Tilburg University

The joint estimation of a non-linear labour supply function and a wage equation using simulated response probabilities

Bloemen, H.G.; Kapteyn, A.J.

Publication date:
1992

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Bloemen, H. G., & Kapteyn, A. J. (1992). *The joint estimation of a non-linear labour supply function and a wage equation using simulated response probabilities*. (CentER Discussion Paper; Vol. 1992-29). Unknown Publisher.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

CBM
R

8414

1992

1992

NP 29

entER

for

economic Research

Discussion paper

No. 9229

**THE JOINT ESTIMATION OF A NON-LINEAR LABOUR
SUPPLY FUNCTION AND A WAGE EQUATION USING
SIMULATED RESPONSE PROBABILITIES**

by Hans G. Bloemen
and Arie Kapteyn

September 1992

ISSN 0924-7815

THE JOINT ESTIMATION OF A NON-LINEAR LABOUR SUPPLY
FUNCTION
AND A WAGE EQUATION USING SIMULATED RESPONSE
PROBABILITIES

by

Hans G. Bloemen and Arie Kapteyn ¹

CentER for Economic Research
Tilburg University
P.O. BOX 90153
5000 LE Tilburg
The Netherlands

August 1992

Abstract

When applying maximum likelihood estimation in jointly estimating a labour supply function and a wage equation, it may be practically impossible, both analytically and numerically, to calculate the required response probabilities especially if the model is non-linear. In practice therefore, usually simplifying assumptions are made, which enhance the tractability of a model, but at the same time undermine its consistency with economic theory. In the present paper we are proposing methods of estimation for mixed discrete-continuous types of models in which the response probabilities are replaced by simulators. To illustrate the methods in practice, we consider a rather simple two-equation labour supply model. The advantage of this is that we can compare the simulation methods with traditional estimators that would not be feasible in a more complicated setting. The various estimation methods are each first tried out on artificial data, to investigate their performance. After that, they are applied to the empirical model. The estimation methods perform well on artificial data although slightly less than ML, but much better than simulated ML. In the empirical application the various estimation methods appear to yield comparable estimates. However, computation of elasticities reveals substantial differences.

¹The authors thank the Organisatie voor Strategisch Arbeidsmarktonderzoek (OSA) for kindly providing the data. Thanks are due to Arthur van Soest for his help and to Peter Kooreman for his useful comments

1 Introduction

This paper investigates the applicability of estimation methods for labour supply models which make use of simulators for the response probabilities. The methods of estimation usually applied, like maximum likelihood and the method of moments, make use of the probabilities of individuals participating in the labour force. If the labour supply model is non-linear and if one wants to incorporate the tax and social security system, thereby assuming that the budget constraints of the individuals are non-convex, the calculation of the participation probabilities may be impossible and there is possibly no other way to estimate the model than making use of simulated moments types of estimators.

McFadden (1989) presents a method of simulated moments estimator for the multinomial response model. The attractiveness of the method is that the number of replications which is used to simulate the response probabilities can be kept fixed to any positive integer without destroying the consistency property of the estimator. In a short time an extensive literature has blossomed in which this approach has been extended and refined. The emphasis has been on computational accuracy and speed in the evaluation of multi-dimensional probabilities, often under normality or closely related assumptions and linearity. See Hajivassiliou (1992) for an overview.

The method of simulated moments estimator, however, is in its simplest form only suitable for discrete response data, whereas in labour supply models the data are usually of a mixed discrete-continuous nature, giving information on whether or not individuals are working and if so, how many hours. Furthermore, any realistic utility consistent model will entail non-linearities and non-normality, so that the various refinements mentioned will not be applicable. Therefore, we want to set up estimation by simulation methods for the mixed discrete-continuous type of model that one typically finds in labour supply analysis.

Different routes can be followed. The most straightforward way is to replace the response probabilities in the likelihood function by simulators, thereby simulating the likelihood function. It can be inferred from Gouriéroux and Montfort (1989) that this method of simulated maximum likelihood (SML) is not consistent for an arbitrary, fixed number of replications to simulate the response probabilities. Lerman and Manski (1981) show that a similar method for the multinomial response model may require huge numbers of replications. To circumvent this problem an alternative is to use the method of simulated scores (MSS) which is based on the simulation of the vector of scores of the log-likelihood function. As in McFadden (1989), point of departure is the property of the likelihood function that under weak regularity conditions the expectation of the score vector equals zero at the true parameter value. This score vector will be replaced by a simulated score vector. An estimator can be obtained by minimizing the length of the simulated score. The score vector will be simulated in such a way that the property of having a zero expectation at the true parameter value carries over to the simulated score vector. There is no unique way to achieve this and therefore we will propose and compare three different methods of estimation. The method of simulated scores is also used by Hajivassiliou (1989).

The methods of estimation will be applied to the joint estimation of a labour supply

function, non-linear in the wage rate, and a wage equation. In this application, we assume a linear budget constraint. This is a rather simple model and in fact it can be estimated by maximum likelihood, using numerical integration. The main purpose of the application is to gain insight in the practical properties of the MSS methods. Hence we have chosen a model simple enough so that ML is feasible and we can compare the performance of the MSS estimators with ML. In a companion paper (Bloemen and Kapteyn (1992)) one of the MSS- estimators is applied to a much more complicated model with random preferences and non-convex budget constraints. In that model ML is not feasible. We will present Monte Carlo results as well as real data estimates. In the Monte Carlo study different MSS estimators are compared with each other as well as with ML and SML with a limited number of replications.

The order of presentation is as follows: In the next section we set out the basic model where labour supply is a (possibly nonlinear) function of the wage rate and nonlabour income. Errors in the wage equation are additive and their nature remains unspecified. In section 3 different simulators for the score vector are proposed, each of them generating an alternative method of estimation. Also, attention will be paid to the statistical properties of the methods. In Section 4 the properties of these estimators and of ML are first investigated by means of some Monte Carlo experiments. Next, the estimators are applied to the analysis of labour supply of Dutch females.

The general finding is that the MSS estimators perform quite well, though of course slightly below ML, whereas simulated ML may perform poorly. We conclude that the MSS estimators proposed present viable routes for the estimation of utility consistent labour supply models.

2 The basic model

Our point of departure is a two-equation model consisting of a labour supply equation and a log-wage equation.

$$h_n^* = h(w_n, \mu_n; \beta) + \epsilon_n \quad (2.1)$$

$$\log(w_n) = w(x_n, \eta) + u_n \quad (2.2)$$

h_n^* is observed and equals the number of hours worked by individual n if the n -th individual is working; h_n^* is unobserved if the n -th individual is non-working,

$$h_n = h_n^* \text{ if } h_n^* > 0 \quad (2.3)$$

$$h_n = 0 \text{ if } h_n^* \leq 0 \quad (2.4)$$

where h_n is the actual number of hours worked by individual n , w_n is the after-tax real wage rate which will be unobserved for a non-working individual, μ_n is non-labour income, x_n is a vector of observable characteristics of individual n , β is a parameter vector with dimension l , η is a vector of parameters with dimension q , ϵ_n and u_n are random disturbances with expectation 0 and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_\epsilon^2 & \sigma_{\epsilon u} \\ \sigma_{\epsilon u} & \sigma_u^2 \end{pmatrix}$$

and joint probability density function $f(\epsilon_n, u_n)$, independent across observations. For ease of notation we introduce the dummy variable d_n with

$$d_n = 1 \text{ if } h_n^* \leq 0 \quad (2.5)$$

$$d_n = 0 \text{ if } h_n^* > 0 \quad (2.6)$$

We start by deriving the joint probability density function of h_n and w_n given x_n and μ_n . First, an expression for the joint density of h_n^* and w_n has to be found. Using the 1-1 transformations $\epsilon_n = h_n^* - h(w_n, \mu_n; \beta)$ and $u_n = \log(w_n) - w(x_n; \eta)$, we can employ the joint density of ϵ_n and u_n to get the density function $g^*(h_n^*, w_n)$ of h_n^* and w_n . The Jacobian of the transformation is

$$\begin{vmatrix} \frac{\partial \epsilon_n}{\partial h_n^*} & \frac{\partial u_n}{\partial h_n^*} \\ \frac{\partial \epsilon_n}{\partial w_n} & \frac{\partial u_n}{\partial w_n} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\frac{\partial h}{\partial w_n} & \frac{1}{w_n} \end{vmatrix} = \frac{1}{w_n}$$

So

$$g^*(h_n^*, w_n) = f(h_n^* - h(w_n, \mu_n; \beta), \log(w_n) - w(x_n; \eta)) \frac{1}{w_n} \quad (2.7)$$

$$-\infty < h_n^* < \infty$$

$$0 < w_n < \infty$$

From this we can derive the mixed discrete-continuous probability density function of h_n and w_n , $g(h_n, w_n | x_n, \mu_n, \theta)$ where θ contains the parameters of β , η and the upper triangular or, equivalently, the lower triangular elements of Σ .

$$g(h_n, w_n | x_n, \mu_n, \theta) = \begin{cases} P(h_n^* \leq 0 | x_n, \mu_n, \theta) & \text{if } h_n = 0 \\ g^*(h_n, w_n | x_n, \mu_n, \theta) & \text{if } h_n > 0, 0 < w_n < \infty \end{cases}$$

where

$$P(h_n^* \leq 0 | x_n, \mu_n, \theta) = \int_0^\infty \int_{-\infty}^0 g^*(h, w | x_n, \mu_n, \theta) dh dw \quad (2.8)$$

For ease of notation this probability will be denoted by $P_n(\theta)$ or by P_n . The wage w is integrated out because for non-working individuals we have no observations on the wage rate. We shall denote the probability of working $1 - P_n(\theta)$ by $\bar{P}_n(\theta)$ or simply by $\bar{P}_n$. We assume that our sample is ordered in such a way that the observations 1 to N_1 refer to non-working individuals and the observations $N_1 + 1$ to N are working individuals.

We now formulate the log-likelihood function of the model.

$$L(\theta | x_n, \mu_n, w_n, h_n, n = 1, \dots, N) = \sum_{n=1}^{N_1} \ln P_n(\theta) + \sum_{n=N_1+1}^N \ln g^*(h_n, w_n | x_n, \mu_n, \theta) \quad (2.9)$$

This is differentiated with respect to θ to derive the first order conditions for a maximum.

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{n=1}^{N_1} \frac{\partial \ln P_n(\theta)}{\partial \theta} + \sum_{n=N_1+1}^N \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} \quad (2.10)$$

$$\frac{\partial L(\hat{\theta}_{ML})}{\partial \theta} = 0 \quad (2.11)$$

where $\hat{\theta}_{ML}$ is the maximum likelihood estimator of θ .

Alternatively, we can rewrite the derivative of the log-likelihood function as

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{n=1}^N \left[d_n \frac{\partial \ln P_n(\theta)}{\partial \theta} + (1 - d_n) \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} \right] \quad (2.12)$$

where d_n is the dummy variable introduced above.

Let θ_0 be the true parameter value. It is easy to show that if the supports of h_n and w_n do not depend on θ ,

$$E \left(\frac{\partial L(\theta_0)}{\partial \theta} \right) = 0 \quad (2.13)$$

which is the result of the fact that the expectation of the derivative of the log-density function with respect to θ at the true parameter value equals zero. In fact, the first order derivative of the log-likelihood function divided by the sample size can be looked upon as a moment estimator of

$$E \left(\frac{\partial \ln g(h, w | x, \mu, \theta)}{\partial \theta} \right) \quad (2.14)$$

evaluated at the parameter value θ .

A procedure which is often followed in estimating this type of model is a two-step procedure. First, the wage equation is estimated using data on working individuals. The resulting estimates of the parameter vector η and the characteristics x of the non-working individuals are used to construct a proxy for the wage variables of the non-working individuals. Second, this constructed data-set is used to estimate the labour supply function, using Tobit-like methods. This method will in general yield inconsistent estimates,

particularly for non-linear labour-supply functions. A correct procedure would be to estimate the model simultaneously. A drawback of simultaneous estimation of the model is the difficulty in calculating the response probabilities analytically whenever the model is non-linear, whereas numerical approximation can be expensive and time-consuming. Our purpose is to develop estimation methods for models of the type (1.1)-(1.2) that allow for simultaneous estimation also if the function h is quite complicated (as for instance in the case, where h represents the outcome of utility maximization under a non-linear and non-convex budget constraint). To this end we make use of simulators for the response probabilities like the ones proposed by McFadden (1989). However, unlike his paper, we shall not restrict ourselves to the discrete response model. We will employ a discrete-continuous type of estimation method. The property (2.13) will be made use of to develop a method of simulated moments type of estimation. We want to replace the response probabilities in (2.12) by simulators and we want to do that in such a way that property (2.13) carries over to the simulated score vector. Then an estimator can be found by minimizing the length of the simulated score vector.

3 Estimation

Three ways of simulating the score of the log-likelihood function are considered. The first method replaces the discrete part of the score by an expression with an instrument matrix and a simulator for the response probability of non-working individuals. The disadvantage of this method is that the consistency of the estimator will depend on the choice of the matrix of instruments. This is due to the fact that the expectation of the simulated score, evaluated at the true parameter value θ_0 , does not equal zero, unless a specific form for the matrix of instruments is chosen which makes use of a consistent estimator for θ_0 . So in order to get a consistent estimator, we need a consistent estimator obtained from a different estimation procedure. Therefore, the first method is only useful to increase efficiency of the first round estimator obtained by one of the next two methods. The second method of estimation also uses a matrix of instruments. The estimator will be consistent, irrespective of the choice of the instruments. To simulate the score, simulators of the response probabilities and their derivatives are needed for each individual, both non-working and working. A second estimation round can be performed to increase the efficiency of the estimators, using an updated version of the matrix of instruments. This method directly extends Mc Fadden's (1989) estimation method for the discrete response model, by adding a continuous component to his objective function. The third method does not rely on a matrix of instruments. Only simulators of the derivatives of the response probabilities are required.

3.1 Method 1

We rewrite the first order derivative of the log-likelihood function in the following way.

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{n=1}^N \left[d_n Z_n (1 - P_n) + (1 - d_n) \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} \right] \quad (3.1)$$

$$\text{where } Z_n := \frac{\frac{\partial P_n}{\partial \theta}}{P_n(1 - P_n)} \quad (3.2)$$

Now the vector Z_n is replaced by an arbitrary vector of instruments $\bar{Z}_n$, which does not depend on θ . The resulting expression is

$$\frac{\partial \bar{L}^1(\theta)}{\partial \theta} = \sum_{n=1}^N \left[d_n \bar{Z}_n (1 - P_n) + (1 - d_n) \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} \right] \quad (3.3)$$

where the superscript 1 refers to the number of the method. Calculating the expectation of expression (3.3), conditional on $\bar{Z}_n$, evaluated at the true parameter value θ_0 yields

$$E \left(\frac{\partial \bar{L}^1}{\partial \theta} | \bar{Z}_n \right) = \sum_{n=1}^N \left[P_n \bar{Z}_n (1 - P_n) + \frac{\partial \bar{P}_n}{\partial \theta} \right] \quad (3.4)$$

(with $\bar{P}_n = 1 - P_n$) which in general doesn't equal zero. However, if we construct $\bar{Z}_n$ in such a way that

$$plim(Z_n) = \frac{\frac{\partial P_n}{\partial \theta}}{P_n(1 - P_n)} \text{ at } \theta_0 \quad (3.5)$$

we have that at the true parameter value θ_0 the simulated score has asymptotic mean equal to zero.

The response probability $P_n(\theta)$ in expression (3.3) is replaced by a frequency simulator or by a so called smooth unbiased simulator $k_n(\theta, v_R^*)$ where v_R^* is a vector of R drawings from some distribution which does not depend on θ . The simulator is unbiased if it has the property

$$E(k_n(\theta, v_R^*)) = P_n(\theta) \quad (3.6)$$

A smooth unbiased simulator can be constructed in the same way as in Monte Carlo importance sampling, see e.g. Hammersley and Handscomb (1979). Take a density function $\gamma(h, w)$ with support coinciding with the bounds of the integrals in expression (2.8) for P_n . The response probability can be rewritten as:

$$P_n(\theta) = \int_0^\infty \int_{-\infty}^0 \tau(h, w | x_n, \mu_n, \theta) \gamma(h, w) dh dw \quad (3.7)$$

with

$$\tau(h, w | x_n, \mu_n, \theta) = \frac{g^*(h, w | x_n, \mu_n, \theta)}{\gamma(h, w)} \quad (3.8)$$

For every n , R random vectors $(h_{(r,n)}, w_{(r,n)})$ are drawn from the density function $\gamma(h, w)$, independently across observations and not depending on the parameter vector θ . By averaging over the drawings a simulator is obtained:

$$k_n(\theta, v_R^*) = \frac{1}{R} \sum_{r=1}^R \tau(h_{(r,n)}, w_{(r,n)} | x_n, \mu_n, \theta) \quad (3.9)$$

where v_R^* consists of the drawings $(h_{(r,n)}, w_{(r,n)})$. The function $\tau(., . | ., ., .)$ is the so called weight function which corrects for the fact that we are drawing random numbers from the distribution with density function $\gamma(., .)$ instead of the true distribution with density function $g^*(., . | ., ., .)$. If $\gamma(., .)$ and $g^*(., . | ., ., .)$ coincide the weight function is identically equal to 1. In our application, described in the next section, we will actually use a slightly different way of simulating the response probabilities by exploiting the normality assumptions and the assumed linearity of the log-wage equation. Therefore, we don't have to choose a density function $\gamma(h, w)$. However, in more complicated applications, as in Bloemen and Kapteyn (1992), weight functions are necessary.

As indicated by McFadden (1989), a simulator for the derivatives of P_n can be constructed in a similar way. An unbiased simulator $m(\theta, v_R^*)$ of the derivatives of P_n is

$$m_n(\theta, v_R^*) = \frac{1}{R} \sum_{r=1}^R \frac{\partial \tau(h_{(r,n)}, w_{(r,n)} | x_n, \mu_n, \theta)}{\partial \theta} \quad (3.10)$$

For the simulation of (3.3) we only need k_n . The simulated score is:

$$K_R^1(\theta) = \sum_{n=1}^N \left[d_n \bar{Z}_n (1 - k_n(\theta, v_R^*)) + (1 - d_n) \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} \right] \quad (3.11)$$

The estimation procedure now becomes: Minimize the length of the simulated score vector:

$$\min_{\theta} s_N(\theta) \quad \text{where} \quad (3.12)$$

$$s_N(\theta) = K_R^1(\theta)' K_R^1(\theta) \quad (3.13)$$

where the matrix of instruments has to be based on the following formula, evaluated in the true parameter point (or at least at a consistent estimate of it):

$$\bar{Z}_n = \frac{m_n(\theta, v_{R_x}^{**})}{k_n(\theta, v_{R_x}^{**})(1 - k_n(\theta, v_{R_x}^{**}))} \quad (3.14)$$

where $m_n(\theta, v_{R_x})$ is an unbiased simulator of the derivatives of $P_n(\theta)$, and $v_{R_x}^{**}$ are drawn independently of the v_R^* which are used in minimizing $s_N(\theta)$. The number of drawings to simulate the vector of instruments is denoted by R_x to indicate that it is not necessarily equal to R , the number of drawings used to construct $k_n(\theta, v_R^*)$ in (3.11). The reason for presenting this method of estimation is that the asymptotic variance of this method is lower than the asymptotic variance of the methods presented in the next two subsections, which will be explained in section 3.4. Therefore, a two-step procedure could be followed: First, obtain a consistent estimate by applying one of the other estimation methods and second, use the estimates to construct the vector of instruments in (3.14) and apply method 1 to increase the efficiency of the estimates.

We are interested in the error we make by replacing the score vector by a simulator. Therefore, the simulated score in (3.11) is rewritten as the sum of the true score in (3.1) and a simulation residual.

$$K_R^1(\theta) = \frac{\partial L}{\partial \theta} + RES_1 \quad (3.15)$$

where

$$RES_1 = \sum_{n=1}^N d_n Z_n (P_n - k_n) + \sum_{n=1}^N d_n (\bar{Z}_n - Z_n) (1 - k_n) \quad (3.16)$$

For ease of notation the arguments are omitted. The first term of (3.16) can be rewritten as

$$\frac{1}{R} \sum_{n=1}^N \sum_{r=1}^R d_n Z_n (P_n - k_{nr}) \quad (3.17)$$

where k_{nr} is the r -th term of k_n . By increasing the number of drawings R this term will tend to zero. In the second term, the vector $\bar{Z}_n$ appears which is a non-linear function of the simulators. Because the two factors in this term are independent by construction, we concentrate on $\bar{Z}_n - Z_n$. If $\bar{Z}_n$ is constructed in a proper way, i.e. by using a consistent estimate for θ , then

$$\text{plim}_{R_x \rightarrow \infty} (\bar{Z}_n - Z_n) = 0 \quad (3.18)$$

so by taking the number of drawings to construct the vector of instruments large enough the second term also can be made arbitrarily small. The variance of the simulation residual determines the loss of efficiency caused by using the simulated score instead of the true score vector. In the next subsection we discuss how the variance of the simulation residual can be influenced by the number of drawings R .

Since the derivation of the asymptotic distribution is equivalent for the three estimators we will treat the asymptotic properties of the estimators at the end of this section.

3.2 Method 2

In order to obtain the second method of estimation we rewrite the score of the log-likelihood function as

$$\frac{\partial L}{\partial \theta} = \sum_{n=1}^N \left\{ Z_n(d_n - P_n) + (1 - d_n) \left[\frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} - \frac{\partial \ln \bar{P}_n}{\partial \theta} \right] \right\} \quad (3.19)$$

with Z_n defined as above. The first component of this expression equals the score of the log-likelihood of the binary response model. If we replace the vector Z_n by an arbitrary vector of instruments $\bar{Z}_n$, independent of θ , the expectation of the resulting expression, conditional on $\bar{Z}_n$, equals zero at the true parameter value θ_0 .

To simulate $P_n(\theta)$ we use the simulator $k_n(\theta, v_R^*)$ defined above. The problem is how to simulate $\frac{\partial \ln \bar{P}_n}{\partial \theta}$. To see why this is a problem we rewrite this expression as

$$\frac{\partial \ln \bar{P}_n}{\partial \theta} = \frac{1}{\bar{P}_n} \frac{\partial \bar{P}_n}{\partial \theta} \quad (3.20)$$

It is not difficult to construct unbiased simulators of $\bar{P}_n$ and $\frac{\partial \bar{P}_n}{\partial \theta}$, as we showed before. However, if we use their simulators $\bar{k}_n(\theta, v_R^*)$ and $\bar{m}_n(\theta, v_R^*)$ to simulate $\frac{\partial \ln \bar{P}_n}{\partial \theta}$, we don't get an unbiased simulator.

$$E \left(\frac{\bar{m}_n(\theta, v_R^*)}{\bar{k}_n(\theta, v_R^*)} \right) \neq \frac{\partial \ln \bar{P}_n}{\partial \theta} \quad (3.21)$$

So the expectation of the simulated score evaluated at θ_0 won't equal zero. It is not clear how to get an unbiased simulator of $\frac{\partial \ln \bar{P}_n}{\partial \theta}$. In order to solve this problem, we will simulate $(1 - d_n) \frac{\partial \ln \bar{P}_n}{\partial \theta}$ instead of $\frac{\partial \ln \bar{P}_n}{\partial \theta}$.

$$E \left[(1 - d_n) \frac{\partial \ln \bar{P}_n}{\partial \theta} \right] = \frac{\partial \bar{P}_n}{\partial \theta} \quad (3.22)$$

We now replace $(1 - d_n) \frac{\partial \ln \bar{P}_n}{\partial \theta}$ by $\frac{\partial \bar{P}_n}{\partial \theta}$. As a result, the original score vector is replaced by

$$\frac{\partial \bar{L}^2}{\partial \theta} = \sum_{n=1}^N \left[\bar{Z}_n(d_n - P_n) + (1 - d_n) \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} - \frac{\partial \bar{P}_n}{\partial \theta} \right] \quad (3.23)$$

Inserting the simulators for the response probabilities and their derivatives in this expression gives the simulated score:

$$K_R^2(\theta) = \sum_{n=1}^N \left[\bar{Z}_n(d_n - k_n(\theta, v_R^*)) + (1 - d_n) \frac{\partial \ln g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} - \bar{m}_n(\theta, v_R^*) \right] \quad (3.24)$$

The estimation procedure becomes: Choose instrument vectors $\bar{Z}_n$ and minimize the length of the simulated score.

$$\min_{\theta} K_R^2(\theta)' K_R^2(\theta) \quad (3.25)$$

We can increase the efficiency of the estimator by rerunning the procedure using the updated vectors of instruments, as described in the preceding subsection.

Again we are interested in the simulation residual. First, (3.19) with Z_n replaced by $\bar{Z}_n$ is compared with (3.24). Then the following residual is obtained:

$$\frac{1}{R} \sum_{n=1}^N \sum_{r=1}^R \left[\bar{Z}_n(P_n - k_{nr}) - \left\{ \bar{m}_{nr} - (1 - d_n) \frac{\partial \ln \bar{P}_n}{\partial \theta} \right\} \right] \quad (3.26)$$

This is the simulation residual corresponding to the comparison of the moments estimator (Z_n replaced by $\bar{Z}_n$) with the simulated moments estimator. The dummy variable can be rewritten as

$$\begin{aligned} d_n &= P_n + \nu_n \\ \text{with } E(\nu_n) &= 0 \\ \text{and } Var(\nu_n) &= P_n(1 - P_n) \end{aligned} \quad (3.27)$$

Inserting this in the residual gives:

$$\frac{1}{R} \sum_{n=1}^N \sum_{r=1}^R \left[\bar{Z}_n(P_n - k_{nr}) - \left\{ \bar{m}_{nr} - \frac{\partial \bar{P}_n}{\partial \theta} \right\} \right] - \sum_{n=1}^N \nu_n \frac{\partial \ln \bar{P}_n}{\partial \theta} \quad (3.28)$$

The variance of the first term of (3.28) can be reduced by increasing the number of drawings R . Suppose that $Var[\bar{Z}_n(P_n - k_{nr}) - \{\bar{m}_{nr} - \frac{\partial \bar{P}_n}{\partial \theta}\}] = \Xi_n$, conditional on the instruments $\bar{Z}_n$. This variance does not depend on r because the drawings are i.i.d. Then the variance of the first term is $\frac{1}{R} \sum_{n=1}^N \Xi_n$. With fixed N , increasing R to infinity results in reducing this variance to zero. The second term is the error which is caused by the fact that $(1 - d_n) \frac{\partial \ln \bar{P}_n}{\partial \theta}$ is simulated by a simulator for $\frac{\partial \bar{P}_n}{\partial \theta}$. The expectation of this term equals zero, whereas the variance equals $\sum_{n=1}^N \bar{P}_n P_n \frac{\partial \ln \bar{P}_n}{\partial \theta} \frac{\partial \ln \bar{P}_n}{\partial \theta'}$. This term of the simulation residual cannot be influenced by the number of drawings. Therefore, this term leads to inefficiency, also for large R .

To compare the efficiency of the method of simulated moments estimator to the maximum likelihood estimator, also the term involving the difference between $\bar{Z}_n$ and Z_n has to be taken into account. The simulation residual then becomes:

$$RES_2 = \frac{1}{R} \sum_{n=1}^N \sum_{r=1}^R \left[Z_n(P_n - k_{nr}) - \left\{ \bar{m}_{nr} - \frac{\partial \bar{P}_n}{\partial \theta} \right\} \right] + \sum_{n=1}^N (\bar{Z}_n - Z_n)(d_n - k_n) - \sum_{n=1}^N \nu_n \frac{\partial \ln \bar{P}_n}{\partial \theta} \quad (3.29)$$

Here the same observations can be made as for method 1.

3.3 Method 3

As opposed to the first two methods of estimation, we won't rewrite the score of the log-likelihood function in a form involving instrument vectors Z_n . Instead, we immediately replace the discrete part of the score by a simulator. In analogy with the preceding method, we replace $d_n \frac{\partial \ln \bar{P}_n}{\partial \theta}$ by $\frac{\partial \bar{P}_n}{\partial \theta}$. This yields the expression

$$\frac{\partial \bar{L}^3}{\partial \theta} = \sum_{n=1}^N \left[\frac{\partial P_n}{\partial \theta} + (1 - d_n) \frac{\partial g^*(h_n, w_n | z_n, \mu_n, \theta)}{\partial \theta} \right] \quad (3.30)$$

The derivative of the probability is replaced by its simulator $m_n(\theta, v_R^*)$, which leads to the following expression for the simulated score:

$$K_R^3(\theta) = \sum_{n=1}^N \left[m_n(\theta, v_R^*) + (1 - d_n) \frac{\partial g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} \right] \quad (3.31)$$

The estimation procedure becomes

$$\min_{\theta} K_R^3(\theta)' K^3(\theta) \quad (3.32)$$

We compare this method of simulating the score with the method described in the preceding subsection. Because $\bar{P}_n = 1 - P_n$, we find that

$$\frac{\partial \bar{P}_n}{\partial \theta} = - \frac{\partial P_n}{\partial \theta} \quad (3.33)$$

The same relation holds for their simulators:

$$m_n(\theta, v_R^*) = -\bar{m}_n(\theta, v_R^*) \quad (3.34)$$

Inserting this in the simulated score, we obtain

$$K_R^3 = \sum_{n=1}^N \left[(1 - d_n) \frac{\partial g^*(h_n, w_n | x_n, \mu_n, \theta)}{\partial \theta} - \bar{m}_n(\theta, v_R^*) \right] \quad (3.35)$$

This is exactly the second component of the simulated score of method 2. Obviously, by replacing the derivative of the log-response probability by a simulator for the derivative of the response probability, some information is lost and $\frac{\partial L_n^2}{\partial \theta}$ and the second component of $\frac{\partial L_n^2}{\partial \theta}$ become indistinguishable. Method 2 has the intuitively appealing property that it includes the minimization of the distance between the binary response indicator d_n and its theoretical expectation P_n . The simulation residual for method 3 can be found in the same way as for method 2 and is given by

$$RES_3 = \frac{1}{R} \sum_{n=1}^N \sum_{r=1}^R \left[m_{nr} - \frac{\partial P_n}{\partial \theta} \right] - \sum_{n=1}^N \nu_n \frac{\partial \ln \bar{P}_n}{\partial \theta} \quad (3.36)$$

The residual contains the same error which is not influenced by the number of drawings R as method 2.

3.4 Asymptotic distribution of the estimators

In the preceding subsections we presented three ways to simulate the first order derivatives of the log-likelihood function. The simulated score vectors of methods 2 and 3 satisfy the property that their expectation, evaluated at the true parameter vector θ_0 , equals zero, whereas the probability limit of the simulated score of method 1 equals zero at the true parameter value θ_0 if the vectors of instruments are constructed in a proper way. Therefore, the length of the expectation of the simulated score is minimized at the true parameter value θ_0 . It is intuitively clear that if we minimize the length of the

simulated score, the resulting parameter vector $\hat{\theta}_R^i$ at which minimization takes place, will converge to the true parameter value θ_0 , or, equivalently, $\hat{\theta}_R^i$ will be a consistent estimator of θ_0 .

We assume that

$$\frac{1}{\sqrt{N}} K_R^i(\theta_0) \xrightarrow{\text{asy}} N(0, V_{R0}^i), i = 1, \dots, 3 \quad (3.37)$$

with V_{R0}^i some positive definite symmetric matrix. Below, we explain that this assumption can be justified on the basis of central limit arguments. Using this assumption and the consistency of $\hat{\theta}_R^i$, apart from usual regularity assumptions on the parameter space and the like, it is possible to show (see, for example, Pakes and Pollard (1989)) that

$$\sqrt{N}(\hat{\theta}_R^i - \theta_0) \xrightarrow{\text{asy}} N(0, (\Gamma^{i'} \Gamma^i)^{-1} \Gamma^{i'} V_{R0}^i \Gamma^i (\Gamma^{i'} \Gamma^i)^{-1}) \quad (3.38)$$

$$\text{where } \Gamma^{i'} = \text{plim} \frac{1}{N} \frac{\partial (\frac{\partial L^i(\theta_0)}{\partial \theta'})}{\partial \theta} \quad (3.39)$$

We now will comment on the assumed normality of the simulated score. The expectation of the simulated score, conditional on the instruments, equals zero at θ_0 for $i = 2, 3$. It was shown that the simulated score could be written as the sum of the true score vector and a simulation residual. It is well-known that under general conditions the distribution of the true score vector divided by the square root of the sample size converges to a normal distribution. The assumed independence across observations and the fact that the simulators are constructed using independent random drawings can be used to apply the Lindeberg-Feller central limit theorem to prove the normality of the simulation residuals (conditional on the instruments) divided by the square root of the sample size. The distribution of the simulated score then converges to the sum of two normal distributions which is in turn a normal distribution. We can do the same for method 1, but not without recalling that the instruments have to be constructed such that (3.5) is satisfied.

Using the expression of the asymptotic covariance matrix and the results of the analysis of the simulation residuals in the preceding subsections it is possible to analyse the efficiency of the estimators by comparing the asymptotic covariance matrices of the simulation estimators with the asymptotic covariance matrix of the maximum likelihood estimator. It is a well known result that

$$\sqrt{N}(\hat{\theta}_{ML} - \theta_0) \xrightarrow{\text{asy}} N(0, \Omega_{ML}) \quad (3.40)$$

$$\text{where } \Omega_{ML} = B^{-1} \quad (3.41)$$

$$B = -\text{plim} \frac{1}{N} \frac{\partial^2 L(\theta_0)}{\partial \theta \partial \theta'} \quad (3.42)$$

To make clear the relation with the asymptotic covariance matrix of the simulation estimators we rewrite Ω_{ML} as

$$\Omega_{ML} = (\Gamma'_{ML} \Gamma_{ML})^{-1} \Gamma'_{ML} V_{ML} \Gamma_{ML} (\Gamma'_{ML} \Gamma_{ML})^{-1} \quad (3.43)$$

where

$$\Gamma'_{ML} = \text{plim} \frac{1}{N} \frac{\partial (\frac{\partial L(\theta_0)}{\partial \theta'})}{\partial \theta} = -B \quad (3.44)$$

which is the equivalent of (3.39), and

$$V_{ML} = plim \frac{1}{N} \sum_{n=1}^N \frac{\partial L_n(\theta_0)}{\partial \theta} \frac{\partial L_n(\theta_0)}{\partial \theta'} \quad (3.45)$$

Recall that $V_{ML} = B$. For method 1, if the instruments are constructed properly and if the number of drawings to construct the instruments tend to infinity $\Gamma^{1'}$ and Γ'_{ML} are equivalent. Then the efficiency comparison reduces to comparing V_{ML} with V_R^1 for this method. The difference between these matrices is given by the covariance matrix of the simulation residuals. In section 3.1 it has been derived that this variance disappears if R tends to infinity. Therefore we can conclude that if the matrix of instruments is constructed on the basis of (3.14) and if both the number of drawings to construct this matrix and the number of drawings to simulate the response probabilities tend to infinity, the covariance matrix of the method 1 estimator and the covariance matrix of the maximum likelihood estimator are asymptotically equal. Of course, this result has only theoretical meaning because the reason why we construct these simulation estimators is to be able to keep the number of drawings fixed and small.

To examine the efficiency of method 2 it has to be noted first that because of replacement (3.22), the simulated score $K_R^2(\theta)$ does not tend to the true score if N is fixed and R tends to infinity. Therefore, we first need to establish the relation between Γ'_{ML} and $\Gamma^{2'}$. From (3.19) and (3.23) it is readily established that

$$\Gamma^{2'} - \Gamma'_{ML} = -plim \frac{1}{N} \sum_{n=1}^N \left[(\bar{Z}_n - Z_n) \frac{\partial P_n}{\partial \theta} + \frac{\partial Z'_n}{\partial \theta} (d_n - P_n) + \nu_n \frac{\partial (\frac{\partial \ln P_n}{\partial \theta'})}{\partial \theta} + \frac{1}{P} \frac{\partial P_n}{\partial \theta} \frac{\partial P_n}{\partial \theta'} \right] \quad (3.46)$$

from which only the first three terms equal zero if the instruments are constructed such that (3.5) is satisfied, i.e. according to formula (3.14) with drawings tending to infinity. From the analysis of the simulation residuals it becomes clear that if the matrix of instruments is constructed according to (3.14) with drawings tending to infinity, and if the response probabilities and their derivatives are simulated with R tending to infinity as well, the asymptotic variance of the score of the likelihood function, evaluated in a consistent estimator is exceeded by X , where

$$X = plim \left(\frac{1}{N} \sum_{n=1}^N P_n P_n \frac{\partial \ln \bar{P}_n}{\partial \theta} \frac{\partial \ln \bar{P}_n}{\partial \theta'} \right) \quad (3.47)$$

which was derived in section 3.2. The same expression can be derived for method 3.

Finally, to estimate the covariance matrix we calculate

$$\hat{\Gamma}_R^i = (\hat{\Gamma}^{i'} \hat{\Gamma}^i)^{-1} \hat{\Gamma}^{i'} \hat{V}_R^i \hat{\Gamma}^i (\hat{\Gamma}^{i'} \hat{\Gamma}^i)^{-1} \quad (3.48)$$

with

$$\hat{\Gamma}^{i'} = \frac{1}{N} \frac{\partial (\frac{\partial L^i(\hat{\theta}_R^i)}{\partial \theta'})}{\partial \theta} \quad (3.49)$$

$$\hat{V}_R^i = \frac{1}{N} \sum_{n=1}^N K_{nR}^i(\hat{\theta}_R^i) K_{nR}^i(\hat{\theta}_R^i)' \quad (3.50)$$

where the index n indicates the n -th component of the simulated score. Expression (3.50) can be calculated by simulation.

4 Monte Carlo and empirical application

We will now illustrate the properties of the various estimators by making specific assumptions about the form of the labour supply function and the log-wage equation. We then compare the estimation methods first by Monte Carlo methods and next by estimating the specification for a sample of 849 married female individuals, drawn in 1985. We'll assume that the preferences of the individuals are described by a utility function introduced by Hausman and Ruud (1984), which implies a labour supply function quadratic in the wage rate and linear in non-labour income. The wage equation is assumed to be log-linear. The disturbances of the labour supply function and the log-wage equation are assumed to be normally distributed.

The specific form of the labour supply function under the assumption of a linear budget constraint becomes:

$$h(w_n, \mu_n, \beta) = \beta_3 + \mu_n^* \beta_2 + w_n \beta_4 \quad (4.1)$$

$$\text{with } \mu_n^* = \beta_1 + \mu_n + w_n \beta_3 + \frac{1}{2} w_n^2 \beta_4 \quad (4.2)$$

Inserting the expression for μ_n^* in the labour supply function and reparameterizing gives:

$$h(w_n, \mu_n, \alpha) = \alpha_1 + \mu_n \alpha_2 + w_n \alpha_3 + \frac{1}{2} w_n^2 \alpha_4 \quad \text{with} \quad (4.3)$$

$$\alpha_1 = \beta_3 + \beta_1 \beta_2$$

$$\alpha_2 = \beta_2$$

$$\alpha_3 = \beta_2 \beta_3 + \beta_4$$

$$\alpha_4 = \beta_2 \beta_4$$

The log-wage equation becomes:

$$\log w_n = \sum_{j=1}^{11} \eta_j x_{nj} + u_n \quad (4.4)$$

where

- $x_{n1} = 1$ for all n ,
- $x_{n2} =$
log of number of persons in individual n 's family,
- $x_{n3} =$
the number of children with age below 6 of individual n ,
- $x_{n4} =$ log-age of individual n ,
- x_{n5} to x_{n8}
are dummy indicators for the level of education of individual n ,
where x_{n5} is the lowest level of education,
 x_{n6} is the next to the lowest education level etc.
For the highest education level no dummy indicator is included

- x_{n9} and x_{n10} are indicators for the type of education received, x_{n9} is a dummy indicator for non-technical and non-commercial type of education, x_{n10} is a dummy indicator for semi-technical and semi-commercial type of education.

For technical and commercial type of education no dummy indicator is included,

- $x_{n11} = x_{n4}^2$
- here w_n is the after tax hourly wage rate; labour supply will be measured in hours per week

The complete model thus reads

$$h_n^* = \alpha_1 + \mu_n \alpha_2 + w_n \alpha_3 + \frac{1}{2} w_n^2 \alpha_4 + \epsilon_n \quad (4.5)$$

$$\log w_n = \sum_{j=1}^{11} \eta_j x_{nj} + u_n \quad (4.6)$$

ϵ_n and u_n are jointly normally distributed with

$$\begin{pmatrix} \epsilon_n \\ u_n \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right) \quad (4.7)$$

$$\Sigma = \begin{pmatrix} \sigma_\epsilon^2 & \sigma_{\epsilon u} \\ \sigma_{\epsilon u} & \sigma_u^2 \end{pmatrix} \quad (4.8)$$

$$h_n = h_n^* \text{ if } h_n^* > 0 \quad (4.9)$$

$$h_n = 0 \text{ if } h_n^* \leq 0 \quad (4.10)$$

We now derive an expression for the response probabilities and their derivatives. Under the normality assumptions the joint probability density function of h_n and w_n is

$$g(h_n, w_n | x_n, \mu_n, \theta) =$$

$$P_n(\theta) \text{ if } h_n = 0 \quad (4.11)$$

$$\frac{1}{2\pi w_n} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \begin{pmatrix} h_n - h(w_n, \mu_n, \alpha) \\ \log w_n - \eta' x_n \end{pmatrix}' \Sigma^{-1} \begin{pmatrix} h_n - h(w_n, \mu_n, \alpha) \\ \log w_n - \eta' x_n \end{pmatrix} \right\} \quad (4.12)$$

$$\text{if } h_n > 0, 0 < w_n < \infty$$

where

$$P_n(\theta) = \int_{-\infty}^{\infty} \Phi \left(-\frac{h[\exp(\eta' x_n + \sigma_u v), \mu_n, \alpha] + \rho \sigma_\epsilon v}{\sigma_{\epsilon|u}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} v^2 \right\} dv \quad (4.13)$$

with

$$\rho = \frac{\sigma_{\epsilon u}}{\sigma_\epsilon \sigma_u}, \sigma_{\epsilon|u}^2 = \sigma_\epsilon^2 (1 - \rho^2)$$

and where $\Phi(\cdot)$ is the standard normal distribution function.

We have written the double integral in a form with the well-known function $\Phi(\cdot)$ and an integral over the standard normal variable v . As a result we only need to simulate the second integral over v . Because the standard normal distribution function doesn't depend on the parameters, we don't have to choose a weighting density function as described for the general case in the preceding section. To simulate the probability P_n we can simply draw from the standard normal distribution and compute the expression under the integral sign.

We obtain the following expression for the smooth unbiased simulator of P_n .

$$k_n(\theta, v_R^*) = \frac{1}{R} \sum_{r=1}^R \Phi \left(- \frac{h[\exp(\eta' x_n + \sigma_u v_{nr}^*), \mu_n, \alpha] + \rho \sigma_\epsilon v_{nr}^*}{\sigma_{\epsilon|u}} \right) \quad (4.14)$$

where the v_{nr}^* are independent random drawings from the standard normal distribution.

To simulate the derivative of P_n with respect to, say, θ_j we first write

$$\frac{\partial P_n}{\partial \theta_j} = \int_{-\infty}^{\infty} \frac{\partial \Phi \left(- \frac{h[\exp(\eta' x_n + \sigma_u v), \mu_n, \alpha] + \rho \sigma_\epsilon v}{\sigma_{\epsilon|u}} \right)}{\partial \theta_j} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} v^2 \right\} dv \quad (4.15)$$

where the derivative in the integrand can be calculated analytically. Then we simulate it by

$$m_{nj}(\theta, v_R^*) = \frac{1}{R} \sum_{r=1}^R \frac{\partial \Phi \left(- \frac{h[\exp(\eta' x_n + \sigma_u v_{nr}^*), \mu_n, \alpha] + \rho \sigma_\epsilon v_{nr}^*}{\sigma_{\epsilon|u}} \right)}{\partial \theta_j} \quad (4.16)$$

4.1 Monte Carlo results

To get an idea about the performance of the estimation methods we have performed some Monte Carlo experiments. First of all, values for the parameters in (4.5), (4.6), (4.7) and (4.8) were chosen. Next, disturbances ϵ_n and u_n , $n = 1, \dots, N$, were generated, under the assumptions (4.7), (4.8) with $\sigma_{\epsilon u} = 0$. The characteristics x_n from the sample were used to generate wages w_n , $n = 1, \dots, N$. Included are the constant term with parameter η_1 , $\log(\text{age})$ with parameter η_2 and the square of $\log(\text{age})$ with parameter η_3 . These wages were used to generate the h_n^* , $n = 1, \dots, N$. Making use of (4.9) and (4.10) and the non-labour income series from the sample, we generated the "observations" h_n .

This generated data set was used to estimate the model with different methods of estimation. First of all the model is estimated with maximum likelihood (ML) using numerical integration. The Gauss-Hermite quadrature formula is used with a number of abscissae equal to 16, see e.g. Stroud and Secrest (1966). The second method of estimation is simulated maximum likelihood (SML) with a number of drawings R equal to 10. Here the probabilities are simulated according to (4.14) and inserted in the likelihood function directly. Although this method is inconsistent for a fixed and small number of drawings, it still would be useful if the asymptotic bias were small. Finally, the model is estimated using the three MSS estimators, abbreviated as MSS1, MSS2 and

MSS3 below. Two different numbers of drawings to simulate the response probabilities are used, i.e. $R = 1$ and $R = 10$. The matrix of instruments is constructed on the basis of formula (3.13) evaluated in the true parameter point, where the number of drawings to simulate the instruments is $R_Z = 10$ for both MSS1 and MSS2. For MSS1 the model is also estimated with $R_Z = 500$ and $R = 1$.

Tables 4.1 presents the results of the Monte Carlo study. The Monte Carlo procedure was repeated 20 times, so that the numbers in the tables refer to means over 20 replications. This modest number of replications has been chosen in view of the rather heavy computational burden of non-linear optimization problems in general. For each parameter the first line presents the mean of the estimates as given by:

$$\bar{\theta}_j = \frac{1}{20} \sum_{i=1}^{20} \hat{\theta}_j^i, \quad j = 1, \dots, N \quad (4.17)$$

N is the number of parameters, whereas the subscript j stands for the j -th component of the parameter vector. The second line gives the sample standard deviation of the estimates:

$$SD_j = \sqrt{\frac{1}{20-1} \sum_{i=1}^{20} (\hat{\theta}_j^i - \bar{\theta}_j)^2}, \quad j = 1, \dots, N \quad (4.18)$$

The third line presents the relative error:

$$\text{rel. err.} = 100\% \times |\bar{\theta}_j - \theta_{0j}| / |\theta_{0j}| \quad (4.19)$$

From the table it can be seen that the ML estimator in general performs best in the sense that it has the lowest standard errors and the lowest relative errors. It is clear that SML with $R = 10$ performs very badly. It performs even worse than all of the three MSS variants with $R = 1$. Therefore, the use of MSS instead of SML is not just something which only has theoretical relevance. Comparing different numbers of drawings of the same MSS estimator it can be seen that the standard deviations decrease with the increase in number of drawings, which is consistent with the analysis of the simulation residuals in chapter 3. For MSS1 the number of drawings R_Z to simulate the matrix of instruments has been increased from 10 to 500. Recall that for this method the construction of the matrix of instruments not only affects the efficiency of the estimator but also its consistency. MSS1 with $R = 1$ and $R_Z = 500$ outperforms MSS1 with $R = 10$ and $R_Z = 10$. The standard deviations are comparable with those of MSS2 with $R = 1$ and $R_Z = 10$, but they are not lower than the MSS2 standard deviations, which questions the use of MSS1. Comparing MSS2 with MSS3 it can be said that for $R = 1$ method 2 has the lower standard deviations and the higher relative errors, whereas for $R = 10$ the differences are rather small, although MSS3 seems to perform slightly better, which may be due to the fact that in MSS2 theoretically an additional inefficiency is introduced by simulating the matrix of instruments with only $R_Z = 10$ drawings.

One may question the practical relevance of the experiments so far with respect to MSS1 and MSS2, since the instruments were computed at the true parameter point which of course is unknown in practice. To see how this affects results, Table 4.2 presents results for MSS1 and MSS2 with instruments based on estimates obtained with MSS3.

Moreover, Table 4.2 also presents the mean standard error of the estimates over the twenty replications so that a comparison with the sample standard deviations is possible. Comparing the table with the corresponding columns in Table 4.1, one observes that MSS1 is a bit more sensitive to the choice of instruments than MSS2, as one would expect. Actually, MSS2 is hardly affected at all by the new instruments. Although MSS1 does give somewhat different estimates now, they are not systematically worse than before, sometimes the relative error is better, and sometimes it is worse. It remains true that the small value for R_Z induces inefficiency. Finally, we observe that the estimated standard errors tend to be of a similar magnitude as the standard deviations, though with some exceptions. The standard errors tend to be a little higher than the standard deviations, hence one could take standard errors take as a somewhat conservative estimate of the inaccuracy of the estimates.

We conclude that for SML the limited number of drawings of 10 is clearly not sufficient for a reasonable performance of the estimator. All of the three MSS estimators perform acceptably even with only one drawing, although it is clear that efficiency can be improved by taking more than one drawing.

TABLE 4.1 MONTE CARLO RESULTS

parameter	θ_0	ML	SML	MSS1
			$R = 10$	$R = 1, R_Z = 10$
α_1	-8.686	-8.961	-5.996	-19.122
SD		0.586	15.689	12.004
rel. err. (%)		3.166	31.0	120.1
α_2	-0.0482	-0.0483	-0.414	-0.0503
SD		0.0146	0.653	0.0126
rel. err. (%)		0.265	760.4	4.443
α_3	3.137	3.098	-2.258	4.090
SD		0.0784	10.249	1.118
rel. err. (%)		1.259	172.0	30.4
α_4	-0.163	-0.156	-0.109	-0.210
SD		0.00850	0.0598	0.0587
rel. err. (%)		4.518	33.2	29.0
σ_ϵ^2	346.921	346.923	254.785	601.509
SD		0.0233	255.331	581.318
rel. err. (%)		0.00552	26.6	73.4
σ_u^2	1.080	1.175	0.448	1.966
SD		0.123	0.572	2.311
rel. err. (%)		8.809	58.6	82.0
η_1	-11.801	-11.860	158.243	-8.983
SD		0.0597	202.316	5.900
rel. err. (%)		0.501	1440.1	23.9
η_2	8.644	8.560	5.600	7.124
SD		0.118	6.436	3.284
rel. err. (%)		0.972	35.2	17.6
η_3	-1.199	-1.189	-3.801	-0.978
SD		0.0383	7.795	0.453
rel. err. (%)		0.844	217.0	18.4

TABLE 4.1 MONTE CARLO RESULTS (continued)

parameter	MSS1 $R = 10, R_Z = 10$	MSS1 $R = 1, R_Z = 500$	MSS2 $R = 1, R_Z = 10$	MSS2 $R = 10, R_Z = 10$
α_1	-22.589	-10.765	-2.517	-6.705
SD	11.920	6.556	6.880	3.407
rel. err. (%)	160.1	23.9	71.0	22.8
α_2	-0.0436	-0.0490	-0.0493	-0.0503
SD	0.0185	0.0136	0.0110	0.0116
rel. err. (%)	9.451	1.752	2.405	4.388
α_3	4.535	3.353	2.624	2.921
SD	1.220	0.685	0.528	0.303
rel. err. (%)	44.6	6.890	16.4	6.890
α_4	-0.221	-0.173	-0.147	-0.156
SD	0.0548	0.0300	0.0194	0.0139
rel. err. (%)	35.6	5.976	9.913	4.340
σ_e^2	366.357	349.480	347.272	347.013
SD	63.981	8.146	0.469	0.163
rel. err. (%)	5.602	0.736	0.101	0.0266
σ_u^2	7.586	2.809	0.995	1.004
SD	16.004	6.503	0.0869	0.0686
rel. err. (%)	602.4	160.1	7.896	7.066
η_1	-8.430	-8.694	-12.602	-12.022
SD	3.342	5.512	5.382	0.630
rel. err. (%)	28.6	26.3	6.783	1.871
η_2	7.299	7.100	9.215	8.859
SD	1.683	2.963	2.970	0.389
rel. err. (%)	15.6	17.9	6.609	2.482
η_3	-1.003	-0.990	-1.291	-1.243
SD	0.227	0.423	0.411	0.0689
rel. err. (%)	16.4	17.4	7.707	3.658

TABLE 4.1 MONTE CARLO RESULTS (continued)

parameter	MSS3 $R = 1$	MSS3 $R = 10$
α_1	-11.786	-8.690
SD	8.002	0.0136
rel. err. (%)	35.7	0.00420
α_2	-0.0445	-0.0506
SD	0.0193	0.0141
rel. err. (%)	7.609	5.125
α_3	3.305	3.091
SD	0.641	0.178
rel. err. (%)	5.355	1.460
α_4	-0.168	-0.161
SD	0.0230	0.0100
rel. err. (%)	3.233	0.982
σ_e^2	342.439	346.921
SD	19.954	0.000762
rel. err. (%)	1.292	0.0475×10^{-3}
σ_u^2	1.116	1.065
SD	0.155	0.0994
rel. err. (%)	3.332	1.427
η_1	-12.031	-11.775
SD	7.450	0.0955
rel. err. (%)	1.946	0.225
η_2	8.651	8.696
SD	4.092	0.167
rel. err. (%)	0.0850	0.599
η_3	-1.185	-1.216
SD	0.565	0.0544
rel. err. (%)	1.127	1.418

TABLE 4.2 ADDITIONAL MONTE CARLO RESULTS, TWO STAGES

parameter	MSS1	MSS2
	$R = 10, R_Z = 10$	$R = 10, R_Z = 10$
α_1	-19.261	-6.146
SD	11.771	4.086
Est. SE.	14.623	9.369
rel. err. (%)	121.8	29.242
α_2	-0.0560	-0.0508
SD	0.0177	0.0121
Est. SE.	0.0106	0.00989
rel. err. (%)	16.208	5.560
α_3	4.068	2.902
SD	1.073	0.366
Est. SE.	1.318	0.607
rel. err. (%)	29.678	7.490
α_4	-0.211	-0.156
SD	0.0528	0.0154
Est. SE.	0.0674	0.0231
rel. err. (%)	29.155	4.575
σ_ϵ^2	647.011	347.050
SD	460.838	0.210
Est. SE.	493.617	103.994
rel. err. (%)	86.501	0.371×10^{-3}
σ_u	5.095	1.0001
SD	12.633	0.0756
Est. SE.	2.673	0.0696
rel. err. (%)	371.8	7.335
η_1	-11.286	-11.650
SD	12.470	10.542
Est. SE.	14.975	11.945
rel. err. (%)	4.371	0.152
η_2	8.814	8.647
SD	7.301	4.928
Est. SE.	8.539	6.734
rel. err. (%)	1.970	0.299×10^{-3}
η_3	-1.224	-1.212
SD	1.020	0.843
Est. SE.	1.191	0.951
rel. err. (%)	2.119	1.098

4.2 Estimation results

We now present the estimates of the model for the real data. The model has been estimated using the three methods described in section 3 and by ML. For each method, the model was estimated under the assumption $\sigma_{eu} = 0$ with two different numbers of drawings; $R = 5$ and $R = 10$. The restriction $\sigma_{eu} = 0$ is relaxed below. The instrument matrices are constructed using $R_Z = 10$ drawings.

In table 4.3 the ML estimates are given. The estimates have the expected sign. The maximum number of hours worked occurs at a wage rate of about 11 guilders per hour, the wage rate reaches its maximum at the age of 35 years.

In table 4.4 the estimates by method 1, using $R = 5$ drawings from the standard normal distribution, and their asymptotic standard errors are presented. To circumvent problems with consistency, we used the estimates of method 3 (see table 4.9) to construct the matrix of instruments. Most of the parameter estimates have the expected sign. Non-labour income has a negative impact on labour supply. The estimate of the linear wage parameter is positive and the estimate of the squared wage parameter is negative. We can calculate that the number of hours worked reaches its maximum at a wage rate of Dfl. 13.5 per hour. Log-family size influences the log-wage rate negatively. The education dummies have the expected negative sign. Moreover, the higher the level of education, the less negative is the parameter estimate of the corresponding education dummy. From the parameter estimates of log-age and its square we can calculate that the wage rate reaches its maximum at the age of 38.

The estimates in table 4.5 are also obtained by applying method 1, but now we have used $R = 10$ drawings to construct the simulators. Again, most of the estimates have the expected sign. The maximum of the number of hours worked with respect to the wage rate is reached at $w = 13.4$. The log-wage rate reaches its maximum with respect to age at 38 years. The estimates don't differ much from the ones in table 4.4, but the standard errors are somewhat lower.

In table 4.6 we present the estimates obtained by method 2. The matrix of instruments is constructed using the estimates in table 4.9. Apart from the other parameter estimates, now also the parameter estimate of $\hat{\eta}_3$ (number of children below the age of 6) has the expected sign. The wage is maximal at the age of 37 and the number of hours supplied is maximal at a wage rate of 54.7.

Table 4.7 shows the method 2 estimates when $R = 10$ drawings are used. Again, the matrix of instruments is constructed using the estimates in table 4.9. The wage equation is maximal at the age of 37, whereas labour supply is maximal at a wage rate of 32.9. The main difference between the method 1 estimates and the method 2 estimates are the parameter estimates of the labour supply function.

Finally, we look at the estimates obtained by method 3. In table 4.8 the results with $R = 5$ drawings are presented. We can make the same remarks about the signs of the estimates as in the previous cases. Wage is maximal at the age of 38 and hours supplied are maximal at the wage rate of 9.4. To obtain the results in table 4.9, we used $R = 10$ drawings to construct the simulators. Most of the standard errors of the estimates are lower than in table 4.8, and whenever they aren't lower, they are only slightly higher.

In table 4.10 we present the estimation results with $R = 50$ drawings. Comparing

tables 4.4 and 4.8, we see that there is not much difference between the results with $R = 10$ and $R = 50$ drawings. Apparently $R = 10$ drawings are in this case sufficient to minimize the effect of the simulation residuals.

Comparing the three methods, we can say that method 1 is the cheapest in C.P.U.-time because it doesn't make use of simulators of the derivatives. Also, it appears to produce the smallest standard errors. However, it only makes sense to use this method when a consistent estimate is available to construct the matrix of instruments. Method 2 is the most expensive in C.P.U.-time.

Table 4.11 gives the results of estimating the model without the restriction of zero correlation between the disturbances by method 3, using 10 drawings to simulate the response probabilities. The estimates of the disturbances' variances and the covariance imply a correlation coefficient of $\rho = 0.082$ which is not significantly different from zero. A comparison of tables 4.11 and 4.9 reveals no big shifts in the parameter estimates.

To get some more feeling for the differences in estimates across methods, we present elasticities of hours worked and of participation with respect to wages. These have been calculated as "aggregate" elasticities in the sense that all wages in the sample have been raised by 5% and then hours and participation probabilities have been predicted for every individual in the sample. The observed changes in the sample averages of these quantities are used to compute the elasticities. The results are given in table 4.12.

Strikingly, ML gives elasticities that are much larger than those implied by the other methods. Method 1 is most similar to ML in this respect. It is hard to interpret these differences. In principle they would call for specification tests. Given the simplistic nature of the model and the illustrative purposes of the estimation we abstain from a specification search.¹

As a final comparison of estimation methods, we present in table 4.13 the likelihood values corresponding to the estimates obtained by the various methods. We now see that MSS2 is closest to ML and MSS3 has the lowest likelihood value.

¹Presumably the most important problem with the present model is that it assumes that anyone who wants to work can do so (if $h^* > 0$, one works a positive amount of hours). This assumption is far too strong, see for instance Blundell, Ham and Meghir (1987) or Kapteyn and Woitties (1989).

TABLE 4.3 MAXIMUM LIKELIHOOD ESTIMATES

parameter	$\hat{\theta}^1$	standard error
α_1	-86.696	10.860
α_2	$-5.605 \cdot 10^{-2}$	0.00936
α_3	12.1126	1.312
α_4	-0.574	0.0625
σ_ϵ^2	558.052	96.753
σ_u^2	0.135	0.0164
η_1	-14.959	3.733
η_2	-0.387	0.0766
η_3	$-8.765 \cdot 10^{-2}$	0.0206
η_4	10.322	2.110
η_5	-0.767	0.119
η_6	-0.771	0.117
η_7	-0.625	0.115
η_8	-0.317	0.112
η_9	0.266	0.410
η_{10}	$5.611 \cdot 10^{-2}$	0.0511
η_{11}	-1.452	0.295

TABLE 4.4 ESTIMATES BY METHOD 1
NUMBER OF DRAWINGS R = 5

parameter	$\hat{\theta}^1$	standard error
α_1	-22.595	9.43
α_2	$-6.194 \cdot 10^{-2}$	0.02
α_3	3.158	1.12
α_4	-0.117	0.05
σ_ϵ^2	309.942	17.53
σ_u^2	$6.819 \cdot 10^{-2}$	0.01
η_1	-11.855	2.43
η_2	-0.118	0.06
η_3	$1.353 \cdot 10^{-2}$	0.03
η_4	8.198	1.39
η_5	-0.522	0.10
η_6	-0.454	0.10
η_7	-0.395	0.10
η_8	-0.156	0.10
η_9	0.163	0.04
η_{10}	$6.796 \cdot 10^{-2}$	0.04
η_{11}	-1.128	0.02

TABLE 4.5 ESTIMATES BY METHOD 1
NUMBER OF DRAWINGS R = 10

parameter	$\hat{\theta}^1$	standard error
α_1	-22.595	9.09
α_2	$-6.209 \cdot 10^{-2}$	0.02
α_3	3.162	1.08
α_4	-0.118	0.05
σ_ϵ^2	309.942	17.12
σ_u^2	$6.817 \cdot 10^{-2}$	0.01
η_1	-11.855	2.37
η_2	-0.118	0.01
η_3	$1.373 \cdot 10^{-2}$	0.03
η_4	8.198	1.36
η_5	-0.521	0.10
η_6	-0.454	0.10
η_7	-0.395	0.10
η_8	-0.156	0.10
η_9	0.163	0.04
η_{10}	$6.831 \cdot 10^{-2}$	0.04
η_{11}	-1.129	0.02

TABLE 4.6 ESTIMATES BY METHOD 2
NUMBER OF DRAWINGS R = 5

parameter	$\hat{\theta}^2$	standard error
α_1	-7.930	17.73
α_2	$-4.815 \cdot 10^{-2}$	0.01
α_3	0.989	2.45
α_4	$-4.815 \cdot 10^{-3}$	0.13
σ_ϵ^2	345.479	22.03
σ_u^2	$7.911 \cdot 10^{-2}$	0.02
η_1	-12.182	2.77
η_2	-0.198	0.10
η_3	$-4.607 \cdot 10^{-2}$	0.05
η_4	8.494	1.62
η_5	-0.579	0.12
η_6	-0.540	0.13
η_7	-0.435	0.11
η_8	-0.187	0.11
η_9	0.188	0.06
η_{10}	$6.186 \cdot 10^{-2}$	0.04
η_{11}	-1.177	0.23

TABLE 4.7 ESTIMATES BY METHOD 2
NUMBER OF DRAWINGS R = 10

parameter	$\hat{\theta}^2$	standard error
α_1	-8.686	19.59
α_2	-4.817 10^{-2}	0.01
α_3	1.098	2.71
α_4	-1.667 10^{-2}	0.14
σ_ϵ^2	346.921	23.92
σ_u^2	7.924 10^{-2}	0.02
η_1	-12.441	2.84
η_2	-0.202	0.01
η_3	-4.507 10^{-2}	0.05
η_4	8.644	1.67
η_5	-0.580	0.12
η_6	-0.541	0.14
η_7	-0.437	0.11
η_8	-0.189	0.11
η_9	0.188	0.06
η_{10}	6.299 10^{-2}	0.04
η_{11}	-1.199	0.24

TABLE 4.8 ESTIMATES BY METHOD 3
NUMBER OF DRAWINGS R = 5

parameter	$\hat{\theta}^3$	standard error
α_1	-6.665	18.43
α_2	-6.680 10^{-2}	0.01
α_3	3.423	3.03
α_4	-0.183	0.01
σ_ϵ^2	269.000	111.58
σ_u^2	6.959 10^{-2}	0.10
η_1	-11.382	2.97
η_2	-0.115	0.09
η_3	4.755 10^{-2}	0.04
η_4	7.902	1.73
η_5	-0.519	0.12
η_6	-0.438	0.12
η_7	-0.388	0.11
η_8	-0.148	0.11
η_9	0.166	0.05
η_{10}	7.139 10^{-2}	0.04
η_{11}	-1.084	0.02

TABLE 4.9 ESTIMATES BY METHOD 3
NUMBER OF DRAWINGS R = 10

parameter	$\hat{\theta}^3$	standard error
α_1	-7.717	13.55
α_2	-6.593 10^{-2}	0.01
α_3	3.137	2.35
α_4	-0.163	0.12
σ_ϵ^2	308.748	125.49
σ_u^2	6.918 10^{-2}	0.01
η_1	-11.368	2.72
η_2	-0.112	0.07
η_3	4.146 10^{-2}	0.03
η_4	7.895	1.57
η_5	-0.517	0.11
η_6	-0.438	0.11
η_7	-0.387	0.10
η_8	-0.148	0.11
η_9	0.165	0.04
η_{10}	7.168 10^{-2}	0.04
η_{11}	-1.084	0.02

TABLE 4.10 ESTIMATES BY METHOD 3
NUMBER OF DRAWINGS R = 50

parameter	$\hat{\theta}^3$	standard error
α_1	-7.716	13.69
α_2	-6.598 10^{-2}	0.01
α_3	3.149	2.39
α_4	-0.164	0.01
σ_ϵ^2	308.748	130.46
σ_u^2	6.932 10^{-2}	0.01
η_1	-11.369	2.76
η_2	-0.111	0.08
η_3	4.324 10^{-2}	0.04
η_4	7.894	1.59
η_5	-0.517	0.11
η_6	-0.437	0.11
η_7	-0.386	0.10
η_8	-0.147	0.11
η_9	0.168	0.05
η_{10}	7.263 10^{-2}	0.04
η_{11}	-1.084	0.02

TABLE 4.11 ESTIMATES BY METHOD 3
NUMBER OF DRAWINGS R = 10
CORRELATED DISTURBANCES

parameter	$\hat{\theta}^s$	standard error
α_1	-8.994	10.872
α_2	$-6.592 \cdot 10^{-2}$	0.01
α_3	3.298	1.84
α_4	-0.172	0.10
σ_ϵ^2	308.869	91.15
σ_u^2	$7.054 \cdot 10^{-2}$	0.01
$\sigma_{\epsilon u}$	0.147	0.58
η_1	-11.728	2.75
η_2	-0.115	0.07
η_3	$3.992 \cdot 10^{-2}$	0.04
η_4	8.093	1.57
η_5	-0.522	0.11
η_6	-0.442	0.11
η_7	-0.390	0.10
η_8	-0.149	0.11
η_9	0.169	0.05
η_{10}	$7.060 \cdot 10^{-2}$	0.04
η_{11}	-1.111	0.22

TABLE 4.12 WAGE ELASTICITIES ACCORDING
TO DIFFERENT ESTIMATION METHODS, R=10

	method 1	method 2	method 3	ML
hours	1.119	0.637	0.518	1.943
participation	0.606	0.133	0.236	1.369

TABLE 4.13 VALUES OF THE LIKELIHOOD FUNCTION
estimates from table:

	likelihood value
4.3 (ML)	-2738.050
4.4 (method 1)	-2881.038
4.7 (method 2)	-2866.351
4.9 (method 3)	-3006.794

5 Conclusions

The main purpose of the paper has been to investigate the usefulness of MSS estimators in mixed discrete-continuous models, with a focus on the kind of model typically encountered in the analysis of labour supply. The experience in this paper appears to be that the estimators perform quite well. In the example considered, the MSS estimators do a little worse than ML, but in more complex situations ML would simply be infeasible. This is not only a matter of computing time, but also due to the fact that in certain situations it is impossible to write down analytically the probability of certain events, whereas the events can still be simulated. Estimation by simulation techniques then turns out to be a useful tool in the analysis of labour supply models, in the sense that these techniques enable us to estimate models which cannot, or only with great difficulty, be estimated with conventional methods like maximum likelihood or the method of moments.

The simulated scores methods presented in this paper perform satisfactory, even with a limited number of drawings. For the use of a limited number of drawings a price has to be paid in the form of a loss in efficiency, but this loss is modest. This is in stark contrast with the method of simulated maximum likelihood with a limited number of drawings, which performs poorly.

The method of simulated scores will be the more useful, the higher is the dimension of integration in the evaluation of response probabilities in the likelihood function. In this context one may think of models of family labour supply with various sources of randomness. In this papers we have only used smooth simulators. In more complex models the use of frequency simulators cannot always be avoided. Their main drawback is their discontinuity in the parameters as a result of which conventional gradient based optimization procedures cannot be used. The downhill simplex method by Nelder and Mead (1965), employed by Bloemen and Kapteyn (1992) for instance, is quite time consuming. This disadvantage, however, is mitigated by the possibility to use a limited number of drawings in the method of simulated scores.

References

- Bloemen, H.G. and A. Kapteyn (1992), *The Estimation of a Utility Consistent Stochastic Labour Supply Model by Means of Simulated Moments*, working paper, Tilburg University.
- Blundell, R.W., J. Ham and C. Meghir (1987), *Unemployment and Female Labour Supply*, *Economic Journal*, 97, pp. 44-64
- Gouriéroux, C. and A. Monfort (1989), *Simulation Based Inference in Models with Heterogeneity*, mimeo, report 8905, INSEE, Paris
- Hajivassiliou, V. (1989), *Simulated Scores for Estimating LDV Models: A Presentation and Comparative Evaluation*, mimeo, Cowles Foundation
- Hajivassiliou, V. (1992), *Simulation Estimation Methods for Limited Dependent Variable Models*, in: Handbook of Statistics, Vol. 11 (Econometrics), North-Holland.
- Hammersley, J.M. and D.C. Handscomb (1964), *Monte Carlo Methods*, London: Methuen.
- Hausman, J. and P. Ruud (1984), *Family Labor Supply with Taxes*, *American Economic Review*, Vol. 74, pp. 242-248
- Kapteyn, A. and I. Woittiez (1989), *Preference Interdependence and Habit Formation in Family Labour Supply*, in: Microeconometrics: Surveys and Applications, ed. by F. Laisney et al., Clevedon, Avon, Tieto Ltd., Chapter 10.
- Lerman, C. and C. Manski (1981), *On the Use of Simulated Frequencies to Approximate Choice Probabilities*, in: Structural Analysis of Discrete Data with Economic Applications, ed. C. Manski and D. McFadden, MIT Press, Cambridge, pp. 305-319
- McFadden, D. (1989), *A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration*, *Econometrica*, Vol. 57, pp. 995-1026
- Nelder, J.A. and R. Mead (1965), *Computer Journal* Vol. 7, pp. 308.
- Pakes, A. and D. Pollard (1989), *Simulation and the Asymptotics of Optimization Estimators*, *Econometrica*, Vol. 57, pp. 1027-1057
- Stroud, A.H. and D. Secrest (1966), *Gaussian Quadrature Formulas*, Prentice-Hall

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9106	C.N. Teulings	The Diverging Effects of the Business Cycle on the Expected Duration of Job Search
9107	E. van Damme	Refinements of Nash Equilibrium
9108	E. van Damme	Equilibrium Selection in 2 x 2 Games
9109	G. Alogoskoufis and F. van der Ploeg	Money and Growth Revisited
9110	L. Samuelson	Dominated Strategies and Common Knowledge
9111	F. van der Ploeg and Th. van de Klundert	Political Trade-off between Growth and Government Consumption
9112	Th. Nijman, F. Palm and C. Wolff	Premia in Forward Foreign Exchange as Unobserved Components
9113	H. Bester	Bargaining vs. Price Competition in a Market with Quality Uncertainty
9114	R.P. Gilles, G. Owen and R. van den Brink	Games with Permission Structures: The Conjunctive Approach
9115	F. van der Ploeg	Unanticipated Inflation and Government Finance: The Case for an Independent Common Central Bank
9116	N. Rankin	Exchange Rate Risk and Imperfect Capital Mobility in an Optimising Model
9117	E. Bomhoff	Currency Convertibility: When and How? A Contribution to the Bulgarian Debate!
9118	E. Bomhoff	Stability of Velocity in the G-7 Countries: A Kalman Filter Approach
9119	J. Osiewalski and M. Steel	Bayesian Marginal Equivalence of Elliptical Regression Models
9120	S. Bhattacharya, J. Glazer and D. Sappington	Licensing and the Sharing of Knowledge in Joint Ventures
9121	J.W. Friedman and L. Samuelson	An Extension of the "Folk Theorem" with Continuous Reaction Functions
9122	S. Chib, J. Osiewalski and M. Steel	A Bayesian Note on Competing Correlation Structures in the Dynamic Linear Regression Model
9123	Th. van de Klundert and L. Meijdam	Endogenous Growth and Income Distribution

No.	Author(s)	Title
9124	S. Bhattacharya	Banking Theory: The Main Ideas
9125	J. Thomas	Non-Computable Rational Expectations Equilibria
9126	J. Thomas and T. Worrall	Foreign Direct Investment and the Risk of Expropriation
9127	T. Gao, A.J.J. Talman and Z. Wang	Modification of the Kojima-Nishino-Arima Algorithm and its Computational Complexity
9128	S. Altug and R.A. Miller	Human Capital, Aggregate Shocks and Panel Data Estimation
9129	H. Keuzenkamp and A.P. Barten	Rejection without Falsification - On the History of Testing the Homogeneity Condition in the Theory of Consumer Demand
9130	G. Mailath, L. Samuelson and J. Swinkels	Extensive Form Reasoning in Normal Form Games
9131	K. Binmore and L. Samuelson	Evolutionary Stability in Repeated Games Played by Finite Automata
9132	L. Samuelson and J. Zhang	Evolutionary Stability in Asymmetric Games
9133	J. Greenberg and S. Weber	Stable Coalition Structures with Uni-dimensional Set of Alternatives
9134	F. de Jong and F. van der Ploeg	Seigniorage, Taxes, Government Debt and the EMS
9135	E. Bomhoff	Between Price Reform and Privatization - Eastern Europe in Transition
9136	H. Bester and E. Petrakis	The Incentives for Cost Reduction in a Differentiated Industry
9137	L. Mirman, L. Samuelson and E. Schlee	Strategic Information Manipulation in Duopolies
9138	C. Dang	The D_x^* -Triangulation for Continuous Deformation Algorithms to Compute Solutions of Nonlinear Equations
9139	A. de Zeeuw	Comment on "Nash and Stackelberg Solutions in a Differential Game Model of Capitalism"
9140	B. Lockwood	Border Controls and Tax Competition in a Customs Union
9141	C. Fershtman and A. de Zeeuw	Capital Accumulation and Entry Deterrence: A Clarifying Note
9142	J.D. Angrist and G.W. Imbens	Sources of Identifying Information in Evaluation Models

No.	Author(s)	Title
9143	A.K. Bera and A. Ullah	Rao's Score Test in Econometrics
9144	B. Melenberg and A. van Soest	Parametric and Semi-Parametric Modelling of Vacation Expenditures
9145	G. Imbens and T. Lancaster	Efficient Estimation and Stratified Sampling
9146	Th. van de Klundert and S. Smulders	Reconstructing Growth Theory: A Survey
9147	J. Greenberg	On the Sensitivity of Von Neuman and Morgenstern Abstract Stable Sets: The Stable and the Individual Stable Bargaining Set
9148	S. van Wijnbergen	Trade Reform, Policy Uncertainty and the Current Account: A Non-Expected Utility Approach
9149	S. van Wijnbergen	Intertemporal Speculation, Shortages and the Political Economy of Price Reform
9150	G. Koop and M.F.J. Steel	A Decision Theoretic Analysis of the Unit Root Hypothesis Using Mixtures of Elliptical Models
9151	A.P. Barten	Consumer Allocation Models: Choice of Functional Form
9152	R.T. Baillie, T. Bollerslev and M.R. Redfearn	Bear Squeezes, Volatility Spillovers and Speculative Attacks in the Hyperinflation 1920s Foreign Exchange
9153	M.F.J. Steel	Bayesian Inference in Time Series
9154	A.K. Bera and S. Lee	Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis
9155	F. de Jong	A Univariate Analysis of EMS Exchange Rates Using a Target
9156	B. le Blanc	Economies in Transition
9157	A.J.J. Talman	Intersection Theorems on the Unit Simplex and the Simplotope
9158	H. Bester	A Model of Price Advertising and Sales
9159	A. Özcam, G. Judge, A. Bera and T. Yancey	The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model
9160	R.M.W.J. Beetsma	Bands and Statistical Properties of EMS Exchange Rates: A Monte Carlo Investigation of Three Target Zone Models Zone Model

No.	Author(s)	Title
9161	A.M. Lejour and H.A.A. Verbon	Centralized and Decentralized Decision Making on Social Insurance in an Integrated Market Multilateral Institutions
9162	S. Bhattacharya	Sovereign Debt, Creditor-Country Governments, and
9163	H. Bester, A. de Palma, W. Leininger, E.-L. von Thadden and J. Thomas	The Missing Equilibria in Hotelling's Location Game
9164	J. Greenberg	The Stable Value
9165	Q.H. Vuong and W. Wang	Selecting Estimated Models Using Chi-Square Statistics
9166	D.O. Stahl II	Evolution of Smart _n Players
9167	D.O. Stahl II	Strategic Advertising and Pricing with Sequential Buyer Search
9168	T.E. Nijman and F.C. Palm	Recent Developments in Modeling Volatility in Financial Data
9169	G. Asheim	Individual and Collective Time Consistency
9170	H. Carlsson and E. van Damme	Equilibrium Selection in Stag Hunt Games
9201	M. Verbeek and Th. Nijman	Minimum MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross Sections
9202	E. Bomhoff	Monetary Policy and Inflation
9203	J. Quiggin and P. Wakker	The Axiomatic Basis of Anticipated Utility; A Clarification
9204	Th. van de Klundert and S. Smulders	Strategies for Growth in a Macroeconomic Setting
9205	E. Siandra	Money and Specialization in Production
9206	W. Härdle	Applied Nonparametric Models
9207	M. Verbeek and Th. Nijman	Incomplete Panels and Selection Bias: A Survey
9208	W. Härdle and A.B. Tsybakov	How Sensitive Are Average Derivatives?
9209	S. Albæk and P.B. Overgaard	Upstream Pricing and Advertising Signal Downstream Demand
9210	M. Cripps and J. Thomas	Reputation and Commitment in Two-Person Repeated Games
9211	S. Albæk	Endogenous Timing in a Game with Incomplete Information

No.	Author(s)	Title
9212	T.J.A. Storcken and P.H.M. Ruys	Extensions of Choice Behaviour
9213	R.M.W.J. Beetsma and F. van der Ploeg	Exchange Rate Bands and Optimal Monetary Accommodation under a Dirty Float
9214	A. van Soest	Discrete Choice Models of Family Labour Supply
9215	W. Güth and K. Ritzberger	On Durable Goods Monopolies and the (Anti-) Coase- Conjecture
9216	A. Simonovits	Indexation of Pensions in Hungary: A Simple Cohort Model
9217	J.-L. Ferreira, I. Gilboa and M. Maschler	Credible Equilibria in Games with Utilities Changing during the Play
9218	P. Borm, H. Keiding, R. Mclean, S. Oortwijn and S. Tijs	The Compromise Value for NTU-Games
9219	J.L. Horowitz and W. Härdle	Testing a Parametric Model against a Semiparametric Alternative
9220	A.L. Bovenberg	Investment-Promoting Policies in Open Economies: The Importance of Intergenerational and International Distributional Effects
9221	S. Smulders and Th. van de Klundert	Monopolistic Competition, Product Variety and Growth: Chamberlin vs. Schumpeter
9222	H. Bester and E. Petrakis	Price Competition and Advertising in Oligopoly
9223	A. van den Nouweland, M. Maschler and S. Tijs	Monotonic Games are Spanning Network Games
9224	H. Suehiro	A "Mistaken Theories" Refinement
9225	H. Suehiro	Robust Selection of Equilibria
9226	D. Friedman	Economically Applicable Evolutionary Games
9227	E. Bomhoff	Four Econometric Fashions and the Kalman Filter Alternative - A Simulation Study
9228	P. Borm, G.-J. Otten and Hans Peters	Core Implementation in Modified Strong and Coalition Proof Nash Equilibria
9229	H.G. Bloemen and A. Kapteyn	The Joint Estimation of a Non-Linear Labour Supply Function and a Wage Equation Using Simulated Response Probabilities

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS

Bibliotheek K. U. Brabant

17 000 01117437 3