A Structural Labour Supply Model with Nonparametric Preferences

Arthur van Soest
Marcel Das
Xiaodong Gong

November 2000

A Structural Labour Supply Model with Nonparametric Preferences

Arthur van Soest
Tilburg University

Marcel Das
Tilburg University

Xiaodong Gong
IZA, Bonn

ABSTRACT

A Structural Labour Supply Model with Nonparametric Preferences[*]

Nonparametric techniques are usually seen as a statistic device for data description and exploration, and not as a tool for estimating models with a richer economic structure, which are often required for policy analysis. This paper presents an example where nonparametric flexibility can be attained in a fully structural model. A structural labour supply model with a nonparametric specification of preferences is introduced, which can be used for the analysis of all sorts of (non-linear) tax and benefits changes. Moreover, the model can deal with several other problems in estimation of structural labour supply models, such as non-convex tax rules, benefits, unobserved wages of non-workers, and model coherency. The utility maximization problem is solved by discretizing the budget set and choosing the optimal leisure and income combination from a finite set of alternatives. The direct utility function is approximated with a series expansion. For a given length of the expansion, the model is estimated by smooth simulated maximum likelihood. The wage equation is estimated jointly with the labour supply model, and measurement errors in wage rates are allowed for. The model is estimated with Dutch data on labour supply of married females, for various lengths of the series expansion. Estimates of labour supply elasticities and effects of a proposed tax reform suggest that the results do not change much once the order of the series expansion is extended beyond two, even though the second order model is statistically rejected against higher order models. Monte Carlo simulations are used to show that the estimation strategy has remarkably good finite sample properties for the size of our sample. On the other hand they lead to some concern about the potential bias to measurement error in the hours variable.

Xiaodong Gong
IZA
P.O. Box 7240
D-53072 Bonn
Germany
Tel.: +49 228 38 94 527
Fax: +49 228 38 94 510
Email: gong@iza.org

---

1 Introduction

Nonparametric regression techniques are usually seen as a statistic device for data description and exploration. They are typically not used for estimating more complex models with a rich economic structure, which are often required for policy analysis. Models for policy analysis, therefore, are often characterized by restrictive functional form assumptions, which are required to make the econometrics tractable, but lack sound economic foundation.

An example is the extensive literature on neo-classical structural labour supply models. This literature goes back to Hausman's kinked budget constraint model, which has been extended in many directions and applied to many data sets. See, for example, Burtless and Hausman (1978), Hausman (1981, 1985), and the surveys of Moffitt (1986, 1990) and Blundell and MaCurdy (1999). It is one of the first type of limited dependent variable models in micro-econometrics where economic theory and econometric specification are directly linked. In most applications of this model, a restrictive specification of preferences is used, particularly in case of nonconvexities in the budget set. The main reason is that an analytic solution of the utility maximization problem then requires explicit expressions for both the direct utility function and the labour supply function (or the indirect utility or expenditure function). Thus the rich economic structure of the model hampers flexibility of the specification.

In the case of piecewise linear convex budget constraints, this problem can be avoided. Using the specific search algorithm for the optimum for this case (see Blomquist, 1983, for example), only the specification of the labour supply curve is necessary, and more flexible specifications come within reach. Blomquist and Newey (1997) have exploited this idea to estimate the labour supply curve nonparametrically, approximating it by a series expansion. This is to our knowledge the only example in this field where economic theory and nonparametrics are combined into one econometric model. Still, this study remains close to the original Hausman framework, and does not allow for, for example, non-convex budget sets, unobserved wages of non-workers, or joint decisions of two spouses. The main reason is that Blomquist and Newey use the labour supply function only, and avoid the need to specify the underlying direct utility function.

In this paper, we build the model around a nonparametric specification of the direct utility function. In this way, we introduce a structural nonparametric labour supply model which can be used for the analysis of all sorts of (non-linear) tax and benefits changes. Following van Soest (1995), we replace the actual budget set by a finite number of points on its boundary, and approximate the utility maximization problem by finding the best point in this finite set. We do not require that the tax and benefits system is piecewise linear or convex, and we show how fixed costs of working, unobserved wages, and non-participation can be incorporated.

Moreover, we show how our model avoids the critique by MaCurdy et al. (1990) that coherency of the model implicitly limits the range of elasticities that can be obtained. This implies that policy outcomes may be driven by implicit constraints on the chosen

model, rather than by the data and the estimates. Thus our framework allows for a nonparametric specification of preferences but also for many other features of structural labour supply models which have been addressed in the literature.

In our framework, the direct utility function is approximated by a nonparametric series approximation in hours and income. Observed individual or family characteristics are incorporated through one or more of the coefficients in the series approximation. Unobserved heterogeneity can be incorporated by treating one or more of the parameters of the utility function as random coefficients. Their random parts can in principle be allowed to have some flexible distribution, for example a mixture of normals or a discrete distribution with several mass points, as in Heckman and Singer (1984).

For given length of the series expansion, our model can be estimated by smooth simulated maximum likelihood. GEV I type errors are added to the utility values of all alternatives in the finite choice set. These errors can be seen as alternative specific random errors in utility evaluation, or can simply be interpreted as a smoothing device (as in Keane and Moffitt, 1998). The wage equation is estimated jointly with the labour supply model, and non-participants are included.

The results can be used to compute labour supply elasticities or to analyze the effects of changes in taxes and benefits on participation and labour supply. The approach will be applied to data from the Dutch Socio-Economic Panel, and we will illustrate the usefulness of our results for policy analysis by studying the consequences of a recently proposed income tax reform for which nonconvexities in the budget constraint are particularly important.

In statistical terms, the method we propose is similar to, for example, the technique of semi-nonparametric maximum likelihood introduced by Gallant and Nychka (1987). Error term distributions and/or the utility function can be specified with some flexible series approximation. Once the number of terms in the series approximations is fixed, ML is performed, with its usual Kullback-Leibner information criterion interpretation. The nonparametric feature is that the number of terms can become large, i.e. increases to infinity with the number of observations, but at a slower rate. We focus on the practical application of these ideas, and not on deriving the appropriate convergence rates or the asymptotics of the estimator if the number of terms in the series approximation tends to infinity. Thus, formally, our hypothesis tests and standard errors are only valid under the assumption that the length of the series approximation is given, and that the utility function is perfectly captured this given length. In this case, standard properties of (parametric) simulated maximum likelihood apply. Comparing results for different lengths of the series expansion will show whether this is an overly restrictive view.

The difference between our framework and many other nonparametric models is that economic theory is used to impose some structure on the model. A utility function is estimated, but labour supply, i.e., the outcome of utility maximization, is observed. One wage rate (at most) is observed for each individual, but the complete budget set is needed for utility maximization. Economic theory does not impose the functional form of the

utility function. It therefore seems natural to use a flexible, nonparametric specification of this utility function, while at the same time retaining the economic structure of the model.

So basically, we combine information in the data with two types of prior information: the nonparametric assumption of utility maximization, and the limitation of the number of terms in the series approximations required due to the finite size of the sample.

The structure of the remainder of this paper is as follows. In Section 2, we introduce the model and discuss the advantages and drawbacks of the discrete approach compared to the continuous approach. In Section 3, we discuss the estimates of the model for Dutch female labour supply data. We focus on the sensitivity of the results for the chosen length of the series approximation. We look at labour supply elasticities, and at the effects on participation and hours worked of a recently proposed tax reform. In Section 4, we discuss several simulation exercises to investigate the quality of our ML estimator, and its sensitivity to misspecification. Section 5 concludes.

2 Model

We present a static neo-classical structural labour supply model. The basic framework is similar to that of van Soest (1995). We describe the model for single decision makers. The model for joint decision makers such as married couples is a straightforward generalization of this, but is not the focus of the current paper. In the application below, we will analyze labour supply of married females, conditional on hours worked by their husbands (and on the husbands' earnings), where the woman is the only decision maker in the model.

Utility

The individual's utility depends on leisure ($TE - h$, where TE is the time endowment), and on total net income of the family (y). Net income is partly determined by the individual's own earnings, but can also contain spouse's earnings, asset income, child allowances, etc.. We follow the majority of labour supply studies, in which 'leisure' is the aggregate of all other time uses except work. The time endowment is a common constant for all individuals, and will drop out of the polynomial expansions. As a consequence, it is equivalent to work with a direct utility function with arguments y and h instead of y and $TE - h$.

The model would be consistent with utility maximization in a life cycle framework with intertemporally additive preference if net income could be replaced by total expenditures (see Blundell and Walker, 1986). In our application, the data do not contain any information on consumption expenditures or savings, so that we remain in a static framework.

The direct utility function is specified as a polynomial in its arguments h and y:

$$U(h, y) = \sum_{p=0,\ldots,K} \sum_{q=0,\ldots,K-p} \alpha(p, q) h^p y^q; \tag{1}$$

K is the order of the polynomial and determines the flexibility of the utility function. If K is allowed to become arbitrarily large, the parameters $\alpha(p, q)$ can be chosen in such a way that $U(h, y)$ can approximate any given function of h and y to any desired accuracy on a given compact set. In this sense, the class of utility functions in (1) is very general, and can be seen as a nonparametric family of utility functions. On the other hand, for a finite sample size, the order of the polynomial that can be used is limited. As usual with nonparametric series expansions, asymptotics will require that K tends to infinity much slower than the number of observations. In practice on the other hand, only small values of K can be used in estimation for the size of the sample which is typically available. We will use $K = 1, 2, 3, 4$ and 5, and compare the results for these different values of K.

We will see below that the economic interpretation of the model requires that the utility function is increasing with income (this is due to the assumption that everyone always chooses a point on the frontier of her budget set rather than in the interior). We will not impose this assumption *a priori* but check whether it is satisfied by our unrestricted estimates. We will also see that the model does not require quasi-concavity of preferences, so that we do not have to impose this either (this is because we use utility maximization over a finite budget set, not requiring tangency conditions). Thus we will not impose any *a priori* restrictions on the utility function in (1).

To allow the utility function to vary with taste shifters such as age and the number of children, the parameters $\alpha(p, q)$ can be made dependent on a vector X of individual and family characteristics. In practice, however, it will be hard to disentangle effects of X via different $\alpha(p, q)$. In the estimates below therefore, only the parameter $\alpha(1, 0)$ –the coefficient of the linear term h in (1) – will be allowed to vary with X: $\alpha(1, 0) = \beta_0 + X'\beta$. This is an attractive choice from the point of view of interpreting the results, since it implies that the marginal utility of leisure varies linearly with X, and the sign of the coefficients in β immediately determines whether the corresponding taste shifter has a positive or a negative impact on the marginal utility of leisure, i.e., a negative or a positive impact on labour supply. On the other hand, there is no theoretical reason not to allow more of the $\alpha(p, q)$ to depend on X, or not to allow for a nonlinear effect of X on these parameters. A specification which is fully nonparametric in X would obviously require this. Our series expansion makes the utility function nonparametric in its arguments h and y, but not in X. Practical limitations due to curse of dimensionality problems and limitations of the data prevented us from experimenting with specifications which are more flexible in X. Moreover, the utility function is only identified up to a monotonic transformation which may depend on X, implying that complete flexibility in terms of X would lead to an unidentified model (cf., e.g., Pollak and Wales, 1979).

Apart from heterogeneity through observed characteristics X, preferences may also vary with unobserved characteristics. To incorporate this, the Hausman type models

have typically allowed for random preferences by adding an error term to one of the parameters of the utility function. We follow the same strategy and, in line with the way X is allowed to enter, assume that unobserved heterogeneity (u_{rp}) enters through the parameter $\alpha(1,0)$:

$$\alpha(1,0) = \beta_0 + X'\beta + u_{rp}; \quad u_{rp} \sim N(0, \sigma_{rp}^2). \tag{2}$$

The normality assumption is mainly made for convenience, and could in principle be replaced by a more flexible distribution. For example, a mixture of normals could be used, or a discrete distribution with a finite number of mass points (cf. Heckman and Singer, 1984). While these are interesting non-parametric extensions of the framework considered here, we will not consider them in our estimates below. The reason is that we already found it hard to accurately estimate the variance σ_{rp} in the normal specification given by (2), and could not find strong support for the presence of unobserved heterogeneity, with the standard error on the estimate of σ_{rp} typically exceeding its point estimate.

Constraints

Labour supply is based upon utility maximization under constraints. An obvious constraint is the budget restriction: to each choice of the number of working hours corresponds a different net income for the family. To determine net family income as a function of the wife's working hours, we need her earnings, other household income (husband's earnings, child benefits, asset income), taxes, and potential unemployment benefits or social security benefits. The components of other household income are observed or can be computed directly from variables observed in the data. To determine the wife's earnings for each number of working hours, we assume that her gross hourly wage rate does not depend on her hours worked. This is a common assumption in most of the structural labour supply literature, although exceptions exist (see Moffitt, 1984, or Tummers and Woittiez, 1991). If wage rates are observed in the data without measurement error, we can then compute gross earnings for each possible number of working hours for those women who work.

For non-workers, we need to predict the before tax wage rate. For this purpose, and to be able to take account of measurement error in observed wage rates, we explicitly need to incorporate a wage equation. To account for selectivity of observed wages in a way consistent with the labour supply model, we estimate the wage equation jointly with the labour supply model.[1] The parameters in the wage equations are then used to predict the wages of non-workers. Because the labour supply model is nonlinear

---

[1]Here we follow Gong and van Soest (2000) rather than Van Soest (1995). The latter estimates the wage equation separately using a standard Heckman selection model (Heckman, 1979), the selection equation which can be seen as a (linearized reduced form) approximation to the selection mechanism implied by the structural model

in wages, it is necessary to take the wage rate prediction errors into account to get consistent estimates of the labour supply model (see the description of the estimation technique given below).

The wage equation we use is given by

$$\text{Log} w = Z' \gamma + u_w + e_w; \tag{3}$$

$u_w \sim N(0, \sigma_u^2)$ (unobserved heterogeneity in wages);
$e_w \sim N(0, \sigma_e^2)$ (measurement error);
u_w, e_w independent of each other and of other error terms in the model.
Here w is the observed wage rate, which possibly contains measurement error e_w. According to simulation results of Blomquist (1996), such measurement error could be quite influential for the elasticities of interest, and this is the reason why we incorporate it. Note that σ_e and σ_u are separately identified because the wage which enters the labour supply part of the model includes u_w but not e_w. Thus a more general interpretation of e_w might be the part of the wage rate which is job or hours specific, and which is not used in the respondents' labour supply decision.

We do not allow for measurement errors on other variables. If women's wages are measured with error, men's wages probably contain error as well, and in our model this would lead to measurement error in other income. This could be included in a similar way, but would require an equation for other income. To keep things simple, we do not do this in the current paper. Another reason for not addressing this issue is that results of Blomquist (1996) suggest that the bias induced by this is much smaller than the bias due to measurement error on the wage rate.[2]

In the sample we use for estimation, all husbands work, and usually earn so much that social assistance benefits for the family do not apply: family incomes excluding the wife's own earnings usually exceeds the official minimum standard of living, which depends on age, marital status and family composition. For the few families in which husband's earnings are so low that this is not the case, we incorporate social assistance benefits: if family income is lower than the official minimum standard of living, it is increased up to the minimum standard of living threshold.

We do not model unemployment insurance benefits. This is difficult to model due to lack of data and due to the static nature of our framework, since unemployment insurance benefits are of temporary nature. The most important type of benefits that enter the budget constraints are child benefits, which do not depend on earnings or labour market status of the two spouses.

Following van Soest (1995), the budget constraint under which the individual maximizes utility is approximated by a finite number of points. We take multiples of 4 hours

---

[2]The specification of the error term in the wage equation is different from that used in earlier papers using a similar framework. In Gong and van Soest (2000), a measurement error is not included explicitly, but the wage error is allowed to correlate with the random preference term u_{rp}. The significant estimate of the covariance is then interpreted as an indication of measurement error. Here we make this more explicit.

(0,4,8,...,60). This gives 16 points for each individual: (h_j, y_j), $j = 0, ..., 15$, where y_j is after tax family income if the wife works $h_j = 4j$ hours per week. Earlier research has shown that the sensitivity of the results for the number of points is small (van Soest, 1995, and Gong and van Soest, 2000). We do not repeat this analysis in the current paper since there is no reason to expect a different result.

Fixed costs of working

Models without fixed costs of working, in which the utility function explains participation as well as hours worked, typically tend to underpredict the number of non-workers. Including fixed costs of working is one way to repair this. Fixed costs of working can also be seen as some gain to not working compared to all the other possibilities, which makes not working more attractive than working few hours per week. The level of the fixed costs may depend on individual and household characteristics Z. We model them as: $FC = Z'\delta$, where Z is a vector of individual and family characteristics.[3] In computing the values of the utility function, we then replace income y_j by $y_j - FC$ if $j > 0$, i.e., if the wife works. Since U is increasing with income, positive fixed costs decrease the utility of working compared to the utility of not working, and therefore decrease the probability of participation.

Fixed costs were used in a similar way by Callan and van Soest (1996) and Euwals and van Soest (1999). An alternative way to explain the lack of part-time jobs is given by Dickens and Lundberg (1993), Tummers and Woittiez (1991), and Van Soest et al. (1990), who model job offer probabilities for part time jobs. Van Soest (1995) uses disutilities of part-time jobs, reflecting search costs of jobs with irregular hours. These methods attain the same goal as the fixed costs: the model can reproduce both the participation rate and average hours worked. The choice between the three seems a matter of taste; we chose fixed cost because it is economically plausible and hardly complicates the model or the estimation procedure.

As explained above, the intuitive explanation why fixed costs are identified is the lack of observations with a small positive number of working hours. While this argument is valid for a restrictive specification of the utility function which limits the way in which utility can vary between points with zero, few, or many hours per week, the argument no longer holds if the specification of the utility function is fully nonparametric. For such a specification, the utility function itself could pick up the gap in the distribution at few hours, by assigning lower utility to such hours values. Thus it seems that the fixed costs are nonparametrically unidentified. In our specifications, the identification problem does not arise, due to the restrictive way in which the taste shifters are allowed to enter the utility function and the fixed costs. Still, this seems a rather unnatural way to obtain identification. The way to avoid this would be not to include fixed costs explicitly in the

---

[3]Unobserved heterogeneity in fixed costs can be allowed for by adding another error term. We experimented with this but it did not lead to any significant improvement.

nonparametric (higher polynomial expansion order) model, and to consider the utility function as an evaluation of preferences in which fixed costs are already captured. The reason why we have not done this is that it makes a fair comparison with the parametric (lower order) model impossible.

Alternative specific error terms

The only error terms included so far are random preferences. In addition, we introduce alternative specific error terms as follows:

$$u(h_j, y_j) = U(h_j, y_j) + \epsilon_j; \quad j = 0, \ldots, 15 \tag{4}$$

We assume that the ϵ_j are iid and follow an extreme value distribution. We assume that the answer to the desired hours question is based upon maximizing $u(h_j, y_j)$ rather than $U(h_j, y_j)$. The error ϵ_j can be seen as the error made in evaluating alternative j. There are several reasons why these errors are incorporated. First, they are needed to give nonzero probability to choices which cannot be optimal for any value of the random preference term. Such choices may very well exist in case of a nonconvex or discontinuous budget set, where some points on the budget frontier may give very low family income compared to adjacent points. In this sense, they play the same role as the optimization or measurement errors in the traditional Hausman (1985) model. Second, it is attractive to include the ϵ_j from a computational point of view: we will see below that they facilitate simulated maximum likelihood estimation by smoothing the approximation of this likelihood, and can thus be seen as a smoothing device. The same interpretation is given to them by Keane and Moffitt (1998). They use the same type of error terms, but, enforcing their interpretation, impose that the ϵ_j have a small variance compared to the variance in $U(h_j, y_j)$. We do not make such an assumption and thus allow for both interpretations. Due to the assumption on the distribution of the ϵ_j the resulting model shows some similarity to the multinomial logit model. The probability that an individual chooses alternative j, conditional on wage rates, tax and benefit rules, exogenous variables, and random preference parameters, is given by:

$$P[j] = exp\{U(h_j, y_j)\} / \sum_k exp\{U(h_k, y_k)\} \tag{5}$$

The summation in the denominator is over the 16 points in the choice set.

$P[j]$ increases with $U(h_j, y_j)$ (given the other $U(h_k, y_k)$). Since U is increasing in income, the utility of working increases with the (before and after tax) wage rate. The utility of not working is not affected by the wage rate. Thus the participation probability increases with the wage. This illustrates that the participation decision is fully incorporated in the structural model, other than in many labour supply studies which are conditional on participation.

9

Coherency

An important issue in the piecewise budget constraint model is coherency. If preferences are not quasi-concave in some relevant region, the model may not have a well-defined unique solution for a non-zero probability set of values of the error terms. This implies that probabilities typically used in likelihoods do not add up to one, and maximizing the likelihood can lead to inconsistent parameter estimates. Van Soest et al. (1993) give an example where the latter is indeed the case. They argue that coherency should be imposed *a priori* before estimation. MaCurdy et al. (1990) show that imposing coherency in the linear labour supply model implies that labour supply cannot be backward bending. Thus imposing coherency conditions limits the flexibility of the specification of preferences. Van Soest et al. (1993) show that quasi-concavity of preferences is sufficient but not necessary to guarantee coherency, and confirm the result that imposing coherency in a restrictive specification leads to bounds on the potential elasticities and policy effects. Blomquist (1995) shows that this problem is not unique to ML-estimation but also plays a role if another estimation technique (IV) is used.

In the current set up, there are two reasons why the coherency problem does not arise and conditions limiting flexibility can be avoided. First, the coherency problem in the Hausman (1985) model is due to the way of solving utility maximization using Kuhn-Tucker's rules, i.e., first order conditions. If preferences are not quasi-concave, there may be multiple solutions to these first order conditions. The multi-regime econometric model based upon the Kuhn-Tucker conditions may then have zero or more than one solutions. In our set up, however, we do not need to rely on tangency conditions or on other results of duality theory, since we simply choose the utility maximum from a finite set of utility values. Due to the assumption of continuously distributed ϵ_j, the probability that two points have optimal utility is zero, and with probability one, our method defines an unambiguous utility optimum. Thus the model has a unique solution with probability one and is coherent, whatever the shape of the utility function.

A second danger exists, however. An implication of the MaCurdy et al. (1990) critique is that a seemingly flexible functional form may not be flexible anymore once quasi-concavity or monotonicity are imposed. Even though we do not impose these conditions explicitly, it might still be the case that the structure of the model implicitly will enforce the estimates to satisfy quasi-concavity. For example, a wrongly shaped utility function would lead to high probabilities of choosing the corners of the budget frontier (0 or TE hours of work), and ML estimates will avoid this shape if the data points are typically not at these corners. Thus even if coherency is not imposed, the question whether quasi-concavity combined with functional form does not limit the range of elasticity values or policy effects remains relevant. In our case, however, we use a flexible functional form of the utility function. Even if we would impose quasi-concavity of preferences on some relevant region of (h, y) space, we would not impose more than that, because our series approximations are able to approximate any quasi-concave utility function arbitrarily. Thus the problem in MaCurdy et al. (1990) that imposing

coherency or quasi-concavity immediately bounds the range of possible elasticities – due to the restrictive functional form – will not occur.

We will not impose quasi-concavity but we will check *ex post* whether estimated preferences are quasi-concave. While this is strictly spoken not even necessary for the interpretation of the model (only utility in the finite choice sets matters), it would help to reconcile our findings with those of the Hausman approach. The only thing we really need for economic interpretation (and meaningful policy simulations) is that utility increases with income y. Since this is not necessary for coherency of the econometric model either, we will check the condition *ex post* without imposing it *a priori*.

Estimation

Due to the multinomial logit nature of the model, estimation by maximum likelihood would be straightforward if random preference terms were observed and all wages were observed without measurement error. In that case, the likelihood would follow directly from (2) and (5), since the $U(h_j, y_j)$ would then be known functions of parameters, explanatory variables, the observed wage rate, and the known random preference term. The likelihood contribution of a given individual would be her wage density (following from (2)) multiplied by the probability in (5). But we do not observe the error terms (including those in the wage equation. As a consequence, the likelihood contribution of a given observation is given by the mean value of the appropriate expression according to (2)) and (5)), with the mean taken over the unobserved errors. This mean is a two-dimensional integral. Such an integral can be approximated by conventional numerical (quadrature) routines, in principle. A convenient alternative which also works for dimensions higher than two, is simulated maximum likelihood: the integral is replaced by a simulated average based upon R independent draws from the (multivariate normal) distribution of the unobserved errors, conditional upon the observed wage rate. Due to the law of large numbers, the approximation will be accurate if R becomes large. With independent draws across observations, it can be shown that the approximation is accurate enough to make simulated maximum likelihood asymptotically equivalent to exact maximum likelihood if R tends to infinity faster than the square root of the number of observations (see, for example, Hajivassiliou and Ruud, 1994). We will use $R = 20$. The sensitivity of the results for the choice of R is analyzed by Gong and van Soest (2000), who find that $R = 20$ is large enough. We will obtain the same conclusion in our Monte Carlo study.

The simulated maximum likelihood procedure is greatly facilitated by the presence of the ϵ_j. Without these, the likelihood contribution conditional on the unobserved error terms would be either 0 or 1. The simulated likelihood would become a discontinuous function of the parameters, its maximization would be numerically much harder, and zero contributions would have to be dealt with. Adding the ϵ_j smooths the likelihood and bounds it away from zero. Adding the ϵ_j could thus be seen as a smoothing device, without giving the ϵ_j any real economic meaning. This is the interpretation of Keane

and Moffitt (1998). They fix the variance of the ϵ_j at some small value, and at the same time, impose a normalization on the systematic part of the utility function. This *a priori* limits the share of the variance of the ϵ_j in the total variance of $u(h_j, y_j)$. We normalize the variance of ϵ_j only, and do not impose an additional scale normalization on the utility function. This corresponds to the view that the ϵ_j could have some meaning as alternative specific errors in the economic model. We let the data decide how important this is. Apart from the normalization of the variance of ϵ_j we impose $\alpha(0,0) = 0$ to fix the location of the utility function.

3   Data and Estimation Results

Data

The data are drawn from May 1995 wave of the Dutch Socio-Economic Panel (SEP). This is a panel consisting of about 5,000 households, which are representative for the Dutch population excluding people living in nursing homes, etc. We focus on married or cohabiting women in the age group 16-64 whose partners have a paid job with observed earnings. We exclude women who are full-time students, receive full-time disability benefits, or receive pensions or other retirement benefits. This leads to a sample of 1794 women; 1100 of these have a paid job.

We tried to construct the dependent labour supply variable in such a way that demand side restrictions do not play a role, i.e., we want to use desired hours instead of actual hours worked. For people who are looking for a(nother) job, we know how many hours they would like to work in their new job. We consider this number rather than their actual hours as their desired labour supply. For those who are not looking for a(nother) job, however, we have no further information, and have to assume that desired hours are given by actual hours. Of the 694 non-workers in the sample, 116 are looking for a job, and thus have positive desired hours. Of the 1100 workers in the sample, 152 are looking for another job with desired hours different from actual hours.

Earnings in the SEP are measured as gross earnings in the year 1994, retrieved from the respondents' tax files. These earnings can only be used to compute an hourly wage rate for the job held at the time in the survey in May 1995 for people who have not changed jobs in 1994 or from January 1995 until May 1995. For those who did change jobs in that period, earnings are set to missing. This concerns 55 observations. Table A1 in the appendix contains the definitions of the variables we use in the analysis, and Table A2 presents some sample statistics. Non-participation among married women in the Netherlands is still substantial, although it has fallen substantially during the past two decades. In our sample (which excludes those who are not available for work such as students or disabled), the net participation rate (excluding those who are looking for work) is 60.9%, while the gross participation rate (including job searchers) is 67.3%.

Estimates

We have estimated the model for $K = 1$ to $K = 5$. In Tables A3 and A4 we present the results for $K = 2$ and $K = 5$. Most parameters in the utility function cannot be interpreted directly. The exceptions are the interactions between hours worked and characteristics, since these parameters determine how the marginal utility changes with characteristics. These results are similar for the two models. They show that age is insignificant, while the presence of children increases the disutility of hours worked, i.e., reduces labour supply. This effect is still stronger when there are children of pre-school age. The impact of children we find is in line with the bulk of the labour supply literature (see Nakamura and Nakamura, 1990, for example). Random preferences play virtually no role: σ_{rp} is small (compared to, for example, the variance of each alternative specific error term which is normalized at $\pi^2/6$), and the standard error of its estimate exceeds the point estimate.

We find that children reduce fixed costs, which seems counterintuitive. Combined with the effect of children on preferences, this finding could mean that for women with children, working a small number of hours per week is particularly attractive. Fixed costs fall with age until approximately age 47.

The wage equation estimates correspond to common findings in the human capital literature, with an increasing age pattern until about age 41, and higher wages for the higher educated. The estimates of the standard deviations σ_u and $\sigma_e u$ suggest that the labour supply decision is largely based on the predicted wage. Most of the unpredicted part of the wage is not used in the labour supply decision, for example because this is measurement error.

All the results discussed above are similar for the second order model and fifth order model. They also appear to be valid for the third and fourth order model (results available upon request from the authors). Only the first order estimates give a different picture, with, for example, an even much larger role of the measurement error e_w compared to u_w.

To choose between the models of different orders, we present their likelihoods and AIC values in Table 1. According to both the AIC criterion and likelihood ratio tests, the first order model is clearly outperformed by all other models. Differences among the higher order models are much smaller, but still, the fifth order model is a significant improvement compared to lower order models, and also yields the best AIC value. In the sequel, we will focus on the first, second and fifth order model, since all results for the intermediate third and fourth order models are in between the results for the second and fifth order model.

Table 1. Likelihood and AIC for different
specifications (real data)

specification	# parameters	likelihood	AIC
1st order	22	-4536.04	5.081
2nd order	25	-4235.75	4.750
3rd order	29	-4234.08	4.753
4th order	34	-4224.57	4.748
5th order	40	-4209.93	4.738

Figure 1 shows some estimated indifference curves for the first, second and fifth order models. Utility is increasing with income in almost all data points, although we did not impose this *a priori*. Utility is also usually increasing with leisure, except at low hours for the fifth order model.

The indifference curves for the first order model are by definition linear, so quasi-concavity is not an issue there. The quadratic indifference curves for the second order model imply quasi-concavity of preferences for all values (h, y). For the fifth order model, the indifference curves are convex in most of the relevant range of (h, y)-space, but not everywhere. In particular, quasi-concavity of preferences is violated at points in (h, y) space with high values of working hours.

Figure 2 presents labour supply curves, constructed in a similar way as traditional labour supply curves. Hours worked is drawn as a function of the woman's own wage rate. The curves concern a benchmark individual (age=40, non-female income=1000, 1 child aged less than 5, and random preference term set to zero). We show expected hours worked in case of a linear budget constraint (no fixed costs, no taxes or benefits), computed as the probability weighted mean of the possible hours values 0,4,8,...,60 in the model. The probabilities are computed from the model estimates (using (5)). Again, the first order model gives very different outcomes from the other models. These other models lead to labour supply curves which are similar to each other in the range of low wage rates, but which show some larger differences for high wage rates. For the second, third and fourth order model, labour supply is everywhere forward bending. For the fifth order model, labour supply is forward bending except at very high wage rate levels.

Elasticities

Figure 2 shows the sensitivity of labour supply to the wage rate for a benchmark individual, but since elasticities vary over the sample, this will not be the most interesting figure for policy analysis. For this purpose, we are more interested in aggregate elasticities. We define the (own) wage elasticity of labour supply of some given group of women as the percentage change in total desired hours of that group if all their before tax wage rates rise by 1%. Unlike some other elasticities used in the literature, this definition takes full account of the impact of the wage rate on the participation decision (with desired hours equal to zero for non-participants). We also show which share of the elasticities is

due to the effect on participation. [4] Our elasticities are based on increasing gross wage rates and leaving the tax system unaffected. Thus they correspond to what Blomquist (1996) calls the Mongrel labour supply function rather than the labour supply curves in Figure 2. Since the benchmark for policy analysis will be the actual tax system and not some linear system, this again seems the most relevant definition from a policy point of view. The way in which net wage rates change is endogenous. On average, they will change by slightly less than 1%, due to the progressive nature of the tax rules.

The elasticities can be computed from model simulations, first using actual (predicted) wage rates, and then with all wage rates increased by 1%. Standard errors are computed by deterministic bootstrapping, repeating the calculations for new parameter values of the model drawn from the estimated distribution of the parameter estimates. In a very similar way, we have also calculated the elasticity and the sensitivity of participation with respect to the husband's wage rate. Since the husband's earnings reflect the bulk of family income other than the woman's own earnings, these are approximately the same as other income elasticities.

The first four columns of Table 2 contain the results for each of the five models, for all women as well as separately for the low and high educated. In line with Figure 2, the elasticities according to the first order model deviate substantially from those according to the other models. The four higher order models, however, lead to similar elasticities, with overlapping confidence intervals. The own wage elasticity is somewhat above one. While this is not out of line with other findings for the Netherlands (cf. Theeuwes, 1988, for example) or other countries (Killingsworth and Heckman, 1986, Blundell and MaCurdy, 1999), it is somewhat higher than recent findings with similar models (cf. Van Soest and Das, 2000, and Vlasblom, 1998). The main reason seems to be that we have allowed for measurement error in the wage rates: if we set σ_e to zero, the own wage elasticity is about half as large. This difference is in line with the standard argument that measurement error biases the coefficients to zero if not properly accounted for. The cross-wage elasticity is about -0.16, in line with earlier findings for the other income elasticity (Theeuwes, 1988). The effects on participation presented in Table 2 are the changes in percentage points if husband's or own wage rates increase by 1%. [5] More than half of the cross-wage elasticity, and somewhat less than half of the own wage elasticity are due to an effect on participation.

Labour supply of low educated women is of particular interest from a policy point of view, since their participation rates are lower and their unemployment rates are higher than for other women. We find that the supply of labour for the low educated is somewhat more sensitive for both wage rates than for the high educated. Again, the various higher order models lead to the same conclusions here.

---

[4] We look at desired hours only and not at actual hours. Participation is defined as having positive desired hours (i.e., gross participation).

[5] The elasticities of participation can be obtained by dividing these numbers by the predicted participation rates, approximately 0.673 for the whole sample, 0.635 for the low educated, and 0.832 for the high educated.

Table 2. Changes in participation, elas. of hours, and policy effects

model	Husband's wage		Own wage		Policy effects	
	change in part. (%-points)	elas. of hrs (%)	change in part. (%-points)	elas. of hrs (%)	change in part. (%-points)	change in hrs (%)
5th order:						
all	-0.110	-0.176	0.532	1.259	-2.99	4.87
	(0.026)	(0.047)	(0.054)	(0.13)	(0.50)	(0.83)
low educated	-0.122	-0.222	0.561	1.344	-3.72	4.68
	(0.029)	(0.048)	(0.059)	(0.14)	(0.56)	(0.89)
high educated	-0.062	-0.051	0.411	1.027	0.01	5.41
	(0.022)	(0.074)	(0.041)	(0.11)	(0.35)	(0.99)
4th order:						
all	-0.099	-0.156	0.500	1.173	-2.87	4.62
	(0.028)	(0.047)	(0.047)	(0.12)	(0.46)	(0.65)
low educated	-0.111	-0.196	0.526	1.258	-3.57	4.48
	(0.031)	(0.051)	(0.051)	(0.13)	(0.53)	(0.70)
high educated	-0.051	-0.049	0.392	0.942	0.02	5.01
	(0.019)	(0.055)	(0.036)	(0.091)	(0.27)	(0.74)
3rd order:						
all	-0.113	-0.167	0.470	1.093	-2.70	4.23
	(0.026)	(0.040)	(0.045)	(0.11)	(0.47)	(0.54)
low educated	-0.126	-0.203	0.490	1.175	-3.38	4.10
	(0.029)	(0.047)	(0.049)	(0.12)	(0.54)	(0.59)
high educated	-0.062	-0.070	0.391	0.879	0.09	4.61
	(0.017)	(0.033)	(0.034)	(0.079)	(0.29)	(0.52)
2nd order:						
all	-0.105	-0.151	0.453	1.063	-2.56	4.27
	(0.020)	(0.034)	(0.043)	(0.10)	(0.34)	(0.55)
low educated	-0.115	-0.180	0.471	1.139	-3.20	4.19
	(0.022)	(0.038)	(0.046)	(0.11)	(0.38)	(0.59)
high educated	-0.063	-0.069	0.383	0.854	0.05	4.50
	(0.013)	(0.029)	(0.033)	(0.071)	(0.25)	(0.48)
1st order:						
all	-0.046	-0.089	0.571	2.062	-1.87	13.70
	(0.0050)	(0.0094)	(0.059)	(0.20)	(0.36)	(1.40)
low educated	-0.051	-0.107	0.589	2.229	-2.50	15.09
	(0.0056)	(0.012)	(0.064)	(0.23)	(0.40)	(1.62)
high educated	-0.028	-0.043	0.494	1.645	0.68	10.23
	(0.0031)	(0.0051)	(0.044)	(0.14)	(0.29)	(0.99)

Simulations are based upon real data.

Bootstrapped standard errors are in parentheses.

Proposed Tax Reform

To illustrate the usefulness of the fully structural model, we have analyzed the potential consequences of a recently proposed tax reform, in which the nonconvexities in the budget set close to zero hours of work play a large role. Such a reform can therefore not be analyzed using the convex budget constraint model of Blomquist and Newey (1997),

or other models not considering the participation decision or not allowing for nonconvex budget sets.

We briefly describe the tax system and the proposed reform. More details are given in van Soest and Das (2000). The budget constraints for a benchmark value of the husband's income (working full time with earnings equal to three times the minimum wage) are presented in Figure 3. The current system has individual taxation of the two spouses, but with one joint feature: the length of the first tax bracket (with a rate of 0%) depends upon earnings of the spouse. If both spouses work and both earn more than Dfl 8,600, then the tax free allowance for both is Dfl 8,600. If the wife has no own income, the husband's tax free allowance is Dfl 16,800, i.e. the wife's tax free allowance is transferred. If the husband earns more than Dfl 8,600, but the wife earns less than Dfl 8,600, the wife can (and, in general, will) largely transfer her allowance to the husband, so that her own tax free allowance becomes Dfl 400 and her husband's allowance will be Dfl 16,800. The transfer possibility creates a disincentive for the woman to earn more than Dfl 8,600 if the husband's earnings are high. This is shown by the solid curve in Figures 3, which gives net family income as a function of the wife's hours of work. There is a dip when the wife's earnings attain the maximum transfer threshold.

In a recent report (Ministry of Finance, 1997), the main ideas are sketched for a complete reform of many features of the tax system of the Netherlands. The main reforms concern increasing taxes on polluting activities, changing some of the VAT rates, and reducing taxes on labour. The latter implies changing the income tax system for private households. The report contains 21 specific proposals for income tax revisions. We only look at the most radical type of reforms, which involves abolishing tax free allowances for two earner families. [6] Only in genuine one earner families, the tax free allowance of the only earner would be increased. Additional tax revenues are used to lower the marginal tax rates, so that the revision as a whole (also accounting for changes in other taxes) would be revenue neutral.

The budget constraint for the benchmark family after the reform is given by the dotted line in Figure 3. There is a discontinuity at 0 hours of work: as soon as the woman starts working, the additional tax free allowance of the husband is lost, and family income falls. Thus the proposed reform creates a disincentive for women to take a small part-time job.

The estimated effects of the policy reform are presented in the final columns of Table 2. Again, the higher order models give similar results, but the first order model does not. Since many women who now have a small part-time job would decide to stop working after the reform, participation would fall. This effect is concentrated in the group of low educated women. On the other hand, net earnings of full-time workers increase, and labour supply is stimulated in the sense that there is a substantial increase in hours

---

[6]We consider the basic version of option 3 in Ministry of Finance (1997). The final proposal which has gone to parliament (Ministry of Finance, 1999) is more refined and avoids the problem with small part-time jobs addressed below.

worked, in spite of the negative effect on participation. Average hours (including zeros) would increase by between 4 and 5%.

4   Monte Carlo Simulations

The goal of this section is threefold. First, we want to establish the finite sample properties of the estimator we have used if the model is correctly specified. There are two reasons here why there could be a problem. First, the model is nonlinear and fairly complicated, and the finite sample bias of maximum likelihood in this type of model is an open issue. Second, we do not use exact maximum likelihood but simulated maximum likelihood, approximating the likelihood using $R = 20$ draws for each observation (with independent draws across observations). While there is some evidence that the results are not very sensitive for the chosen value of R for this type of model, this, at least in theory, remains a potential source of bias for every new version of the model or every new dataset.

The second purpose of the Monte Carlo simulations is to analyze the consequences of using a too restrictive model for the estimates of the elasticities of interest and of the policy effects. In the previous section, we have seen that the fifth order model is preferred to any of the lower order models using the AIC criterion or likelihood ratio tests. Still, using the real data, we did not find much difference between elasticities according to the various higher order models, suggesting that a second order model would be sufficient to get reasonable estimates of the parameters of interest. We want to confirm this finding in a situation where we know that the order of the series expansion is the only source of misspecification.

Third, the issue of measurement error in hours worked has until now not been addressed. In the original Hausman model, measurement error is explicitly incorporated, but in our model, it is not. We want to check whether the presence of measurement error in hours worked can bias the results for the parameters of interest in our model.

We used the exogenous variables in the sample and the estimates of the fifth order model in Table A4, to generate 100 new data sets. Using equation (3), wage rates (including measurement errors) are generated for all observations. Using equation (5) (and errors drawn from the GEV 1 distribution), optimal hours are then generated (i.e., the utility maximizing element of $\{0, 4, ..., 60\}$ is determined). Wage rates of non-workers are then set to missing. The 100 new data sets are generated independently (conditional on the exogenous variables), and each of them consists of 1794 (independent) observations.[7] For each of the 100 new data sets, we estimated the fifth order and the second

---

[7]We do not generate actual hours or involuntary unemployment, and we do not generate missing wage rates for workers. Thus for all those in the new data sets who want to work, the wage rate is observed. In the actual data, there are some workers with unobserved wage rate and some job searchers with known desired hours but unknown wage rate. Such observations will reduce the accuracy of the model estimates. As a consequence, the Monte Carlo simulations may somewhat overestimate the

order model. (Measurement errors on hours worked will be discussed below.) Each of these estimations required numerical optimization of a simulated likelihood function, and thus requires substantial computational effort. This is why we have restricted the Monte Carlo study to 100 new data sets, which is rather low compared to other Monte Carlo studies.[8] We believe, however, that the results are stable enough not to expect that a larger number would lead to very different results.[9] Based upon each set of estimates, we repeated the calculations of elasticities and policy effects as presented in Table 2. In Table 3, we summarize these results. We focus on the sample as a whole, and do not distinguish between education levels. For each elasticity or policy effect, we present the mean and the "sample" standard deviation over the 100 new data sets.

The right hand column refers to the fifth order model. This can be used to answer the first question: is there a serious bias in the simulated ML estimates, either due to the small sample or due to the small value of R? We know that the true elasticities and policy effects are their point estimates for the fifth order model in the first row of Table 2. The result is striking: the numbers in the second column of Table 3 are quite close to those in the first row of Table 2. There is no evidence of any systematic bias in the point estimates of the elasticities. There is a somewhat larger difference in the estimate of the policy effect. Why the policy effects are harder to estimate than the elasticities is not quite clear; the former is probably a still more complicated function of the parameters of the model than the latter.

---

accuracy of results that can be obtained with the real data.

[8]Blomquist (1996), for example, uses 400 replications. He has much simpler models (with fewer parameters), and only 602 observations, however.

[9]We also experimented with 75 new data sets, and obtained very similar conclusions.

Table 3. Results of a Monte-Carlo study

	Estimated model	
	2nd order	5th order
wage effects		
husband's wage		
- change in participation (in %-points)	-0.110	-0.115
	(0.025)	(0.027)
- elasticity of hours worked (in %)	-0.155	-0.178
	(0.045)	(0.049)
own wage		
- change in participation (in %-points)	0.471	0.524
	(0.048)	(0.058)
- elasticity of hours worked (in %)	1.11	1.23
	(0.12)	(0.14)
policy effects		
- change in participation (in %-points)	-2.27	-2.63
	(0.40)	(0.50)
- change in hours worked (in %)	4.79	4.95
	(0.78)	(0.96)
likelihood	-4332.08	-4303.55
	(51.5)	(51.2)
AIC	4.86	4.84
	(0.057)	(0.057)

Note: 100 data sets are generated using $5^{t}h$ order model
estimates (cf. Table A4). Means and standard deviations
(in parentheses) over these 100 are reported.

The standard deviations in the second column of Table 3 can be compared to the estimated standard errors in the first row of Table 2. Standard errors in Table 2 are derived from the estimated asymptotic distribution of the (simulated) maximum likelihood estimates of the parameters in the model. They could be inaccurate due to finite sample bias, due to too small choice of R,[10] or due to model misspecification.[11] Comparing them to the standard deviations in Table 3 suggests that the standard errors in Table 2 are somewhat underestimated. Again however, the differences for the elasticities are quite small. For the policy effect on hours worked, the difference is somewhat larger, but still quite modest. The conclusion can only be that the performance of the simulated

---

[10]We have estimated the standard errors using the common ML technique, thus implicitly assuming that R tends to infinity at a rate of at least the square root of the number of observations. This is the condition under which ML and simulated ML are asymptotically equivalent (Hajivassiliou and Ruud, 1994, p. 2419).

[11]We have used the outer product estimator of the information matrix; more robustness to misspecification might be obtained by also using the Hessian.

ML estimator in case of no misspecification appears to be remarkably good. We find no evidence of either a finite sample bias, or a bias due to small R.

The first column of Table 3 shows the bias due to using a second order specification of preferences rather than the fifth order specification. The differences with column 3 and with the true values given in Table 2 are rather small. They are quite similar to the differences obtained when the real data are used, see the fourth panel in Table 2. Thus although the fifth order model is the true model and the second order model is misspecified, the bias on the elasticities and the policy effects induced by this misspecification is very limited. This conclusion obviously might be specific to our empirical example, where preferences apparently can be approximated by a second order utility function reasonably well.

We repeated this exercise for the first, third and fourth order model. Results are not presented to save space, but they all confirm the findings in Table 2. The third and fourth order model yield similar results as the fifth and second order model. The first order model, however, appears to be too restrictive for the data at hand, and leads to grossly biased estimates of elasticities and policy effects. This should not come as a surprise of course, since a linear utility function is extremely restrictive.

The bottom panel of Table 3 shows the average log likelihood and AIC criterion values and their "sample" standard deviations over the 100 new data sets. On average, the likelihood of the fifth order model is substantially larger than that of the second order model. It appears that for 99 out of the 100 new data sets, a likelihood ratio test rejects the second order model in favour of the fifth order model. Similarly, although the difference is always small, the AIC value of the second order model is worse (i.e., larger) than the AIC value of the fifth order model in (the same) 99 data sets. This confirms the conclusion from Table 1, that from a statistical point of view, the fifth order model outperforms the second order model (and all other models). From an economic point of view, however, the conclusion is more or less opposite: both Table 2 and Table 3 show that the second order model is already quite capable of reproducing the relevant elasticities and policy effects.

Measurement errors in hours worked

The traditional Hausman model has two types of errors in the labour supply equation: random preferences, and optimization or measurement (of hours worked) error. In our labour supply model, random preferences are incorporated, and the GEV I errors could be seen as alternative specific utility evaluation errors, i.e., a form of optimization error. They cannot be seen as measurement error on (desired) hours worked, however. To investigate whether neglecting measurement error on hours worked could bias the results, we have generated new data sets including such measurement error, and re-estimated the model with these new data.

It is not clear what would be a reasonable size of the measurement error. Estimates of the Hausman model do not lead to common findings on this point. Since different

questions are used in different surveys (annual hours versus weekly hours, for example), there is not much reason to expect that measurement errors are always of the same size. In the simulations below, we used measurement errors with mean zero and standard deviations 2 and 4 (hours per week).

The new data sets including measurement errors on hours worked are constructed from the data sets used for the other Monte Carlo simulations. First, hours without measurement error are generated (as discussed above), then a measurement error drawn from $N(0, 4)$ or $N(0, 16)$ is added, then hours are again rounded to a multiple of 4. As before, we only considered the second and fifth order models. The results are presented in Table 4. Again, the elasticities and policy effects can be compared with the true values used to generate the data, given in the top panel of Table 2. They can also be compared to the figures in Table 3, based upon the same 100 new data sets without measurement errors.

We find that the wage elasticities of hours worked are not affected very much by the measurement errors. This result we would intuitively expect in a linear model, where measurement error on the dependent variable does not affect consistency (although it reduce efficiency).

The results for the effects on participation, however, are less in line with what a linear model would predict. If the measurement error is substantial (with standard deviation 4 hours per week), the estimated effects on participation of changing the women's or their husbands' wages, are strongly biased towards zero, with a bias of more than 50% of the absolute value of the true effect. This is not due to less accuracy: the standard deviations do not exceed those in Table 3. A closer look at the parameter estimates suggests that this problem is related to identifying fixed costs. Fixed costs estimates for the data with measurement error, lead to a large number of people with negative predicted fixed costs, while fixed costs were almost always positive according to the earlier estimates.

Table 4. Monte-Carlo with measurement err. in hours worked

	Std. deviation=2		Std. deviation=4	
	Estimated model		Estimated model	
	2nd order	5th order	2nd order	5th order
wage effects				
husband's wage	-0.085	-0.060	-0.044	-0.029
- change in part. (%-points)	(0.019)	(0.025)	(0.010)	(0.019)
	-0.145	-0.152	-0.127	-0.119
- elas. of hours worked (%)	(0.041)	(0.052)	(0.040)	(0.046)
own wage	0.329	0.320	0.195	0.178
- change in part. (%-points)	(0.039)	(0.042)	(0.024)	(0.026)
	1.11	1.08	1.04	0.94
- elas. of hours worked (%)	(0.14)	(0.15)	(0.14)	(0.14)
policy effects				
- change in part. (%-points)	-3.28	-3.06	-2.63	-2.17
	(0.53)	(0.63)	(0.51)	(0.52)
- change in hours worked (%)	7.94	8.13	10.2	9.89
	(1.11)	(1.27)	(1.41)	(1.45)
likelihood	-4417.78	-4393.69	-4362.11	-4339.74
	(46.1)	(47.8)	(44.3)	(46.2)
AIC	4.95	4.94	4.89	4.88
	(0.051)	(0.053)	(0.049)	(0.051)

Note: 100 data sets are generated using 5th order model estimates
(cf. Table A4). Means and standard deviations (in parentheses) over these
100 are reported; measurement error is added to hours worked with
either standard deviation equal to 2 or standard deviation equal to 4.

The policy effect is another nonlinear and complicated function of the parameters. It is related to the elasticities in the sense that both are driven by the sensitivity of labour supply to financial incentives. Still, for the policy effects, we find results which are not in line with the results for the wage changes. While the true policy effect on participation is reproduced reasonably well, the policy effect on hours worked is grossly overestimated (by more than 100% in the large measurement error case).

5 Conclusions

In this paper we have shown how nonparametric features can be built into a fully structural econometric model which is useful for policy analysis. We have combined a framework with a rich economic structure (utility maximization under a complex budget constraint) with a nonparametric specification of the key element in this framework (the direct utility function). We have taken a static neo-classical labour supply model, one of the most popular structural models in (cross-section) econometrics over the past two

decades. We have replaced the direct utility function by a flexible polynomial expansion, able to accurately approximate any utility function in a given compact set of relevant hours income combinations. Using the direct utility function only is made possible by treating the labour supply decision as a discrete choice problem, approximating the budget frontier by a finite set. The model also deals with many other problems in the structural labour supply literature, like nonparticipation, fixed costs of working, unobserved wage rates, measurement error in wage rates, and model coherency.

Using data for women with partner in the Netherlands, we use the framework to investigate female labour supply elasticities and the consequences of a recently proposed reform of income taxes which implies disincentives for small part-time jobs and incentives to work full-time. We find that the fifth order model statistical outperform the lower order models, but the second order model already gives almost the same outcomes as the fifth order model for the elasticities and the policy effects.

We then proceed with some Monte Carlo simulations. The most striking result is that if the model is correctly specified, the performance of our estimator is excellent: there is no finite sample bias on the elasticities and policy effects, and the asymptotic standard errors are very good approximations of the true standard errors. This finding seems in contrast with many earlier Monte Carlo Studies (such as Blomquist, 1996), though it should be realized that our sample size may be relatively large (1794 observations, versus 602 in Blomquist (1996)). The second finding confirms that with the real data: the true model used in the Monte Carlo simulations (the fifth order model as estimated using the real data), deviates enough from lower order models to reject these lower order models using LR tests or AIC criterion values in almost all Monte Carlo replications. On the other hand, every model of order at least two is a good enough approximations to reproduce the policy parameters of interest quite accurately. This means that from an economic point of view, it seems sufficient to use a second order model. Whether this conclusion remains valid in other situations for different data sets, obviously remains to be investigated. The final set of Monte Carlo simulations gives a less positive image: measurement error on the hours of work variable can seriously bias the estimates of policy effects, although they do not affect the wage elasticities estimates of hours worked.

The model considered here is obviously just one example, and it is kept as simple as possible. Numerous directions of extension exist. Some of these can straightforwardly be incorporated in the current framework. The economic model can be extended to joint decision making of husband and wife, gross wage rates varying with hours worked, cost of child care. Econometric extensions include allowing for more flexible distributions of error terms in the wage equation or in random preferences, or explicitly allowing for measurement error. The current results give some confidence that even with a limited number of observations such extensions could already be useful.

References

Blomquist, N. (1983), "The effect of income taxation on the labour supply of married men in Sweden," *Journal of Public Economics* 22, 169-197.

Blomquist, N. (1995), "Restrictions in labor supply estimation: is the MaCurdy critique correct?," *Economics Letters* 47, 229-235.

Blomquist, N. (1996), "Estimation methods for male labour supply functions: how to take account of non-linear taxes," *Journal of Econometrics* 70, 383-405.

Blomquist, N. and W. Newey (1997), "Nonparametric estimation of labor supply functions generated by piece wise linear budget constraints," *Working paper* no. 1997-24, Uppsala University.

Blundell, R. and T. MaCurdy (1999), "Labor supply: a review of alternative approaches," *Handbook of Labor Economics* Vol. 3A, O. Ashenfelter and D. Card (eds.), North-Holland, Amsterdam, 1559-1695.

Blundell, R. and I. Walker (1986), "A life cycle consistent empirical model of labour supply using cross-section data," *Review of Economic Studies* 53, 539-558.

Burtless, G. and J. Hausman (1978), "The effect of taxes on labor supply," *Journal of Political Economy* 86, 1103-1130.

Callan, T. and A. van Soest (1996), "Family labour supply and taxes in Ireland, mimeo," Tilburg University.

Dickens, W. and S. Lundberg (1993), "Hours restrictions and labor supply," *International Economic Review* 34, 169-192.

Euwals, R. and A. van Soest (1999), "Desired and actual labour supply of unmarried men and women in the Netherlands," *Labour Economics* 6, 95-118.

Gallant, R. and D. Nychka (1987), "Semi-nonparametric maximum likelihood estimation", *Econometrica*, 55, 363-390.

Gong, X. and A. van Soest (2000), "Female labour supply and family structure in Mexico City," *mimeo*, Tilburg University.

Hajivassiliou, V. and P. Ruud (1994), "Classical estimation methods for LDV models using simulation," in: R. Engle and D. McFadden (eds.), *Handbook of Econometrics* Vol. IV, North-Holland, New York, 2384-2443.

Hausman, J. (1981), "Labor supply", in *How taxes affect economic behavior*, Brookings Institution, Washington DC.

Hausman, J. (1985), "The Econometrics of nonlinear budget sets," *Econometrica* 53, 1255-1283.

Heckman, J. (1979), "Sample selection bias as a specification error," *Econometrica* 47, 153-161.

Heckman, J. and B. Singer (1984), "A method for minimizing the impact

of distributional assumptions in econometric models for duration data," *Econometrica* 52, 271-320.

Keane, M. and R. Moffitt (1998), "A structural model of multiple welfare program participation and labor supply," *International Economic Review* 39, 553-589.

Killingsworth, M. and J. Heckman (1986), "Female labor supply: a survey," in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics* Vol. 1, North-Holland, Amsterdam.

MaCurdy, T., D. Green, and H. Paarsch (1990), "Assessing empirical approaches for analyzing taxes and labor supply," *Journal of Human Resources* 25, 415-490.

Ministry of Finance (1997), *Belastingen in de 21-ste eeuw: een verkenning*, SDU Publishers, The Hague.

Ministry of Finance (1999), *Wet inkomstenbelasting 2001*, SDU Publishers, The Hague.

Moffitt, R. (1984), "Estimation of a joint wage-hours labor supply model," *Journal of Labor Economics* 2, 550-556.

Moffitt, R. (1986), "The econometrics of piecewise linear budget constraints: a survey and exposition of the maximum likelihood method," *Journal of Business and Economic Statistics* 4, 317-327.

Moffitt, R. (1990), "The econometrics of kinked budget constraints," *Journal of Economic Perspectives*, 4, 119-139.

Nakamura, A. and M. Nakamura (1990), "The econometrics of female labor supply and children," *Econometric Reviews* 11, 1-71.

Pollak, R. and T. Wales (1979), "Welfare comparisons and equivalence scales," *American Economic Review*, Papers and Proceedings 69, 216-221.

Theeuwes, J. (1988), "Arbeid en Belastingen," in *Belastingheffing en Belastinghervorming, Koninklijke Vereniging voor Staathuishoudkunde*, Stenfert Kroese, Leiden, 111-143.

Tummers, M. and I. Woittiez (1991), "A simultaneous wage and labour supply model with hours restrictions," *Journal of Human Resources* 26, 393-423.

Van Soest, A. (1995), "Discrete choice models of family labor supply," *Journal of Human Resources* 30, 63-88.

Van Soest, A., I. Woittiez and A. Kapteyn (1990), "Labour supply, income taxes and hours restrictions in the Netherlands," *Journal of Human Resources* 25, 517-558.

Van Soest, A., P. Kooreman and A. Kapteyn (1993): "Coherency and regularity of demand systems with equality and inequality constraints," *Journal of Econometrics* 57, 161-188.

Van Soest, A. and M. Das (2000), "Family labour supply and proposed tax

reforms in the Netherlands," *CentER Discussion Paper* nr. 20, Tilburg University.

Vlasblom, J.D. (1998), "Differences in labour supply and income of women in the Netherlands and the Federal Republic of Germany," PhD thesis, University of Utrecht.

Figure 1. Indifference curves benchmark individual
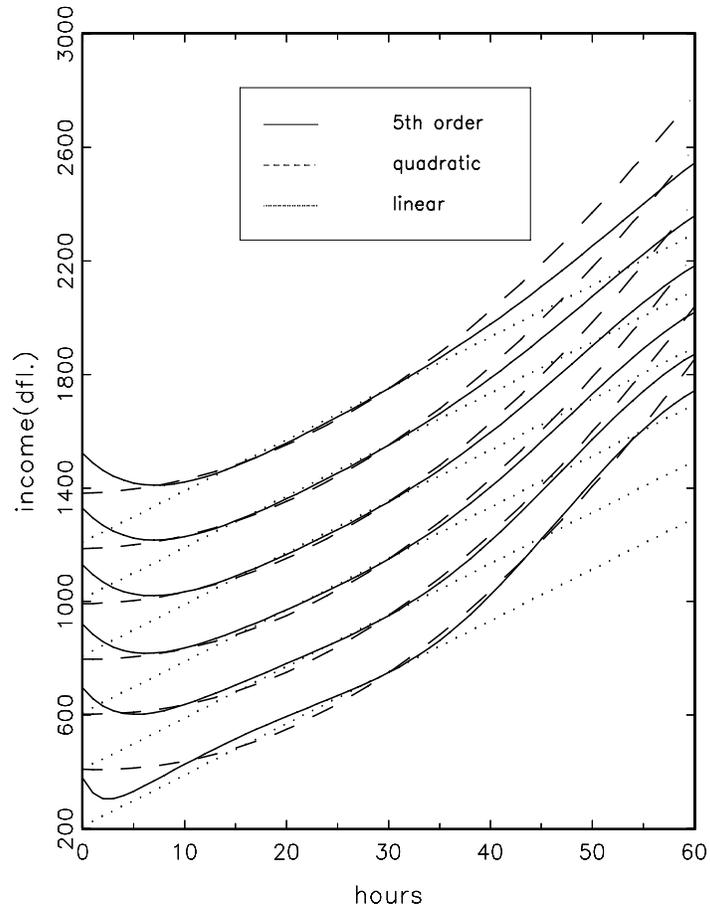
Figure 2. Labour supply curves benchmark individual
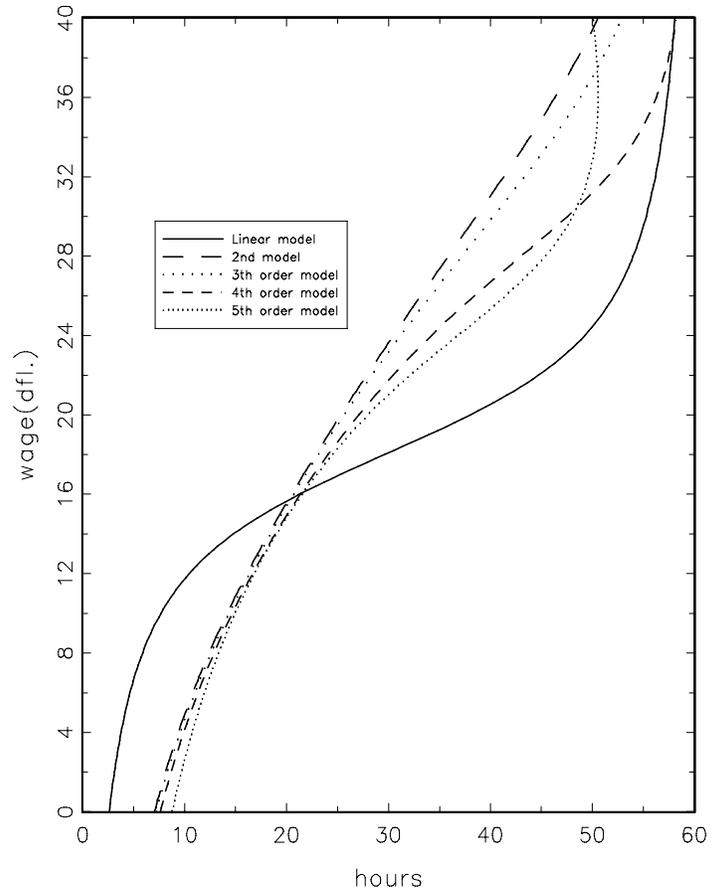
Figure 3. After tax family income as a function of the hours worked by the wife. The husband earns three times the minimum wage and the wife's hourly wage is equal to 1.5 times the minimum (hourly) wage.
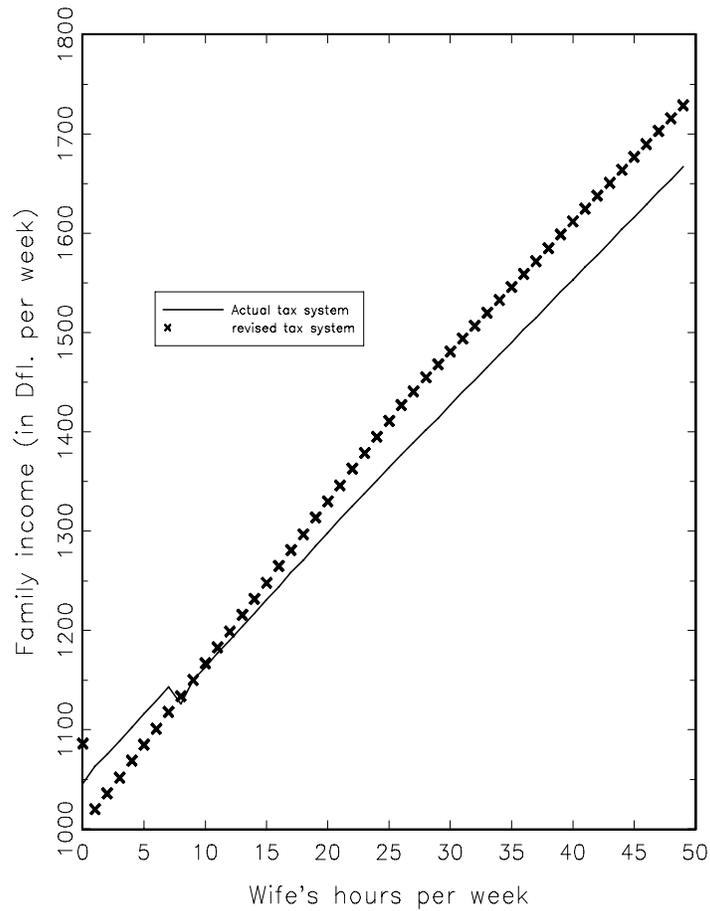
Appendix:

<div align="center">

Table A1. Definitions of variables

</div>

w_m	gross hourly wage rate husband
h_m	actual working hours husband (per week)
w	gross hourly wage rate wife
h_a	actual working hours wife (per week)
h	desired working hours wife (per week)
oth	other income husband and wife
y_{gross}	gross family income (excluding child benefits)
	$= w_m * h_m + w * h + oth$
y	net family income (including child benefits) in thousands
$dedu1$	1 if wife has primary education only (ref. category:).
$dedu2$	1 if wife has lower vocational trianing; 0 otherwise.
$dedu3$	1 if wife has intermediate vocational training; 0 otherwise.
$dedu4$	1 if wife has highschool training; 0 otherwise.
$dedu5$	1 if wife has higher vocational training; 0 otherwise.
$dedu6$	1 if wife has university degree; 0 otherwise.
age	age of wife (in years)
$nkid18$	number of kids with age 18 or less
$dkid05$	1 if kid with age less than 5 is present
$chbf$	child benefits (per week)

<div align="center">

Table A2. Summary statistics

</div>

Variable	Mean	Standard dev.
w_m	31.44	12.44
h_m	41.71	9.44
w	13.77	13.97
h	14.54	15.14
y_{gross}	1658.7	715.4
- female earnings	334.59	396.6
- male earnings	1296.83	569.7
- oth	27.29	118.3
$dedu2$	0.269	
$dedu3$	0.425	
$dedu4$	0.150	
$dedu5$	0.021	
$dedu6$	0.026	
age	38.22	8.87
$nkid18$	1.19	1.17
$dkid05$	0.278	
$chbf$	50.80	47.45

Table A3. Estimates of the 2nd order model

Parameters	Estimates	
utility function:		
y^2	0.839	(1.74)
$(h/10)^2$	-0.447	(-18.6)
$y*(h/10)$	-0.156	(-2.72)
y	10.5	(6.70)
$(h/10)$	3.32	(5.01)
$(h/10)*nkid18$	-0.482	(-11.8)
$(h/10)*dkid05$	-0.545	(-5.83)
$(h/10)*(age/10)$	-0.469	(-1.34)
$(h/10)*(age/10)^2$	-0.009	(-0.20)
σ_{rp}	0.0298	(0.29)
fixed costs:		
$constant$	0.759	(4.40)
$nkid18$	-0.0383	(-4.34)
$dkid05$	-0.0139	(-0.73)
$age/10$	-0.220	(-2.77)
$(age/10)^2$	0.0235	(2.45)
wage equation		
$constant$	1.84	(7.75)
$age/10$	0.490	(4.15)
$(age/10)^2$	-0.0597	(-3.98)
$dedu2$	0.0645	(2.40)
$dedu3$	0.257	(10.6)
$dedu4$	0.480	(14.0)
$dedu5$	0.617	(11.0)
$dedu6$	0.436	(6.06)
σ_u^w	0.183	(6.46)
σ_e^w	0.408	(29.63)

$t-$values are in the parentheses

Table A4. Estimates of the 5th order model

Parameters	Estimates	
utility function:		
y^5	0.0261	(0.08)
$y^4 * (h/10)$	-0.0381	(-0.22)
$y^3 * (h/10)^2$	-0.111	(-0.76)
$y^2 * (h/10)^3$	-0.0697	(-1.36)
$y * (h/10)^4$	-0.0182	(-0.84)
$(h/10)^5$	0.0228	(3.71)
y^4	0.0985	(0.03)
$y^3 * (h/10)$	0.663	(0.55)
$y^2 * (h/10)^2$	1.42	(1.76)
$y * (h/10)^3$	0.423	(1.47)
$(h/10)^4$	-0.336	(-3.98)
y^3	-3.58	(-0.37)
$y^2 * (h/10)$	-4.39	(-1.25)
$y * (h/10)^2$	-4.44	(-2.59)
$(h/10)^3$	1.69	(3.63)
y^2	15.3	(1.01)
$(h/10)^2$	-2.49	(-1.77)
$y * (h/10)$	5.08	(2.08)
y	-6.84	(-0.60)
$(h/10)$	1.46	(0.54)
$(h/10) * nkid18$	-0.492	(-11.5)
$(h/10) * dkid05$	-0.510	(-5.22)
$(h/10) * (age/10)$	-0.296	(-0.77)
$(h/10) * (age/10)^2$	-0.0281	(-0.57)
σ_{rp}	0.0106	(0.10)

$t-$values are in the parentheses

Table A4. (continued)

Parameters	Estimates	
fixed costs:		
$constant$	0.715	(4.40)
$nkid18$	-0.0300	(-3.96)
$dkid05$	-0.0045	(-0.28)
$age/10$	-0.167	(-2.39)
$(age/10)^2$	0.0184	(2.18)
wage equation:		
$constant$	1.91	(7.78)
$age/10$	0.464	(3.66)
$(age/10)^2$	-0.0565	(-3.51)
$dedu2$	0.0508	(1.33)
$dedu3$	0.231	(5.99)
$dedu4$	0.437	(8.87)
$dedu5$	0.564	(8.44)
$dedu6$	0.410	(5.48)
σ_u^w	0.165	(5.68)
σ_e^w	0.414	(32.1)

$t-$values are in the parentheses

IZA Discussion Papers

No	Author(s)	Titel	Area	Date
111	V. Sorm K. Terrell	Sectoral Restructuring and Labor Mobility: A Comparative Look at the Czech Republic	1/4	2/00
112	L. Bellmann T. Schank	Innovations, Wages and Demand for Heterogeneous Labour: New Evidence from a Matched Employer-Employee Data-Set	5	2/00
113	R. Euwals	Do Mandatory Pensions Decrease Household Savings? Evidence for the Netherlands	3	2/00
114	G. Brunello A. Medio	An Explanation of International Differences in Education and Workplace Training	2	2/00
115	A. Cigno F. C. Rosati	Why do Indian Children Work, and is it Bad for Them?	3	2/00
116	C. Belzil	Unemployment Insurance and Subsequent Job Duration: Job Matching vs. Unobserved Heterogeneity	3	2/00
117	S. Bender A. Haas C. Klose	IAB Employment Subsample 1975-1995. Opportunities for Analysis Provided by the Anonymised Subsample	7	2/00
118	M. A. Shields M. E. Ward	Improving Nurse Retention in the British National Health Service: The Impact of Job Satisfaction on Intentions to Quit	5	2/00
119	A. Lindbeck D. J. Snower	The Division of Labor and the Market for Organizations	5	2/00
120	P. T. Pereira P. S. Martins	Does Education Reduce Wage Inequality? Quantile Regressions Evidence from Fifteen European Countries	5	2/00
121	J. C. van Ours	Do Active Labor Market Policies Help Unemployed Workers to Find and Keep Regular Jobs?	4/6	3/00
122	D. Munich J. Svejnar K. Terrell	Returns to Human Capital under the Communist Wage Grid and During the Transition to a Market Economy	4	3/00
123	J. Hunt	Why Do People Still Live in East Germany?	1	3/00
124	R. T. Riphahn	Rational Poverty or Poor Rationality? The Take-up of Social Assistance Benefits	3	3/00
125	F. Büchel J. R. Frick	The Income Portfolio of Immigrants in Germany - Effects of Ethnic Origin and Assimilation. Or: Who Gains from Income Re-Distribution?	1/3	3/00
126	J. Fersterer R. Winter-Ebmer	Smoking, Discount Rates, and Returns to Education	5	3/00

181	E. Wasmer Y. Zenou	Space, Search and Efficiency	2	8/00
182	M. Fertig C. M. Schmidt	Discretionary Measures of Active Labor Market Policy: The German Employment Promotion Reform in Perspective	6	8/00
183	M. Fertig C. M. Schmidt	Aggregate-Level Migration Studies as a Tool for Forecasting Future Migration Streams	1	8/00
184	M. Corak B. Gustafsson T. Österberg	Intergenerational Influences on the Receipt of Unemployment Insurance in Canada and Sweden	3	8/00
185	H. Bonin K. F. Zimmermann	The Post-Unification German Labor Market	4	8/00
186	C. Dustmann	Temporary Migration and Economic Assimilation	1	8/00
187	T. K. Bauer M. Lofstrom K. F. Zimmermann	Immigration Policy, Assimilation of Immigrants and Natives' Sentiments towards Immigrants: Evidence from 12 OECD-Countries	1	8/00
188	A. Kapteyn A. S. Kalwij A. Zaidi	The Myth of Worksharing	5	8/00
189	W. Arulampalam	Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages	3	8/00
190	C. Dustmann I. Preston	Racial and Economic Factors in Attitudes to Immigration	1	8/00
191	G. C. Giannelli C. Monfardini	Joint Decisions on Household Membership and Human Capital Accumulation of Youths: The role of expected earnings and local markets	5	8/00
192	G. Brunello	Absolute Risk Aversion and the Returns to Education	5	8/00
193	A. Kunze	The Determination of Wages and the Gender Wage Gap: A Survey	5	8/00
194	A. Newell F. Pastore	Regional Unemployment and Industrial Restructuring in Poland	4	8/00
195	F. Büchel A. Mertens	Overeducation, Undereducation, and the Theory of Career Mobility	5	9/00
196	J. S. Earle K. Z. Sabirianova	Equilibrium Wage Arrears: A Theoretical and Empirical Analysis of Institutional Lock-In	4	9/00
197	G. A. Pfann	Options to Quit	1	9/00

198	M. Kreyenfeld C. K. Spiess G. G. Wagner	A Forgotten Issue: Distributional Effects of Day Care Subsidies in Germany	3	9/00
199	H. Entorf	Rational Migration Policy Should Tolerate Non-Zero Illegal Migration Flows: Lessons from Modelling the Market for Illegal Migration	1	9/00
200	T. Bauer G. S. Epstein I. N. Gang	What are Migration Networks?	1	9/00
201	T. J. Dohmen G. A. Pfann	Worker Separations in a Nonstationary Corporate Environment	1	9/00
202	P. Francois J. C. van Ours	Gender Wage Differentials in a Competitive Labor Market: The Household Interaction Effect	5	9/00
203	J. M. Abowd F. Kramarz D. N. Margolis T. Philippon	The Tail of Two Countries: Minimum Wages and Employment in France and the United States	5	9/00
204	G. S. Epstein	Labor Market Interactions Between Legal and Illegal Immigrants	1	10/00
205	A. L. Booth M. Francesconi J. Frank	Temporary Jobs: Stepping Stones or Dead Ends?	1	10/00
206	C. M. Schmidt R. Baltussen R. Sauerborn	The Evaluation of Community-Based Interventions: Group-Randomization, Limits and Alternatives	6	10/00
207	C. M. Schmidt	Arbeitsmarktpolitische Maßnahmen und ihre Evaluierung: eine Bestandsaufnahme	6	10/00
208	J. Hartog R. Winkelmann	Dutch Migrants in New Zealand: Did they Fare Well?	1	10/00
209	M. Barbie M. Hagedorn A. Kaul	Dynamic Efficiency and Pareto Optimality in a Stochastic OLG Model with Production and Social Security	3	10/00
210	T. J. Dohmen	Housing, Mobility and Unemployment	1	11/00
211	A. van Soest M. Das X. Gong	A Structural Labour Supply Model with Nonparametric Preferences	5	11/00

An updated list of IZA Discussion Papers is available on the center's homepage *www.iza.org*.