 contains the empirical specification and the econometric methodology, while a general discussion of the problem of estimating models from incomplete samples, is contained in Appendix B. In Section 4 we discuss the empirical results. Finally in Section 5 we offer some concluding remarks. Other technical results are relegated to appendices.

2. MODELLING LABOUR SUPPLY IN THE PRESENCE OF JOB SEARCH

2.1. *Labour supply and on-the-job search*

In this section we present a model of labour supply including on-the-job search which is used as the basis for our empirical specification in Section 3. The theoretical framework is a generalization of Burdett and Mortensen (1978) which explicitly takes into account savings behaviour, but the emphasis here is on deriving an empirically tractable model. The resulting labour supply model, which is estimable on cross-section data is consistent with intertemporal optimization under uncertainty and with the presence of on-the-job search. The maintained assumptions are that job offers are indexed by the hourly wage rate only and are drawn from some non-degenerate wage-offer distribution.

2. The evidence on this issue is circumstantial. In the data appendix we present a distribution of hours which shows that there is relatively little clustering of hours at particular points. This can be contrasted to data from other countries (see for example Bourguignon and Magnac (1990)). In addition Blundell, Duncan and Meghir (1991) carry out a specification test where, under the alternative, the hours information for individuals working between 19-21 hours and 38-40 hours is not used. Instead all that is assumed is that these individuals have positive labour supply. The test does not reject the null of hours flexibility.

Given the offer, the individual can choose hours of work. Moreover preferences are intertemporally additively separable. Finally we assume perfect capital markets, that is, the individual can borrow and save at a given rate of interest any amount she wishes.

In our model search takes place by devoting time to this activity. The more time devoted the larger the probability of an offer. There is some controversy as to whether people classifying themselves as job seekers actually have a greater probability of a job offer, which is our maintained hypothesis here. In the context of unemployed job seekers Flinn and Heckman (1982) present some evidence in favour of this hypothesis. Moreover, the fact that transitions may also take place without search does not affect the specification of our empirical model. Finally we assume that search does not yield utility as an individual activity. This assumption has some empirical content and we discuss it below.

Given the above, the optimization problem for an individual at period t can be expressed by the following

$$\max_{C_t, l_t, S_t, A_{t+1}} \{U_t(C_t, l_t) + E_t[\beta_t V_{t+1}(A_{t+1}) \mid S_t]\} \tag{2.1}$$

where C_t is consumption l_t is total leisure time, S_t is time spent searching, β_t is the personal discount rate and A_{t+1} are end of period t assets which evolve according to the standard difference equation

$$A_{t+1} = (1+r_t)(A_t - C_t + w_t(T - l_t - S_t)). \tag{2.1a}$$

where w_t is the real wage, r_t is the real rate of interest and T is time endowment. $V_{t+1}(\cdot)$ denotes intertemporal utility at the beginning of period $(t+1)$. Moreover, by conditioning on S_t in (2.1) we emphasise that different opportunities may arise as a result of search activity. For example, in this model a job seeker will face *ceteris paribus* a higher wage profile. The expectations operator E_t is taken with respect to the distribution of future prices, interest rates, wages and possible exogenous layoffs, conditional on information available at time period t.

The first useful conclusion drawn from solving (2.1) is that for workers $(l_t + S_t < T)$ the marginal rate of substitution between consumption and leisure within each period depends only on current income and the real wage rate and not on search time.[3] That is,

$$\frac{\partial U_t/\partial l_t}{\partial U_t/\partial C_t} = w_t \qquad (j = 1, \dots, n). \tag{2.2}$$

This would not have been the case if search yielded utility unless it was separable. Note that (2.2) continues to be valid in the presence of liquidity constraints that are not directly linked to hours of work (or consumption). For example (2.2) is still valid if we impose $A_{t+1} \geqq 0$; such constraints affect the Euler equation but not the within-period marginal rate of substitution.[4]

We now obtain a labour supply schedule corresponding to (2.2). Define the within-period budget identity

$$w_t l_t + C_t \equiv w_t(T - S_t) + \mu_t \tag{2.3}$$

3. The argument made here extends to the more general case where we consider many goods rather than a single composite commodity, consumption. Clearly the formulation used here, where we just look at leisure as a function of consumption and the real wage, implies certain preference restrictions.

4. As a referee pointed out, in this context liquidity constraints could be related to the present value of earnings. The individual could alleviate them by searching for a better wage offer. This would *not* invalidate (2.2). But if current hours of work can be used to alleviate current or future liquidity constraints then (2.2) needs to be generalised. Weber (1990) presents evidence that for the period covered by our data in the U.K. earnings-related liquidity constraints are not binding in general.

where μ_t is a measure of other income which reflects net dissaving at the end of period t. Hence solving (2.2) subject to (2.3) we obtain a demand equation for leisure as a function of w_t and other income net of search costs i.e. $\mu_t - w_t S_t$. Using the total time constraint $T = l_t + h_t + S_t$, h_t being hours worked, we see that $\mu_t = C_t - w_t h_t$ and hence is observed in the data, if consumption and earnings are measured.

Using the total time constraint again, the hours of work equation is defined by

$$h_t = T - l_t(w_t, \mu_t - w_t S_t) - S_t \qquad (S_t \geqq 0) \tag{2.4}$$

where w_t and μ_t are now defined in real terms. Thus search activity will have a direct negative effect on hours of work and, to the extent that leisure is a normal good, search time interacted with the wage rate (which can be interpreted as lost income) will have a positive effect on hours of work. Hence the labour supply function of job seekers will have a higher wage derivative and a lower intercept conditional on search time S_t. The implications of (2.4) are that given search time S_t and other income μ_t, both defined by the solution to the optimization problem (2.1), current labour supply depends only on the current wage. The fact that on-the-job search creates an additional link between periods over and above (2.1a) has been fully accounted for under the assumption that workers can adjust their hours of work at the start and during the employment spell.

While (2.4) can be interpreted as a labour supply model accounting for search activity, it is consistent with other interpretations. Suppose, for example, that time spent searching for a new job was just another use of non-market time ("leisure") yielding utility. If this activity was weakly separable in the period utility function the labour supply function would take the form (2.4) (see for example, Pollak (1971) and Browning and Meghir (1991)). This should be contrasted to the case where total time spent in non-market activities only matters in the utility function, an assumption which implicitly characterizes most labour supply models. In the general case where search activity yields utility *per se* and is not separable, search time S_t would enter (2.4) in a general way affecting possibly all income and wage effects.

If search time does not yield utility, or is weakly separable from other non-market activities it will enter the labour supply function as specified in (2.4). This combined hypothesis is testable. Job search activity will have implications on wage growth since the basic underlying hypothesis is that people searching on the job are doing so in order to improve their wage. This hypothesis is not testable on our data. Thus we can specify and estimate a labour supply model that accounts for the observed job search activity, but we cannot test the overall validity of the search model through its implications on wage growth. On the other hand, we consider it an advantage that the model is consistent with a search theory interpretation (which we maintain), while at the same time being robust to a variety of other interpretations.

2.2. *The choice to search and the choice to participate*

The optimization model in (2.1) can be solved in principle for search time S_t. The first-order condition will be a function of future expectations and, in general, no set of observable statistics will be sufficient to control for the entire information set. We have thus decided to use a reduced-form approach in estimating an equation for job search. The strategy we follow derives from Blundell, Ham and Meghir (1988) and is the following.

First we define a reduced-form participation indicator function. This is positive for all workers and non-working seekers. It is negative for all non-workers who are not seeking work. This equation reflects the decision to participate in the labour force and is a function of variables that account for preferences, for fixed costs and for search costs. Defining the probability of participation as P and an employment probability index P^E, then the probability that a person is employed is PP^E, unemployed but seeking work is $P(1-P^E)$ and a non-participant $1-P$. We finally define a probability that an individual is searching on-the-job, P^S. Again this will be a reduced-form function with no *a priori* restrictions on it. Its precise form, as well as the form for P will be empirically determined. Thus our model is structural as far as the labour supply equation is concerned but reduced-form with respect to the decision to participate and to search for a (new) job. Overall the specification is consistent with intertemporal optimization under uncertainty as well as with job search activity and fixed costs. Details on the identification and estimation of the model are discussed in Section 3.

3. DATA, EMPIRICAL SPECIFICATION AND ECONOMETRIC METHODOLOGY

The data used in this study has been drawn from two independent sources: the U.K. Labour Force Survey (LFS) and the U.K. Family Expenditure Survey (FES) both for 1983. The FES is a continuing survey covering the whole year and contains approximately 7000 observations, on households. The LFS takes place over one month only and covers 70,000 households. The two surveys are complementary in that they both contain detailed demographic characteristics, education, skill etc. but only the LFS contains information on on-the-job search. The LFS contains no income or consumption information. Although both surveys contain hours information we use only the hours information contained in LFS where both hours and search behaviour are jointly observed. This allows us to test for the exogeneity of search. Thus wage and "other income" information will be obtained from FES while hours and search activity information come from LFS. In this way we exploit for the first time the rich labour market data available in the LFS for the estimation of a structural labour supply model consistent with intertemporal optimization regarding participation and on-the-job search decisions, under uncertainty. The sub-sample we draw on relates to married women of working age. A brief description of the variables used from both surveys is provided in the Data Appendix.

The LFS data for workers contains information on whether the individual is searching for alternative employment. We assume that individuals who say they are searching actually spend time in this activity. We do not observe the time spent searching but, given our assumption, this can be inferred (up to scale) by using the binary observations as fitted values from a discrete-choice model. The coefficient on search activity in the labour supply function will reflect the underlying variance of search time. Let S_i^* be a latent variable denoting desired search hours and described by the equation

$$S_i^* = b'x_i + u_{si} \tag{3.1}$$

and let h_i^* and h_i denote desired and observed hours, respectively. Then the specification for the labour supply model is

$$\begin{aligned} h_i^* = a_1(z_i) + a_2(z_i)\log w_i + a_3(z_i)\mu_i + a_4(z_i)[\mathbb{I}(S_i^*>0)S_i^*]w_i \\ + a_5[\mathbb{I}(S_i^*>0)S_i^*] + u_{hi} \end{aligned} \tag{3.2}$$

where $\mathbb{I}(A)$ is one when A is true and zero otherwise. In (3.2) z_i is a set of demographic variables, w_i is the marginal after tax wage rate and μ_i is "other income" defined by the budget identity

$$\mu_i = C_i - w_i h_i$$

C_i being observed household consumption. The interaction terms we have included in (3.2) imply that income and substitution effects are allowed to vary across different demographic groups. The model presented in Section 2.1 implies that $a_4(z_i) = a_3(z_i)a_5$.[5] The within-period indirect utility function corresponding to the labour supply equation used to derive (3.2) from the generic form (2.4) is given by Stern (1986) and is

$$U^*(w, \mu) = [a_1(z) + u_h + a_2(z)\log w + a_3(z)\mu][\exp(a_3(z)w)/a_3(z)] - [(a_2(z)/a_3(z))Ei(a_3(z)w)],$$

where $Ei(x) = \int_{-\infty}^{x} e^t/t dt$. Hence we interpret the error term u_h as reflecting random variation in the parameter $a_i(z)$ of the utility function across the population. The integrability conditions for this model are $a_2(z) \geqq a_3(z)wh$.

To complete the model we specify reduced-form equations for the wage rate

$$\log w_i = c'x_i + u_{wi}, \tag{3.3}$$

for other income

$$\mu_i = d'x_i + u_{\mu i} \tag{3.4}$$

and for labour force participation

$$I_i = \delta' x_i + u_{Ii}, \tag{3.5}$$

I_i being positive for participants. We assume that the errors in (3.1) to (3.5) are $N(0, \Sigma)$. This implies that $E(w_i | x_i) = \exp(c'x_i + \frac{1}{2}\sigma_w^2)$ a fact which will be useful in what follows.

In the absence of search costs and fixed costs and with a degenerate wage offer distribution, (3.5) would be defined by the labour supply equation (3.2) since all those with a given market wage above their reservation wage ($h_i^* > 0$) would also be labour market participants. But in general there is no reason for this to be the case and we do not impose any restrictions between (3.5) and (3.2).

The model we presented is still valid in the context of a single search intensity with some modifications. The search time term $\mathbb{I}(S_i^* > 0)S_i^*$ is replaced by $S_i = \mathbb{I}(S_i^* > 0)$. Then $a_5 S_i$ are interpreted as fixed time-costs of search and $(a_4/a_5)S_i w_i$ is the market value of this time. Finally in this context equation (3.1) is interpreted as the reduced form for the gain in lifetime utility from search in the current period: an individual searches if this gain is positive.

3.1. *Identification*

An issue that requires some discussion relates to the identification of such a system of equations. First if h, w, μ and S were all observable on the same sample it would be possible (although not necessarily satisfactory) to identify all the parameters using exogeneity assumptions. Moreover, given other over-identifying restrictions such exogeneity assumptions could be tested. Since (h, I, S) are observed in one sample and (w, μ) in another, (3.2) can only be identified provided suitable exclusion restrictions are available. Thus we assume that male and female education and demand-side variables (regional vacancies and redundancies) do not enter (3.2) directly. In addition, we have made the identifying assumption that the errors in the reduced-form search equation (3.1) and in the participation equation (3.5) are uncorrelated. These restrictions are sufficient

5. In fact, the theoretical model also implies that $a_5 = -1$, but since only the sign of S_i^* is observed a_5 cannot be distinguished empirically from the standard deviation of S_i^*.

to identify (3.2) in the case when w, μ, S and I are all endogenous for the labour supply equation. Moreover we also considered and tested the use of skill dummies as additional instruments.

Finally, economic theory suggests that both the realised wage rate w and other income μ should be correlated with on-the-job search activity S as well as with all the moments of the truncated wage offer distribution. It would thus be interesting, in principle, to identify and estimate wage and income effects in a structural search-intensity equation. There are two problems with this. First it is difficult to imagine any plausible exclusion restrictions in (3.1) that would allow us to identify an income and wage effect. Demand-side variables, education, and demographic characteristics are in principle determinants of search activity since they reflect search costs, preferences and expectations.

The second problem relates to the interpretation of the wage coefficient, if identified by some exclusion restriction. What economic theory suggests is that individuals with bad draws from the wage-offer distribution would be more likely to search on-the-job. Hence a relevant explanatory variable is $\log w - E(\log w \mid x)$. By introducing $\log w$ in the equation and instrumenting it we can only identify the effect of $E(\log w \mid x)$ on search activity. While this may well be a relevant explanatory variable its interpretation is by no means clear.[6] Nevertheless in our empirical results we do present a search equation with a wage effect, identified by excluding the variables relating to the husband.

The model as specified splits the Labour Force Survey into four parts as shown in Figure 1 below. In the brackets we show the size of each cell. In that figure f^s is the conditional density function for hours of work given the individual is a labour force participant and a job seeker, f^{NS} the conditional density function of hours given the individual is a participant and not seeking on-the-job. P is the probability of labour force participation and P^S the probability of seeking work while employed. P^E is the employment probability. This probability, which relates to job arrival rates and layoff rates, is assumed independent (conditional on the observables) of P^S, f^{NS} and f^S.[7] The wage and unearned income information coming from the Family Expenditure Survey, together with the identifying restrictions mentioned above allow us to disentangle each of these components. Thus P is identified by "comparing" the stock of non-participants to the rest of the population. P^S is identified by "comparing" the employed job seekers to the employed non-seekers. The densities of hours of work can be identified from the variation of hours within each cell. Moreover as far as these densities are concerned there are

Employed Job Seekers $f^S P P^S P^E$ (581)	Employed Non Seekers $f^{NS} P(1-P^S)P^E$ (10954)
Unemployed Job Seekers $(1-P^E)P$ (1184)	Non-participants $1-P$ (10581)

FIGURE 1

6. In the context of an explicit structural search equation this problem may not arise since suitable functional form restrictions may identify the effect of the current wage on search intensity.

7. Note that Figure 1 does not imply itself any restrictions on the state probabilities P^E and P^S which can be interpreted as conditional probabilities. It is only in the interpretation of regression results that the distinction between conditional and unconditional matters.

cross-cell restrictions which originate from the structure of the model and which we exploit. Finally the employment probability P^E can be identified from the comparison of the employed workers to the unemployed job seekers. Since P^E is peripheral to this study we do not estimate it.[8]

3.2. *Estimation and diagnostic tests*

Given equations (3.1)–(3.5) the model can be estimated over the two separate samples maximizing a combined log-likelihood criterion of the type described in Appendix B.1. Yet, given the size of our sample and the number of parameters to be estimated, it is computationally practical to use a two-stage technique, at the expense of some efficiency loss. An advantage of the two-step estimator is that the results can be more easily replicated by other researchers. Thus the marginal wage equation (3.3) and the other income equation (3.4) have been estimated using the FES for 1983. Then using LFS (1983) we use a probit between working seekers and non-seekers to estimate (3.1). Since we have assumed that $E(u_s u_I)=0$ we can estimate this equation separately from the participation equation. Next we use a probit to estimate the participation equation (3.5). In line with the discussion of Section 2 we classify all workers and all unemployed job seekers as participants ($I>0$) and the rest as non-participants. This approach differs from traditional studies that classify unemployed job seekers with the non-participants.

As far as the labour supply function is concerned, we replace $\mathbb{I}(S^*>0)S^*$ by the dummy variable $\mathbb{I}(S^*>0)$. This in effect assumes that there is only one search intensity. Below we also discuss a variable search intensity specification. For the single search intensity case the conditional expectation of hours, given the observables in the LFS, is

$$E(h_i^*|z_i, x_i, I_i>0, S_i) = a_1(z_i)+a_2(z_i)c'x_i+a_3(z_i)d'x_i + a_4(z_i)S_i \exp(c'x_i+\tfrac{1}{2}\sigma_w^2)+a_5S_i+a_6\lambda_i^P+a_7\lambda_i^s+a_8S_i\exp(c'x_i)\lambda_i^s \tag{3.6}$$

where

$$S_i=\mathbb{I}(S_i^*>0)$$
$$\lambda_i^P=\phi(\delta'x_i)/\Phi(\delta'x_i),$$
$$\lambda_i^S=\begin{cases}\phi(b'x_i)/\Phi(b'x_i) & \text{if } S_i=1\\ -\phi(b'x_i)/(1-\Phi(b'x_i)) & \text{if } S_i=0.\end{cases} \tag{3.7}$$

These terms are generalised residuals controlling for the possible endogeneity of participation and search in the hours of work equation.[9] Note moreover, that conditioning on $I_i>0$ is sufficient since the probability of being in employment, given $I_i>0$, is assumed to be independent of u_{hi} as mentioned above. Hence

$$h_i=E(h_i^*|z_i, x_i, I_i>0, S_i)+v_i \tag{3.8}$$

8. The lack of correlation of the employment index with hours of work was tested by Blundell, Ham and Meghir (1988) on FES data and the hypothesis was accepted.

9. Note that $E(S_iw_i|S_i, x_i)=S_i\exp(c'x_i+\frac{1}{2}\sigma_{wo}^2)E(\exp(\rho u_{si})|S_i)$, where σ_{w0}^2 is the conditional variance of the log-wage given S_i^* and ρ is the correlation between u_s and u_w. However given this expression is non-linear in ρ, which cannot be estimated from elsewhere, we used the first-order approximation $E(S_iw_i|x_i, S_i)\approx S_i\exp(c'x_i+\frac{1}{2}\sigma_w^2)+\rho S_i\exp(c'x_i+\frac{1}{2}\sigma_w^2)E(u_{si}|S_i)$. The original function can be approximated arbitrarily closely by adding higher-order conditional moments of u_{si} (see Lee (1982)). For our purpose of testing for the exogeneity of search the first-order approximation is sufficient.

where

$$v_i = u_{hi} + a_2 u_{wi} + a_3 u_{\mu i} - a_6 \lambda_i^P - a_7 \lambda_i^S - a_8 S_i \exp(c'x_i)\lambda_i^s. \tag{3.9}$$

Since wages, "other income" and hours of work are not observed in the same sample and since u_{hi}, u_{wi} and $u_{\mu i}$ may be correlated we cannot identify $E(u_{hi}^2)$, but only $E(u_{hi} + a_2 u_{wi} + a_3 u_{\mu i})^2$.

Given consistent parameter estimates for b, c, σ_w^2, δ and d the remaining parameters in (3.6) can be estimated by ordinary least squares on the sample of workers. The standard errors of these parameters must then be adjusted to take into account that we are conditioning on estimated parameters and that v_i in (3.8) is heteroscedastic since it is a function of the generalised residuals (see among others Lee (1982), Pagan (1986)). The derivation of the standard errors is presented in Appendix B.4.

In the variable search intensity case the conditional expectation $E(S_i^* w_i \mid S_i, x_i)$ is required. Since neither S_i^* nor w_i are directly observable in the same survey and since no information is available in the data relating to their joint distribution, the estimation of this model would require further prior information relating to the moments of the joint distribution.

Finally, we consider a set of diagnostic tests that will help to evaluate the statistical properties of the model. First, testing whether λ^P and λ^S are significant in (3.6) amounts to an exogeneity test on the participation and search decisions respectively (see also Smith and Blundell (1986)). Next we use normality tests on the reduced-form equations for participation and on-the-job search based on the third- and fourth-order generalised residuals of the probit equations (see Bera, Jarque and Lee (1984) and Gourieroux, Monfort, Renault and Trognon (1984)). Moreover we also present a test of over-identifying restrictions for the structural on-the-job search equation. To evaluate the assumed preference specification we compute a Wald statistic for non-linearity of the labour supply equation in the wage, by testing the significance of the predicted squared log-wage in the labour supply equation. Finally we experimented with several instrument sets. A complete discussion of these experiments is presented in the next section.

4. RESULTS

We now turn to the empirical results. We first consider the compatibility of the two data sets. *A priori*, there is no reason for the FES and the LFS to be incompatible since: (a) they are collected by the same government agency from the same population and (b) the definition of most of the variables in the two questionnaires is the same. In the Data Appendix we present simple descriptive statistics for all variables used in the analysis, for both the LFS and the FES. Moreover, since there are some differences in the mean of hours in the two surveys we present the percentiles for hours in both surveys. However, the differences in the unconditional distribution do not matter so long as the conditional distribution of hours are the same. In order to compare the lower-order moments of the conditional distribution of *observed* hours we have run regressions of hours (including the zeros) and hours-squared on a set of conditioning variables, observed in both surveys and for both workers and non-workers. We test for the equality of slope coefficients in the two regressions respectively. For the regression of hours we get a test statistic of 34 which is distributed as χ^2 with 15 degrees of freedom. For the hours-squared regression the equivalent statistic is 39 (again 15 df.). These test statistics are quite acceptable and hence we believe that for our purposes the two surveys are compatible. The only difficulty arises with the definition and construction of the skill/occupational variables. In the FES these are based on self-categorization while in the LFS these are constructed on the basis of the answers to other questions. Moreover the categories used are not identical. This implies that the skill variables may not be appropriate instruments for "linking" the two data sets. Since we have other identifying information we can test whether the skill variables are valid instruments.

We now turn to the actual results. We first estimate, using the FES, a log-wage equation on the sample of workers only and an equation for other income on the whole sample. To test for selection bias in the wage equation we included the inverse Mill's ratio identified using asset income in the FES participation equation. The t-ratio for the selection term was 1·5.[10] The resulting parameters with heteroscedasticity-consistent standard errors are presented in Table A.2 of Appendix A.

Using the LFS we estimate a reduced form probit for on-the-job search (column 4 of Table A.1 Appendix A). The Normality test for this equation is 4·7 (χ^2_2) which is quite acceptable. Using the minimum-distance method (as described in Appendix B.1) to impose a number of exclusion restrictions, we estimate the search equation presented in Table 4.1. This includes a wage and unearned income effect. The exclusion restrictions imposed are listed at the bottom of the table.[11] The test of over-identifying restrictions is 24·3 and is distributed asymptotically chi-squared with 19 degrees of freedom. Thus the model is not at odds with the assumed statistical properties.

In Table 4.1 education is the age at which the woman left full-time education; vacancies and redundancies are regional measures of the vacancy and redundancy rate while the dummies $D_j (j = 1, \ldots, 4)$ point to the age group of the youngest child of the household. The age groups are 0–2, 3–5, 5–10 and 11+. When $D_j = 0$ for all j then there are no children in the household. At most one of the D_js can be one for any household.

The results show strong age, education, demographic and wage effects. Moreover, of the demand-side variables redundancies seem to have a significant negative effect while vacancies are not very significant. The redundancy rate reflects job arrival rates. When the redundancy rate is high the returns to job search will be low. The fact that education has an independent effect over and above the wage may also reflect better job arrival rates for educated individuals. On the other hand education may be correlated with lower replacement rates (even given the wage); this would provide an incentive for educated people to seach on-the-job rather than quit. Less educated individuals may find it more productive to quit in order to search for an alternative employment. The drop of the probability of on the job search with age may also reflect lower job arrival rates but also greater attachment to their current employment for reasons not captured by the included variables . The presence of young children does not seem to have a significant effect, but women whose younger child is at school ($D3 > 0$ or $D4 > 0$) are more likely to search on-the-job. Finally we find a strong and negative effect of the wage. Since this variable is instrumented and the wage residual is unobservable we can only capture the effects of the mean of the wage-offer distribution and not the effect of the actual position relative to the mean. Such a "surprise" term is not identified. The negative wage effect here reflects the higher opportunity cost of search time for high-wage individuals. The other income variable does not have a significant effect.

We re-estimate this search equation, excluding all skill dummies. The alternative reduced forms are presented in Appendix A, Table A.1 column 2 and Table A.2 columns 1 and 2. The minimum-distance results are presented in column B of Table 4.1. The main qualitative conclusions remain unaffected. Although there is some loss in precision, the differences are significant.

Next we discuss the labour supply results. These have been estimated without using skill dummies as instruments.

TABLE 4.1

On-the-job search

	A		B	
	Parameter	Standard error	Parameter	Standard error
Intercept	−1·240	0·180	−1·540	0·190
Age	−0·130	0·024	−0·140	0·025
Education	1·280	0·150	0·670	0·230
Vacancies	−0·360	0·190	−0·140	0·200
Redundancies	−1·000	0·430	−0·490	0·440
D1	0·089	0·084	0·080	0·085
D2	0·052	0·087	0·160	0·098
D3	0·165	0·078	0·290	0·087
D4	0·091	0·080	0·240	0·087
Log wage	−0·818	0·150	−0·110	0·255
Other income	0·0006	0·0016	−0·0024	0·0016
Non-Normality (2)	4·7		5·2	
Skewness (1):	4·6		5·0	
Kurtosis (1):	4·1		4·7	

Test of over-identifying restrictions
Model A: (19 Degrees of freedom): 24·30
Model B: (13 Degrees of freedom): 11·08

Notes.
Model A uses Skill Dummies as instruments while Model B does not.
Age: (Age − 40)/10
Education: (Age at the end of full time Education − 15)/10
Other income (μ) is measured in pounds per week and the wage in pounds per hour.
Exclusion Restrictions: Model A: Skill Dummies, Regional Unemployment, Number of Children, Age squared, Education Squared, Education × (D1 D2 D3 D4), Age × Education Male Education × Female Education plus all male characteristics. Model B: As above except for the Skill Dummies that are not in any of the reduced forms.
Reduced forms: Model A: Table A.1, column 4 and Table A.2 columns 3 and 4. Model B: Table A.1, column 2 and Table A.2 columns 1 and 2

10. The other income equation is estimated on the whole sample, except when we experiment with Skill Dummies as instruments. These are observed only for working women.

11. We would like to thank John Ham for suggesting the usefulness of the minimum-distance method when combining data sets. In Appendix B.1 and B.3 we relate the minimum distance method to our overall methodology for combining data sets.

To implement the estimator we first estimate a labour force participation equation by a probit between labour force participants (workers and unemployed job seekers) and non-participants. This reduced-form equation has been estimated using the LFS and the results are presented in the first column of Table A.1 in Appendix A.

The participation equation contains age, education, demographic characteristics as well as demand-side variables. Overall the results are very sensible. Age has a negative effect, the presence of children in any age group has a negative effect, the presence of pre-school children (D1 and D2) having the strongest effect by far. Female education has a strong positive effect (mainly a wage effect) while male education reduces the probability of participation (mainly an income effect). Finally note that the chi-square diagnostic tests presented at the bottom of Table A.1 do not indicate mis-specification.

The participation equation and the reduced-form search equation (columns 1 and 2 of Table A.1) are used to construct the generalised residuals λ^P and λ^S which are used to correct for selectivity and for the possible endogeneity of search respectively, when estimating the labour supply equation. The reduced-form parameters of the log-wage equation and other income equations (the first two columns of Table A.2) are used with

LFS variables to impute log wage and other income values for each working household in the LFS. The standard errors need to be corrected and this is done for all models (see Appendix B.4).

The labour supply results are presented in Table 4.2. In the first column we present a neoclassical labour supply model which does not include any search variable. This model is interesting in itself since it is consistent with the presence of fixed costs of work, search costs for the unemployed and intertemporal optimization under uncertainty. Both the log wage and unearned income are instrumented.[12] An interesting result is the strong and positive wage effect. The implied labour supply elasticity at the sample means is about 0·36 with a standard error of 0·06. This elasticity is higher than any other elasticity estimated for the U.K. and more in line with those obtained in some U.S. studies. Earlier results on smaller samples and using a different methodology, have found elasticities that are much closer to zero and often negative. Yet the results are typically much less precise. We carried out a one degree of freedom test of non-linearity in the log-wage term. The t-value for log-wage squared was 1·2. The implied elasticity at sample means for this alternative model was 0·42. The remaining properties of the model are unaltered.[13]

The interactions of the log-wage with Age and the child dummies are not as important but in general the presence of children tends to increase the wage elasticity except in the case of young children (note that only one of the child dummies can be equal to one for a household since these variables point to the age group of the youngest child).

Turning now to the effects of the other income measure we see that it is negative and significant. The elasticity of the sample means is about $-0\cdot13$. The effect of age is not particularly significant, but the presence of pre-school children ($D1=1$ or $D2=1$) increases significantly the income coefficient. In Table 4.3 we present some wage and other income elasticities.

The overall effect of age on labour supply is negative. Since the other income measure captures the life-cycle effects this is interpreted as a cohort effect. The presence of children less than 11 years of age has a negative effect on labour supply (for most wages and other income) while the labour supply of women whose youngest child is older than 11 years ($D4=1$) is higher at relatively high wages (everything else kept constant). Finally the selectivity term λ^P is positive as expected. Yet the effect is not as strong as one might anticipate. Overall the basic labour supply model that comes out of this analysis makes sense and provides particularly interesting results.

We estimated the same labour supply model using female skill dummies as instruments for the wage and other income. The wage effect is higher in this case but the main difference occurs at the other income coefficient which is now three times the size. In fact the wage elasticity at sample means is 0·5 while the other income elasticity is $-0\cdot37$. These differences are significant. A two degree of freedom Hausman test focusing only on the other-income coefficient and the log-wage coefficient is 25 (χ^2_2). Hence skill dummies are inappropriate as instruments for labour supply. Although this could imply that skill is endogenous for labour supply, in our context this rejection is probably due to the incompatibility in the definitions of the skill variables across the two data sets.

Finally, we considered whether female education is a valid exclusion restriction. Including education in the labour supply equation reduced precision but did not have significant effects on the original parameters estimates. The t-value for the coefficient on

12. Clearly their exogeneity cannot be tested or imposed since wages and other income are not observed in the same data set as hours.

13. The test was carried out by estimating a separate log-wage squared reduced form on the FES and imputing a log-wage squared term for women in the LFS and including this as a regressor in the hours equation.

TABLE 4.2

Labour supply and on-the-job Search

	No. search	Search Exogenous		Search Endogenous	
Intercept	30·714	30·724	30·712	30·800	30·802
	(1·193)	(1·201)	(1·202)	(1·229)	(4·196)
Age	−3·3983	−3·4527	−3·4483	−3·5252	−3·5259
	(0·563)	(0·568)	(0·568)	(0·580)	(1·791)
D1	−4·7524	−4·5783	−4·5774	−4·5253	−4·5253
	(3·461)	(3·477)	(3·478)	(3·501)	(13·37)
D2	−11·590	−11·273)	−11·265	−11·011	−11·010
	(3·030)	(3·047)	(3·048)	(3·088)	(15·63)
D3	−12·681	−12·474	−12·472	−12·220	−12·219
	(2·272)	(2·312)	(2·313)	(2·398)	(7·624)
D4	−1·4621	−1·3439	−1·3464	−1·2206	−1·2197
	(2·711)	(2·726)	(2·727)	(2·757)	(18·68)
Log Wage	9·4261	9·6463	9·6677	9·9074	9·9071
	(1·632)	(1·651)	(1·652)	(1·736)	(6·401)
Age × Log Wage	1·6232	1·5136	1·5033	1·3390	1·3389
	(1.100)	(1.111)	(1.112)	(1.117)	(4.010)
D1 × Log Wage	−1·8492	−1·9163	−1·8969	−2·0313	−2·0334
	(2·951)	(3·019)	(3·022)	(3·113)	(8·979)
D2 × Log Wage	3·4171	3·2677	3·2656	3·0974	3·0967
	(4·207)	(4·276)	(4·275)	(4·368)	(10·11)
D3 × Log Wage	8·3448	8·7107	8·7171	9·0948	9·0964
	(4·582)	(4·705)	(4·705)	(4·824)	(12·44)
D4 × Log Wage	6·6349	6·6955	6·7025	6·7918	6·7917
	(4·604)	(4·658)	(4·656)	(4·722)	(14·01)
Other Income (μ)	−0·0420	−0·0429	−0·04297	−0·04439	−0·04440
	(0·017)	(0·017)	(0·017)	(0·017)	(0·062)
Age × Other Income	0·0032	0·0039	0·00393	0·00502	0·00502
	(0·007)	(0·007)	(0·007)	(0·007)	(0·028)
D1 × Other Income	−0·0844	−0·0870	−0·08728	−0·08787	−0·08786
	(0·036)	(0·036)	(0·036)	(0·036)	(0·162)
D2 × Other Income	−0·0523	−0·0547	−0·05485	−0·05596	−0·05596
	(0·033)	(0·033)	(0·033)	(0·033)	(0·186)
D3 × Other Income	−0·0229	−0·0257	−0·02581	−0·02847	−0·02848
	(0·029)	(0·029)	(0·029)	(0·029)	(0·088)
D4 × Other Income	−0·0602	−0·0608	−0·06084	−0·06128	−0·06129
	(0·029)	(0·030)	(0·030)	(0·030)	(0·228)
Search (S)		−3·2382	−2·8874	−7·7960	−7·8481
		(0·466)	(1·978)	(6·763)	(28·48)
Wage × S			−0·18684		0·01482
			(1·048)		(4·472)
λ^s				2·1404	2·1518
				(3·174)	(11·17)
λ^P	0·9044	1·0235	1·0261	1·1281	1·1284
	(1·187)	(1·196)	(1·196)	(1·208)	(4·288)

Notes.

(1) The instruments for the wage, other income and search used for this estimation DO NOT include Skill Dummies. The relevant reduced forms are: log wage and other income: Table A.2 columns 1 and 2, Participation and on-the-job search: Table A.1 columns 1 and 2.

(2) Standard errors in parentheses.

(3) Sample size: 11535 (Workers Only).

(4) Age: (Age − 40)/10.

(5) μ is measured in pounds per week and the wage in pounds per hour.

TABLE 4.3

Wage and other income elasticities
(Hours = 26, Other income = 80, log w = 0·50, Age = 40)

Household Type	Uncompensated wage	Compensated wage	Other income
No. of Children	0·37	0·44	−0·13
D1 = 1	0·29	0·50	−0·40
D1 = 0, D2 = 1	0·50	0·66	−0·30
D1 = D2 = 0, D3 = 1	0·71	0·82	−0·21
D1 = D2 = D3 = 0, D4 = 1	0·62	0·77	−0·32

education was 1·8 and hence we maintain this restriction. The above conclusions are valid for all models presented in the paper (The tests were in fact performed with search both included and excluded).

We now turn to the models that include variables relating to on-the-job search. The main results that we discussed above do not change: ignoring on-the-job search does not bias the other labour supply parameters. Hence we now concentrate on the parameters of the search variables. We first consider the issue of whether search can be treated as exogenous. The test for exogeneity of search is just a test for the significance of λ^S in the search equation. The t-ratio for this coefficient is 0·67.[14] Moreover none of the coefficients of the model change significantly when we instrument search (see the last two columns of Table 4.2). On the basis of these results it is valid to treat search as exogenous for the purposes of estimating the hours equation.

In the second column of Table 4.2 we have simply included a search dummy. The third column relates to the single search intensity model discussed earlier. The results from the first two indicate that job seekers do work less than non-seekers at any given wage and other income as suggested by the theoretical model. This effect is very precisely estimated for the model including just a search dummy and it amounts to about −3·2 hours a week. When we instrument search we obtain an even larger effect (−7·8 hours) but this difference is not significant. The interaction of search with the wage although negative and contrary to the theoretical predictions, is not significant.[15] We illustrate the importance for the purposes of inference of correcting standard errors in Table 4.4.

Finally, there are other interpretations that one could give to the search variable. For example, the job seekers may be individuals who are dissatisfied with their number of hours and cannot change them in their current job. This would be consistent with a model where job offers arrive as fixed wage/hours packages. If that were the case, the appropriate empirical strategy, given our data, would be to allow all parameters of the labour supply function to be different for the two groups (seekers and non-seekers) (see Ham 1982)).

We tried this and the test statistic that all parameters (apart from the intercept) are equal was 14 which is distributed as a χ^2 variable with 18 degrees of freedom. When we include the intercept in the test, the test statistic becomes 46 (19 degrees of freedom), confirming that the main differences are in the level. Given these results, a "constraints" interpretation would imply that if at all constrained the job seekers are on average underemployed. Overall the data does not reject the theory underlying our model but there is not sufficient information to discriminate definitively this model from one where

14. The term $S_t \exp(c'x_t)\lambda_t^s$ proved to be empirically irrelevant so we did not include it in the results we presented.

15. Similar results were obtained when we used skill variables as instruments.

TABLE 4.4

Comparison of corrected and uncorrected standard errors
Model with search dummy instrumented

	Corrected	Uncorrected
Intercept	1·229	0·815
Age	0·580	0·390
D1	3·501	2·481
D2	3·088	2·445
D3	2·398	1·579
D4	2·757	2·016
Log Wage	1·736	1·016
Age × Log Wage	1·117	0·688
D1 × Log Wage	3·113	1·661
D2 × Log Wage	4·368	2·047
D3 × Log Wage	4·824	2·415
D4 × Log Wage	4·722	2·218
Other Income	0·017	0·010
Age × Other Income	0·007	0·005
D1 × Other Income	0·036	0·027
D2 × Other Income	0·033	0·027
D3 × Other Income	0·029	0·021
D4 × Other Income	0·030	0·024
Search (S)	6·763	5·458
λ^s	3·174	2·554
λ^P	1·208	0·870

offers arrive as fixed wage/hours packages. As discussed by Altonji and Paxson (1987) it would be possible to investigate the actual structure of job offers if we had mobility data. In particular we would want to observe people changing jobs and people changing hours within the job, as well as the pre and post change wage. On the other hand we can say that the neoclassical model, adapted for the observed search activity, is not rejected by our data.

5. CONCLUSIONS

In this paper an empirical model of labour supply that is consistent with on-the-job search has been developed and successfully identified and estimated by combining two different sources of data: The U.K. Family Expenditure Survey, which contains accurate information on income and expenditure and the larger U.K. Labour Force Survey, which concentrates on data relating to hours and job search behaviour. Moreover, since hours of work are observed in both surveys we have been able to provide formal statistical evidence on the compatibility of the two samples for the purposes of estimating a labour supply equation.

The main empirical results are as follows:

(1) There is strong evidence that job search has a direct negative effect on hours. However, the positive income effect through the wage rate, predicted by the theory, is not confirmed by our results using this data.
(2) Ignoring on-the-job search does not bias the wage and income effects.
(3) We establish a strong positive wage effect and a negative income effect on women's hours of work. Particularly interesting is the magnitude of the labour supply elasticity which at sample means is about 0·4.

DATA APPENDIX

TABLE A

	LFS				
	Workers		Non-workers		FES
	Nonseekers	Seekers	Nonseekers	Seekers	Workers
Hours	28·48	25·76			26·228
	(12·79)	(13·02)			(11·92)
Log Wage					0·51
					(0·43)
Other Income					80·587
					(59·48)
Predicted Log Wage	0·52	0·53			
	(0·23)	(0·25)			
Predicted other Income	78·4	77·4			
	(24·0)	(25·0)			
Education	0·11	0·18	0·09	0·11	0·124
	(0·19)	(0·23)	(0·17)	(0·17)	(0·21)
Husband's Education	0·13	0·18	0·13	0·12	0·129
	(0·22)	(0·23)	(0·22)	(0·19)	(0·25)
Age	−0·30	−0·74	−0·46	−0·74	−0·321
	(1·07)	(0·88)	(1·09)	(0·96)	(1·00)
Husband's Age	−0·04	−0·49	−0·18	−0·49	−0·083
	(1·12)	(0·93)	(1·13)	(1·01)	(1·03)
Clerical	0·51	0·51			0·444
	(0·49)	(0·50)			(0·49)
Skilled	0·04	0·02			0·046
	(0·20)	(0·16)			(0·20)
Unskilled	0·18	0·20			0·268
	(0·39)	(0·40)			(0·44)
Husband Clerical	0·11	0·08	0·08	0·09	0·087
	(0·31)	(0·27)	(0·27)	(0·28)	(0·28)
Husband Skilled	0·31	0·30	0·24	0·29	0·357
	(0·46)	(0·45)	(0·42)	(0·45)	(0·47)
Husband Unskilled	0·15	0·13	0·12	0·16	0·155
	(0·35)	(0·34)	(0·33)	(0·36)	(0·36)
Regional Unemployment	15·46	15·71	15·59	16.31	15.563
	(3·49)	(3·36)	(3·47)	(3·27)	(3·47)
Regional Vacancies	0·60	0·61	0·61	0·61	0·622
	(0·09)	(0·09)	(0·09)	(0·10)	(0·13)
Regional Redundancies	0·11	0·11	0·11	0·12	0·113
	(0·04)	(0·04)	(0·04)	(0·04)	(0·05)
Number of Children	0·87	1·00	1·46	1·20	1·028
	(1·05)	(1·03)	(1·12)	(1·03)	(1·07)
D1	0·05	0·05	0·27	0·23	0·046
	(0·21)	(0·23)	(0·44)	(0·42)	(0·21)
D2	0·06	0·07	0·17	0·12	0·062
	(0·24)	(0·26)	(0·38)	(0·32)	(0·24)
D3	0·22	0·29	0·22	0·25	0·235
	(0·41)	(0·45)	(0·41)	(0·43)	(0·42)
D4	0·13	0·12	0·06	0·06	0·209
	(0·33)	(0·33)	(0·25)	(0·25)	(0·40)
Sample Size	10,954	581	10,581	1184	1215

Notes.

LFS = U.K. Labour Force Survey 1983

FES = U.K. Family Expenditure Survey 1983.

Age = (Age − 40)/10, Education: (Age at end of Education −15)/10.

D_j = Dummy which is equal to one when the youngest child is in the j-th a group. Age groups: 0–2, 3–4, 5–10, 11+. Other income is measured in pounds per week and the wage rate is in pounds per hour. Details on all constructed variables in the LFS and FES available from the authors.

TABLE B

Percentiles of the hours distribution in the two surveys

	All workers										
	MIN	M5	M10	M15	M20	M25	M30	M35	M40	M45	M50
FES	1	6	9	12	15	16	18	20	22	25	28
LFS	1	8	10	14	16	18	20	22	25	28	31
	M55	M60	M65	M70	M75	M80	M85	M90	M95	MAX	
FES	30	35	35	36	37	38	38	39	40	70	
LFS	35	35	37	38	38	39	40	41	45	84	

Mi = The *i*-th % percentile.

APPENDIX A

TABLE A.1

The participation and the on-the-job search reduced form probit equations

	Participation-1	Job-search-1	Participation-2	Job-search-2
Intercept	0·965 (0·097)	−1·820 (0·225)	0·654 (0·099)	−1·96 (0·234)
Age	−0·090 (0·023)	−0·119 (0·061)	−0·090 (0·024)	−0·10 (0·062)
Age Squared	−0·125 (0·013)	0·008 (0·034)	−0·121 (0·013)	0·01 (0·035)
Husb. Age	−0·090 (0·022)	−0·042 (0·054)	−0·049 (0·022)	−0·04 (0·055)
Husb. Age Sq.	−0·012 (0·011)	0·009 (0·029)	−0·005 (0·011)	0·01 (0·029)
Education	1·617 (0·166)	0·060 (0·377)	1·766 (0·167)	0·46 (0·392)
Education Sq.	−1·22 (0·279)	0·617 (0·572)	−1·210 (0·281)	0·41 (0·582)
Husb. Educ.	−0·188 (0·139)	0·346 (0·333)	0·290 (0·143)	0·40 (0·350)
Husb. Educ. Sq.	−0·268 (0·220)	−0·345 (0·522)	−0·497 (0·223)	−0·46 (0·533)
Reg. Unem.	0·009 (0·005)	0·006 (0·013)	0·013 (0·006)	0·001 (0·013)
Reg. Vacancies	−0·334 (0·101)	−0·212 (0·235)	−0·324 (0·102)	−0·20 (0·237)
Reg. Redund's	−0·753 (0·474)	−0·742 (1·00)	−0·982 (0·478)	−0·59 (1·09)
Number of Children	−0·114 (0·014)	−0·050 (0·041)	−0·098 (0·015)	−0·06 (0·041)
D1	−1·490 (0·049)	0·073 (0·147)	−1·514 (0·050)	0·07 (0·148)
D2	−1·186 (0·050)	0·253 (0·134)	−1·211 (0·050)	0·24 (0·134)
D3	−0·520 (0·043)	0·338 (0·103)	−0·542 (0·043)	0·32 (0·104)
D4	−0·101 (0·042)	0·278 (0·101)	−0·115 (0·043)	0·26 (0·102)
Husb. Clerical			0·307 (0·032)	−0·22 (0·074)
Husb. Skilled			0·409 (0·022)	−0·06 (0·054)
Husb. Unskilled			0·390 (0·028)	−0·05 (0·067)
Clerical				0·23 (0·060)
Skilled				0·18 (0·125)
Unskilled				0·35 (0·075)
F. Ed × F. Age	−0·056 (0·069)	−0·321 (0·140)	−0·073 (0·069)	−0·31 (0·140)
M. Ed × M. Age	−0·215 (0·045)	0·115 (0·109)	−0·265 (0·046)	0·11 (0·110)
Med × F. Ed	−0·107 (0·237)	−0·212 (0·458)	−0·179 (0·239)	−0·21 (0·463)
F. Ed × D1	−0·046 (0·164)	−0·028 (0·384)	−0·044 (0·165)	0·01 (0·386)
F. Ed × D2	−0·108 (0·170)	−0·419 (0·391)	−0·093 (0·171)	−0·36 (0·389)
F. Ed × D3	−0·164 (0·136)	0·024 (0·239)	−0·150 (0·137)	0·07 (0·240)
F. Ed × D4	0·383 (0·193)	0·014 (0·316)	0·408 (0·194)	0·03 (0·317)
Log-likelihood	−13,709	−2,222·6	−13,476·5	−2,197·0
Non-Normality (2)	3·5	4·7	0·2	5·2
Skewness (1)	0·3	4·6	0·074	5·0
Kurtosis (1)	3·3	4·1	0·078	4·7
Employed Participants = 11,535				
Unemployed Participants = 1,184				
Employed Job Seekers = 581				
Non-Participants = 10,581				

Note. The default skill group are the professional and Managerial Workers.

TABLE A.2

The log-wage and other income equations

	Log-wage-1 Coeff (St. error)	Other income-1 Coef (St. error)	Log-wage-2 Coeff (St. error)	Other income-2 Coeff (St. error)
Intercept	0·776 (0·097)	70·5 (10·3)	0·987 (0·097)	44·1 (14·2)
Age	0·045 (0·033)	7·98 (3·66)	0·025 (0·030)	5·3 (5·01)
Age Squared	−0·061 (0·016)	−0·22 (2·03)	−0·051 (0·015)	−1·7 (2·94)
Husb. Age	−0·001 (0·032)	3·33 (3·16)	0·003 (0·029)	2·5 (4·45)
Husb. Age Sq.	0·021 (0·016)	−3·92 (1·69)	0·022 (0·015)	−1·1 (2·39)
Education	0·814 (0·177)	5·99 (23·5)	0·554 (0·168)	1·1 (27·5)
Education Sq.	−0·038 (0·301)	−40·1 (39·3)	−0·112 (0·282)	17·1 (49·3)
Husb. Educ.	0·325 (0·110)	52·2 (11·9)	0·184 (0·110)	3·1 (18·0)
Husb. Educ Sq	−0·160 (0·072)	−11·7 (7·91)	−0·090 (0·067)	12·7 (9·43)
Reg. Unem.	−0·006 (0·004)	−1·28 (0·54)	−0·003 (0·004)	−0·9 (0·76)
Reg. Vacancies	−0·212 (0·086)	33·5 (8·96)	−0·169 (0·083)	47·3 (12·5)
Reg. Redunds	−0·526 (0·321)	29·4 (38·0)	−0·675 (0·313)	63·0 (53·6)
Number of Children	−0·006 (0·022)	8·04 (1·71)	0·006 (0·021)	12·9 (2·79)
D1	0·019 (0·107)	17·7 (6·08)	−0·003 (0·107)	11·5 (13·2)
D2	−0·089 (0·075)	19·2 (6·36)	−0·091 (0·072)	12·0 (10·9)
D3	−0·045 (0·058)	14·4 (5·76)	−0·037 (0·054)	7·7 (8·23)
D4	−0·101 (0·047)	14·1 (5·07)	−0·095 (0·044)	9·8 (6·82)
Husb. Clerical			−0·023 (0·042)	−7·9 (6·11)
Husb. Skilled			−0·016 (0·016)	−13·3 (4·08)
Husb. Unskill.			−0·051 (0·034)	−15·2 (5·10)
Clerical			−0·275 (0·032)	15·2 (4·78)
Skilled			−0·293 (0·052)	−3·9 (7·03)
Unskilled			−0·340 (0·038)	8·8 (5·12)
F. Ed × F. Ag	0·076 (0·065)	11·0 (13·5)	0·066 (0·061)	2·8 (18·7)
M. Ed × M. Ag	−0·053 (0·050)	17·8 (8·46)	−0·031 (0·046)	15·7 (10·0)
Med × FD. Ed	−0·243 (0·261)	6·8 (33·8)	−0·158 (0·252)	−17·0 (43·3)
F. Ed × D1	0·766 (0·447)	21·2 (23·0)	0·713 (0·450)	−29·8 (41·3)
F. Ed × D2	0·428 (0·207)	33·8 (30·3)	0·445 (0·251)	−39·0 (37·6)
F. Ed × D3	−0·169 (0·146)	32·1 (23·3)	−0·171 (0·137)	−2·3 (25·4)
F. Ed × D4	0·151 (0·159)	29·8 (34·5)	0·035 (0·153)	−1·5 (36·2)
σ	0·38	53·9	0·36	53·9
R^2	0·23	0·182	0·30	0·18
Sample	Workers	All	Workers	Workers

Notes.

1. The other income-1 equation has been estimated on the whole sample. The wage equations, as well as the other income equation-2 which contains the female skill dummies have been estimated on the workers only sample. For a discussion of the selectivity tests see main text.
2. All characteristics that are not explicitly referred to as husband's characteristics, relate to the wife.

APPENDIX B

B.1. *Estimating models from combined samples with missing variables*

A moment-estimation problem defines the parameter of interest θ by a set of restrictions of the form

$$E_p g(x, \theta) = 0. \tag{B.1}$$

θ is identified if there is a unique element of the parameter space θ satisfying (B.1) for every probability distribution P in the family of distributions under consideration. Then the true value θ_0 is estimated by minimizing some criterion function $Q_N(\theta)$ based on a sample of size N, which may be of the M-estimator form (cf. Gourieroux, Monfort and Renault (1987)) or alternatively it can be a GMM or minimum-distance criterion. The class of criterion functions which are able to produce consistent estimates of θ_0 will depend on the form of the constraints (B.1) defining the parameter of interest.

In our model the set of defining constraints that we can write from the observed variables in a given sample is not enough to identify the parameter of interest. However a second sample including observations on additional variables is available, which then provides the complementary number of restrictions that are needed to identify the parameter of interest. On other occasions, the restrictions arising from the variables in a single sample may be enough to identify the parameter of interest, but the availability of a second sample with additional observables may add further restrictions, thus making possible more efficient inferences, and the testing of additional over-identifying restrictions.

Let x_{1i} and x_{2i} be two sub-vectors of x_i which may contain some elements in common. We observe N_1 realizations of x_1 and N_2 realizations of x_2 the two random samples being mutually independent, i.e. no sample information is available on cross-moments of variables which are not contained in both x_1 and x_2. We are interested in estimation problems that can be represented in the form

$$E_P g_1(x_1, \theta)=0 \tag{B.2a}$$

$$E_P g_2(x_2, \theta)=0 \tag{B.2b}$$

where we assume that there is a unique value of θ in Θ which satisfies both sets of equations. So, restrictions that involve interactions between variables which are not in both x_1 and x_2 cannot be exploited to help identification of θ (e.g. covariance restrictions between these components). The estimators we consider are the minimizers of criterion functions of the form

$$\frac{1}{N} Q_N(\theta)=\lambda \frac{1}{N_1} Q_1(\theta)+(1-\lambda)\frac{1}{N_2} Q_2(\theta)$$

where $\lambda = N_1/N$, $N=N_1+N_2$, and Q_1 and Q_2 are criteria associated respectively to (B.2a) and (B.2b).

Given suitable regularity and identification conditions (see Appendix B.2) we can establish the consistency and the asymptotic normality of the minimizer of $Q_N(\theta)$, $\hat{\theta}$ say. A consistent estimate of avar $(\hat{\theta})$ is obtained as

$$\text{a}\,\hat{\text{v}}\text{ar}\,(\hat{\theta})=[N_1A_1(\hat{\theta})+N_2A_2(\hat{\theta})]^{-1}[N_1B_1(\hat{\theta})+N_2B_2(\hat{\theta})][N_1A_1(\hat{\theta})+N_2A_2(\hat{\theta})]^{-1}$$

where $A_j(\theta)$ and $B_j(\theta)$, $j=1, 2$ are Hessian and outer-product matrices, respectively as defined in Appendix B.2. In the usual way, an estimator is said to be asymptotically efficient relative to a given class (Q_1, Q_2) if $A_1(\theta_0)=B_1(\theta_0)$ and $A_2(\theta_0)=B_2(\theta_0)$.

Sometimes it may be possible to write $\pi_1=\pi_1(\theta)$ and $\pi_2=\pi_2(\theta)$ where although θ is not identified separately from a single sample, π_1 is identified from sample 1 and π_2 from sample 2. In such cases one can choose minimum-distance criteria of the form

$$Q_j=(\hat{\pi}_j-\pi_j(\theta))'A_j(\hat{\pi}_j-\pi_j(\theta))$$

where $\hat{\pi}_j$ is a consistent and asymptotically normal estimate of π_j obtained from the j-th sub-sample. Since θ is identified from the reduced-form coefficients π_1 and π_2, estimates of θ that rely exclusively on $\hat{\pi}_1$ and $\hat{\pi}_2$ are consistent and asymptotically normal. Furthermore if $\hat{\pi}_1$ and $\hat{\pi}_2$ are efficient they contain all the relevant sample information about θ so that the resulting estimates of θ are also efficient.

Suppose that $\hat{\pi}_j=\text{argmin}\, s_j(\pi_j)$. If the s_j are efficient criteria, the minimizer of $s_1(\pi_1(\theta))+s_2(\pi_2(\theta))$ is asymptotically equivalent to the optimal two-step MD estimator of θ. However, for inefficient criterion functions, the two-step estimator of θ that uses the optimal MD procedure relative to $\hat{\pi}_1$ and $\hat{\pi}_2$ in the second step will always be at least as efficient as the direct estimator (See Appendix B.3).

B.2. *Regularity conditions for combined criterion functions*

We make the following assumptions:

(i) $\lim_{N\to\infty} \lambda=\bar{\lambda}$ and $0<\bar{\lambda}<1$. This ensures that one sample is not asymptotically irrelevant relative to the other.

(ii) (Identification) $Q_N(\theta)$ is twice-differentiable and

$$\text{plim}_{N\to\infty} N^{-1}Q_N(\theta)=\bar{Q}(\theta) \quad \text{uniformly in } \theta\in\Theta$$

where

$$\bar{Q}(\theta)=\bar{\lambda}\bar{Q}_1(\theta)+(1-(1-\bar{\lambda})\bar{Q}_2(\theta)$$

attains a unique global minimum at θ_0 and

$$\partial\bar{Q}_1(\theta_0)/\partial\theta=\partial\bar{Q}_2(\theta_0)/\partial\theta=0$$

(i.e. θ_0 is a stationary point for both $\bar{Q}_1$ and $\bar{Q}_2$).

(iii) $\text{plim}_{N\to\infty} N^{-1}\partial^2 Q_T(\hat{\theta})/\partial\theta\partial\theta' = A(\theta_0) = \lim_{N\to\infty} E(N^{-1}\partial^2 Q_T(\theta)/\partial\theta\partial\theta')$, where $A(\theta_0)$ is a finite non-singular matrix, for any $\hat{\theta}$ such that

$$\text{plim}_{N\to\infty}\hat{\theta} = \theta_0.$$

Also $A(\theta_0) = \bar{\lambda}A_1(\theta_0) + (1-\bar{\lambda})A_2(\theta_0)$ but A_1 and A_2 will in general be singular matrices.

(iv)
$$N_j^{-1/2}\partial Q_j(\theta_0)/\partial\theta \overset{d}{\to} N(0, B_j(\theta_0)), \qquad (j=1,2)$$

where

$$B_j(\theta_0) = \lim_{N_j\to\infty} E\{N_j^{-1}(\partial Q_j(\theta_0)/\partial\theta)(\theta Q_j(\theta_0)/\partial\theta)'\}$$

so that

$$N_j^{-1/2}\partial Q_N(\theta_0)/\partial\theta \overset{d}{\to} N(0, B_j(\theta_0)),$$

with

$$B(\theta_0) = \bar{\lambda}B_1(\theta_0) + (1-\bar{\lambda})B_2(\theta_0)$$

in view of the independence between the two scores.

Under these conditions (e.g. see Amemiya (1985, p. 111)) it can be proved that

$$\sqrt{N}(\hat{\theta}-\theta)\overset{d}{\to} N((0, C(\theta_0)),$$

where

$$C(\theta_0) = [\bar{\lambda}A_1(\theta_0) + (1-\bar{\lambda})A_2(\theta_0)]^{-1}[\lambda B_1(\theta_0) + (1-\bar{\lambda})B_2(\theta_0)][\bar{\lambda}A_1(\theta_0) + (1-\bar{\lambda})A_2(\theta_0)]^{-1}.$$

B.3. *Efficiency comparisons for estimators based on inefficient reduced-form criteria*

Let $\hat{\pi}$ be a consistent and asymptotically normal unconstrained estimator of the coefficient vector $\bar{\pi}$ defined to be the minimizer of some criterion function $s(\pi)$. Assume that $\bar{\pi}$ depends on a set of constraint parameters $\bar{\theta}$, $\bar{\pi} = \pi(\bar{\theta})$, and that standard regularity and identification conditions are satisfied.

Let the asymptotic variance matrix of $\sqrt{N}\,(\hat{\pi}-\bar{\pi})$ be

$$\text{avar}\,(\hat{\pi}) = V = A^{-1}BA^{-1}$$

where $A = \text{plim}\, N^{-1}[\partial^2 s(\bar{\pi})/\partial\pi\delta\pi']$ and B is the asymptotic variance matrix of $N^{-1/2}\partial s(\bar{\pi})/\partial\pi$. We say that $s(\pi)$ is an efficient criterion function if $A = B$.

Now we consider the relative efficiency of two alternative estimators of $\bar{\theta}$. The first one, $\tilde{\theta}$, is the minimizer of $s[\pi(\theta)]$. The second estimator, $\hat{\theta}$, is the optimal MD estimator based on $\hat{\pi}$.

In order to derive the asymptotic distribution of $\tilde{\theta}$ note that

$$\text{plim}\,\frac{1}{N}\frac{\partial^2 s[\pi(\bar{\theta})]}{\partial\theta\partial\theta'} = D'AD$$

and that

$$\frac{1}{\sqrt{N}}\frac{\partial s[\pi(\bar{\theta})]}{\partial\theta} = \frac{1}{\sqrt{N}}D'\frac{\partial s(\bar{\pi})}{\partial\pi}\overset{d}{\to} N(0, D'BD),$$

where $D = (\partial\bar{\pi}/\partial\theta')$. Hence from a first-order expansion of $\partial s[\pi(\hat{\theta})]/\partial\theta$ about $\bar{\theta}$ in the usual way we obtain

$$\text{avar}\,(\tilde{\theta}) = (D'AD)^{-1}(D'BD)(D'AD)^{-1}.$$

On the other hand $\hat{\theta}$ minimizes

$$c(\theta) = [\hat{\pi}-\pi(\theta)]'\hat{V}^{-1}[\hat{\pi}-\pi(\theta)],$$

where $\hat{V}$ is a consistent estimate of V. MD theory tells us that

$$\text{avar}\,(\hat{\theta}) = (D'V^{-1}D)^{-1} = (D'AB^{-1}AD)^{-1}.$$

Finally note that

$$[\text{avar}\,(\hat{\theta})]^{-1} - [\text{avar}\,(\tilde{\theta})]^{-1} = D'A[B^{-1} - D(D'BD)^{-1}D']AD,$$

which is non-negative definite. So that $\text{avar}\,(\hat{\theta}) \leqq \text{avar}\,(\tilde{\theta})$, with equality when $A = B$. Efficiency gains of this kind have been discussed by a number of people, including Cragg (1981), Chamberlain (1982) and White (1982). The case we discussed in Appendix B.1 is when $s(\pi) = s_1(\pi_1) + s_2(\pi_2)$.

B.4. Computing the standard errors

Here we derive the standard errors for the most general model we estimate in the paper. The standard errors for the other models are simple special cases of those that follow.

The model we estimate has the form

$$y_1 = z_i' a_1 + (\hat{c}' x_i) a_2(z_i) + (\hat{d}' x_i) a_3(z_i) + a_4 S_i + \hat{w}_i S_i a_5 + a_6 \lambda_i^s(\hat{b}) + a_7 \lambda_i^P(\hat{\delta}) + v_i, \tag{B.3}$$

where

$$\hat{w}_i = \exp(\hat{c}' x_1 + \tfrac{1}{2}\sigma_w^2), \qquad \lambda_i^p(\hat{\delta}) = \phi(x_i'\hat{\delta})/\Phi(x_i'\hat{\delta})$$

and

$$\lambda_i^s(\hat{b}) = \begin{cases} \phi(x_i'\hat{b})/\Phi(x_i'\hat{b}) & \text{if } S_i > 0 \\ -\phi(x_i'\hat{b})/(1-\Phi(x_i'\hat{b})) & \text{if } S_i \leq 0 \end{cases}$$

$\phi(\cdot)$ being the standard normal density function and $\Phi(\cdot)$ the standard normal distribution function. We omit from consideration the term $S_i\lambda_i^s$ since it proved empirically irrelevant.

The error term in (B.3) can be approximated to first order by the following expression

$$v_i \simeq u_i + (a_2' z_i)[x_i'(\hat{c} - c)] + (a_3' z_i)[x_i'(\hat{d} - d)] + (a_5 \hat{w}_i S_i)[x_i', \tfrac{1}{2}]\begin{pmatrix} \hat{c} - c \\ \hat{\sigma}_w^2 - \sigma_w^2 \end{pmatrix}$$

$$+ a_6 q_{si} x_i'(\hat{b} - b) + a_7 q_{pi} x_i'(\hat{\delta} - \delta), \tag{B.4}$$

where $q_i = \partial\lambda_i/\partial(x_i'\delta) = -\lambda(\lambda + x_i'\delta)$. Since the probits for participation and on-the-job search have been estimated on a different sample from the wage and unearned income equation, $\hat{b}$ and $\hat{\delta}$ are not correlated with $\hat{c}$ and $\hat{d}$. In computing the standard errors we simplify the computations by ignoring the correlation between those estimators that are correlated. We could make sure this was true by estimating each of the reduced form equations on a different sub-sample but we have not done this. Moreover we ignore the variance of $\hat{\sigma}_w^2$. For presentational simplicity we have assumed that all reduced forms contain the same regressors. Given the above, we group terms to obtain

$$v_i \approx u_i + [a_2' z_i + \hat{w}_1 S_i a_5](x_i'(\hat{c} - c)) + [a_3' z_i] x_i'(\hat{d} - d) + [a_6 q_{si}] x_i'(\hat{b} - b) + [a_7 q_{pi}] x_i'(\hat{\delta} - \delta), \tag{B.5}$$

Since the equation (B.3) contains two hazard functions (λ) the error term u_i in (B.4) is heteroscedastic. This is true irrespective of the fact that the hazard rates are estimated. Denote Euu' by Σ which is assumed to be a diagonal matrix. Let the matrices V_1, V_2, V_3 and V_4 denote the asymptotic covariances for $\hat{c}$, $\hat{d}$, $\hat{b}$ and $\hat{\delta}$ respectively. Clearly V_1 and V_2 will be estimated on one sample while V_3 and V_4 on another. Then (allowing now for the fact that each reduced form may include different variables) we have that Evv' is given by

$$\Omega = \Sigma + \textstyle\sum_{j=1}^4 A_j X_j V_j X_j' A_j \tag{B.6}$$

where A_j are diagonal matrices with the terms in the square brackets in (B.5) as elements. Each element corresponds to a different observation in the sample in which the labour supply equation is estimated. Denote by Q the matrix of observations of all the explanatory variables (including the generated ones) on the right-hand side of (B.3). Thus the estimated covariance matrix of the estimator of all the unknown coefficients in (B.3) obtained by regressing y on Q is

$$V(\hat{\beta}) = (Q'Q)^{-1} Q'\hat{\Omega} Q (Q'Q)^{-1}. \tag{B.7}$$

The estimate of Ω, $\hat{\Omega}$ is obtained by replacing all unknown parameters by consistent estimates. In place of Σ we use diag $(\hat{v}_1^2)$. An equivalent to this would be to use the exact form of heteroscedasticity implied by the inclusion of the hazard functions. However, the method we use is simpler to implement.

Acknowledgements. We should like to thank two anonymous referees for their detailed comments and patience. We are indebted to John Ham for extensive comments and helpful suggestions. Thanks are also due to Richard Blundell, Jim Heckman, Arie Kapteyn, Tom MaCurdy, Chris Pissarides, Mark Stewart, Andrew Trigg and the participants at the Warwick Summer Workshop 88 for useful comments and discussion. We should also like to thank Andrew Trigg for his excellent research assistance. Part of this research was undertaken while Costas Meghir was at the Institute for Fiscal Studies. The hospitality and financial assistance of the IFS is gratefully acknowledged. We thank the IFS, the Department of Employment and the ESRC Data Archive at Essex for providing the data. Finance for this research was provided by the Leverhulme Trust. All errors and shortcomings are our own.

REFERENCES

ALTONJI, J. G. (1986), "Intertemporal Substitution in Labour Supply: Evidence from Micro Data", *Journal of Political Economy*, **94**, S176-S215.

ALTONJI, J. G. and PAXSON, C. H. (1987), "Labour Supply Preferences, Hours Constraints and Hour-Wage Trade-offs" (NBER Working Paper, No., 2121).

AMEMIYA, T. (1985) *Advanced Econometrics* (Oxford: Blackwell).

BERA, A. K., JARQUE, C. M. and LEE, L. F. (1984), "Testing for the Normality Assumption in Limited Dependent Variable Models", *International Economic Review*, **25**, 563-578.

BLUNDELL, R. W., DUNCAN, A. and MEGHIR, C. (1991), "Taxation in Empirical Labour Supply Models: Lone Parents in the UK", (Institute for Fiscal Studies, mimeo). *Economic Journal* (forthcoming).

BLUNDELL, R. W. and WALKER, I. (1986), "A Life-Cycle Consistent Empirical Model of Family Labour Supply using Cross-Section Data", *Review of Economic Studies*, **53**, 539-558.

BLUNDELL, R. W., HAM, J. C. and MEGHIR, C. (1988), "Unemployment, Discouraged Workers and Female Labour Supply" (UCL Discussion Paper 88/18)

BOURGUIGNON, F. and MAGNAC, T. (1990), "Labour Supply and Taxation in France", *Journal of Human Resources*, **25**, 358-389.

BOUND, J., GRILICHES, Z. and HALL, B. (1986), "Wages, Schooling and IQ of Brothers and Sisters: Do the Family Factors Differ?", *International Economic Review*, **27**, 77-105.

BROWNING, M. and MEGHIR, C. (1991), "The Effects of Male and Female Labor Supply on Commodity Demands", *Econometrica*, **59**, 925-951.

BURDETT, K. and MORTENSEN, D. T. (1978), "Labour Supply under Uncertainty", in Ehrenberg, R. G. (ed.); *Research in Labour Economics, 12.* (Greenwich, Conn.: JAI Press).

CHAMBERLAIN, G. (1982), "Multivariate Regression Models for Panel Data", *Journal of Econometrics*, **18**, 5-46.

COGAN, J. F. (1981), "Fixed Costs and Labour Supply", *Econometrica*, **49**, 945-963.

CRAGG, J. (1983), "More Efficient Estimation in the Presence of Heteroscedasticity of Unknown Form", *Econometrica*, **51**, 751-763.

FLINN, C. and HECKMAN, J. (1982), "New Methods for Analysing Structural Models of Labour Force Dynamics", *Journal of Econometrics*, **18**, 115-168.

GOURIEROUX, C., MONFORT, A., RENAULT, E. and TROGNON, A. (1987), "Generalised Residuals", *Journal of Econometrics*, **34**, 5-32.

GOURIEROUX, C., MONFORT, A. and RENAULT, E. (1987). "Consistent M-Estimators in a Semi-Parametric Model" (mimeo).

HAM, J. C. (1982), "Estimation of a Labour Supply Model with Censoring due to Unemployment and Underemployment", *Review of Economic Studies*, **49**, 335-354.

HAM, J. C. AND HSIAO, C. (1984), "Two-stage Estimation of Structural Labour Supply Parameters using Interval Data from the 1971 Canadian Census,", *Journal of Econometrics*, **24**, 133-158.

HARTLEY, H. O. (1958), "Maximum Likelihood Estimation from Incomplete Data", *Biometrics*, **27**, 783-823.

HAUSMAN, J. A. (1980), "The Effects of Wages, Taxes and Fixed Costs in Women's Labour Force Participation", *Journal of Public Economics*, **14**, 161-194.

HECKMAN, J. J. (1974), "Effects of Child Care Programs on Women's Work Effort", *Journal of Political Economy*, **82**, S136-163.

HECKMAN, J. J. and MACURDY, T. E. (1980), "A Life-Cycle Model of Female Labour Supply", *Review of Economic Studies*, **47**, 47-74.

JORGENSON, D. W., LAU, L. J. and STOKER, T. M. (1982), "The Transcendental Logarithmic Model of Aggregate Consumer Behaviour", in *Advances in Econometrics, Vol. 1.* (Greenwich, Conn.: JAI Press).

LEE, L. F. (1982), "Some Approaches to the Correction of Selectivity Bias", *Review of Economic Studies*, **49**, 355-372.

MADDALA, G. S. (1971), "The Likelihood Approach to Pooling Cross Section and Time Series Data", *Econometrica*, **39**, 939-953.

MADDALA, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics* (Cambridge: Cambridge University Press).

MROZ, T. A. (1987), "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions", *Econometrica*, **55**, 765-799.

NELSON, F. D. and OLSON, L. (1978), "Specification and Estimation of a Simultaneous equation Model with Limited Dependent Variables", *International Economic Review*, **19**, 695-709.

PAGAN, A. (1986), "Two-Stage and Related Estimators and Their Applications", *Review of Economic Studies*, **53**, 517-538.

POLLAK, R. A. (1971), "Conditional Demand Functions and the Implications of Separability", *Southern Economic Journal*, **37**, 423-433.

SMITH, R. J. and BLUNDELL, R. W. (1986), "An Exogeneity Test for the Simultaneous Equation Tobit Model", *Econometrica*, **54**, 679-685.

STERN, N. (1986), "On the Specification of the Labour Supply Functions", in Blundell, R. W. and Walker, I. (eds.), *Unemployment, Search and Labour Supply* (Cambridge: Cambridge University Press).

WEBER. G. (1990), "Earnings-Related Borrowing Restrictions: Empirical Evidence from a Pseudo Panel for the UK (IFS Working Paper W90/15).

WHITE, H. (1982), "Instrumental Variables Regression with Independent Observations, *Econometrica*, **50**, 483-500.