FIEF WORKING PAPER SERIES

No. 147

IDENTIFYING INTERDEPENDENT BEHAVIOR IN AN EMPIRICAL MODEL OF LABOR SUPPLY

By

Thomas Aronsson*
Sören Blomquist**
and
Hans Sacklén***

ABSTRACT
In this paper we test a particular form of interdependent behavior, namely the hypothesis that individuals' choices of hours of work are influenced by the average hours of work in a social reference group. There are problems to empirically disentangle the effects of interdependent behavior and preference variation across groups. We show that panel data or data from several points in time are needed. In the empirical analysis we combine cross-section data from 1973, 1980 and 1990. Our results support the hypothesis of interdependent behavior. The implication is that conventional tax policy predictions, in which preference interdependencies are neglected, will tend to underestimate the effect of a tax reform on hours of work. Our point estimates suggest that conventional calculations would capture only about a third of the actual change in hours of work.

JEL-classification: C51, H31, J22
Keywords: labor supply, interdependent preferences, nonlinear taxes

October 28, 1998

---

* Department of Economics, Umeå University.
** Department of Economics, Uppsala University.
*** Trade Union Institute for Economic Research.

We are grateful for helpful comments from three anonymous referees.

1. INTRODUCTION

In economics we usually assume that individuals only interact via the market. Other social sciences study direct interactions, both bilateral and in groups. Blomquist (1993) shows that if there are direct interactions, and we do not account for this, then our predictions of the effect of changes in the economic environment can be seriously biased. In the present paper we test the hypothesis of interdependent behavior in an empirical model of labor supply, and we also try to quantify its importance.

The significance of taking account of interdependent behavior, if it exists, can be illustrated by a simple example. Suppose the true data generating process is given by $h_i = const + \alpha w_i(1-\tau) + \beta y_i + \eta\bar{h}$, where h_i, w_i and y_i are hours of work, the wage rate and nonlabor income for individual i, $\bar{h}$ the arithmetic mean of hours of work and τ a proportional tax rate. That is, individual i:s hours of work are influenced by the average hours of work in the economy. The assumed data generating process implies that the average hours of work in the economy is given by $\bar{h} = const + \alpha\bar{w}(1-\tau) + \beta\bar{y} + \eta\bar{h}$. The effect of a change in τ is given by $d\bar{h}/d\tau = -\alpha\bar{w}/(1-\eta)$. If we do not realize that behavior is interdependent, and falsely include $\eta\bar{h}$ in the constant, we would calculate the effect as $d\bar{h}/d\tau = -\alpha\bar{w}$; that is, the effect would be underestimated by a proportional factor $1/(1-\eta)$. If η is, say, 0.5 this implies that the effect of a tax change would be underestimated by 50%. Most recent studies of the effect of taxes on labor supply have estimated functional forms where it is possible that an effect like $\eta\bar{h}$ is "hidden" in the constant. If there is interdependent behavior, earlier predictions of the effect of tax reform on labor supply may have been biased.

Several recent studies have attempted to test the hypothesis of interdependent behavior. Andreoni and Scholz (1990) use American cross-section data to test whether interdependent preferences can explain patterns of charitable giving. There are quite a few studies in which Dutch data have been used. Woittiez (1990) estimates models of labor supply for both males and females. Alessie and Kapteyn (1991) estimate a version of the Almost Ideal Demand System using a two-year panel. Kapteyn et al. (1997) estimate the linear expenditure system for a cross-section of households. Woittiez and Kapteyn (1998) study female labor supply. All these studies claim that their results support the hypothesis of interdependent behavior.

Most earlier empirical studies of interdependent behavior have used cross-section data, and the interdependence hypothesis has usually been phrased in terms of individuals being influenced by the behavior within a set of social reference groups. As we argue in section 2, it might be difficult to properly identify interdependent behavior using such data. The reason being that it is difficult to separate preference variation across groups from preference interdependence if only cross-section data are available. In the context of labor supply, if we use the mean of hours of work in the reference group as an explanatory variable in the labor supply function, then there is a risk that we simply catch preference variation across groups. This general conclusion is also in line with the findings of Manski (1993). Manski actually distinguishes between three effects, which he denotes endogenous effects, correlated effects and exogenous effects. The first two effects coincide with what we call interdependent behavior and preference variation. The exogenous effects imply that the behavior of an individual also varies with the exogenous characteristics of the reference group to which the individual belongs. We shall argue below that panel data or repeated cross-section data allow us to identify both interdependent behavior (endogenous effects) and preference variation across groups (correlated effects), if we make the identifying assumption that the exogenous effects are absent.

In the present study we use a data source which contains data from 1973, 1980 and 1990. During this time period the tax schedule changed considerably. In our data we therefore obtain a variation in average hours of work not only between social reference groups, but also a variation over time for each reference group. We use the maximum likelihood method to estimate the parameters of the model. Except for the inclusion of interdependent behavior, the technique is similar to the one used in Blomquist and Hansson-Brusewitz (1990).

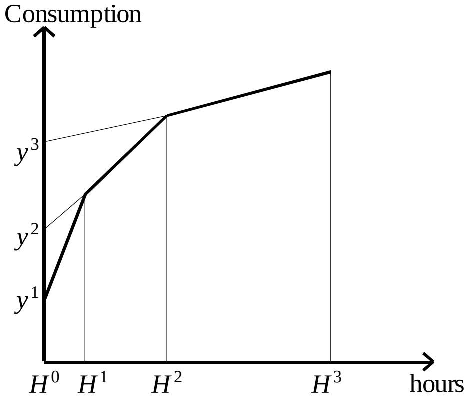
The rest of the study is organized as follows. In section 2 we present the model and the corresponding likelihood function. We also discuss what kind of data are required in order to test the hypothesis of interdependent behavior. In section 3 we perform a set of Monte Carlo simulations. Section 4 describes our data and the empirical results. Section 5 concludes.

2. A labor supply model with interdependent behavior

2.1 The model

The Swedish tax system is progressive, with the marginal tax increasing in a stepwise fashion as taxable income increases. Individuals' budget sets therefore are piecewise linear with a general shape as illustrated in figure 1. In this figure H^3 is the upper physical limit an individual can work and H^0 a lower limit.[1] y^1, y^2 and y^3 are the intercepts of the extended linear segments. In the labor supply literature these intercepts are often denoted "virtual incomes". The slopes of the linear segments can be interpreted as marginal wage rates. The kink points in the interior of the budget set are denoted by H^1 and H^2. Each segment is defined by the marginal wage rate w^s, virtual income y^s, lower starting point H^{s-1} and upper ending point H^s. As will become clear in section 2.3 below, this characterization of the budget set is quite useful when it comes to solving an individual's constrained optimization problem.

Figure 1. A piecewise linear budget constraint



Let us briefly describe what we will refer to as the standard labor supply model. The vast majority of recent empirical labor supply studies rely on the implicit assumption that individuals' preferences are not affected by other individuals' behavior. That is, changes in wages, income or taxes are believed to affect hours of work only via the individual's own budget constraint, wheras the map of indifference curves remains the same. Assuming that individual i maximizes a strictly quasi-concave utility function subject to a piecewise linear budget set, the desired hours of work will be a nonlinear function of the utility function parameters, the tax system and the individuals' gross wage rate and pretax nonlabor income. Suppressing the dependence on the utility parameters, we can summarize this as $h_i^* = f(W_i, Y_i; \tau)$, where h_i^* denotes desired hours, W_i is the

1 In reality, budget sets may consist of a large number of segments. For example, the 1981 Swedish tax and transfer system generated budget sets consisting of up to 27 segments.

gross wage rate, Y_i pretax nonlabor income and τ a vector of tax parameters. (Note that we use capital letters for pretax economic variables). Finally, taking measurement/optimization errors into account by adding a normally distributed random term ε_i (i.i.d. and $E(\varepsilon_i)$=0), observed hours of work, $\tilde{h}_i$, are assumed to be given by $\tilde{h}_i = h_i^* + \varepsilon_i$.[2]

Beginning in his doctoral thesis [Kapteyn (1977)], Kapteyn has in a series of articles studied interdependent preferences. He argues that individuals compare themselves to a social reference group. This idea can be formalized in several ways. In the present study we formulate a model where individuals' preferences are affected by the average hours of work in a social reference group, although the mode or the median might be equally resonable choices to represent the "social norm".[3] We also allow for the possibility that preferences may differ across groups. Retaining the hypothesis of utility maximization subject to a piecewise linear budget set, our modified version of the standard model then takes the form

$$h_i^{*j} = f(W_i^j, Y_i^j, \bar{h}^j; \tau, \mu^j), \qquad i=1,..., N^j \quad ; \quad j=1,...,K, \qquad (1)$$

where j indices social reference groups, μ^j is a group specific preference parameter and N^j the number of individuals in reference group j. $\bar{h}^j$ is the arithmetic mean of observed working hours in social reference group j, i.e. $\bar{h}^j = (1/N^j)\sum_{i=1}^{N^t}\tilde{h}_i^j$. Following the standard model by allowing for measurement/optimization errors, observed hours are assumed to be given by $\tilde{h}_i^j = h_i^{*j} + \varepsilon_i^j$, where $\varepsilon_i^j \sim$ N.I.D$(0,\sigma_\varepsilon^2)$. The econometric specification used for our empirical application is described in section 2.3 below.

2.2 Data requirements for identification

The fact that variations in $\bar{h}^j$ can arise because of preference variation across groups implies that cross-section data can not be used to identify interdependent behavior. To show this we use a simple linear model where we for simplicity abstract from taxes and exclude nonlabor income.[4] In the following we will also assume that we know to which group each individual belongs. As Manski (1993) points out, this is a crucial identifying assumption. Let data be generated by

---

2 If the sample contains corner observations (at zero hours or at he upper physical limit), then the assumed data generating process is usually stated in terms of a simple Tobit structure.

3 For simplicity we neglect the possibility that the variance of the hours distribution might be of importance.

4 In the empirical application below, we shall, of course, incorporate nonlabor incomes as well as taxation into the analysis.

$$\tilde{h}_i^j = \mu^j + \alpha W_i^j + \eta \bar{h}^j + \varepsilon_i^j, \qquad i=1,..., N^j \quad ; \quad j=1,...,K, \qquad (2)$$

where $\tilde{h}_i^j$ and W_i^j are, respectively, the hours of work and the wage rate of individual *i* in group *j*. In this model there is both interdependent behavior, captured by $\eta \bar{h}^j$, and a variation in preferences across groups. The latter is represented by the group specific terms μ^j in eq. (2). For each group we can derive the reduced form:

$$\tilde{h}_i^j = \frac{\mu^j}{1-\eta} + \frac{\eta\alpha}{1-\eta}\bar{W}^j + \alpha W_i^j + \varepsilon_i^j \qquad (3)$$

Within each group there is no variation in $(\frac{\mu^j}{1-\eta} + \frac{\eta\alpha}{1-\eta}\bar{W}^j)$. We denote the expression in parentheses by c^j and rewrite (3) as

$$\tilde{h}_i^j = c^j + \alpha W_i^j + \varepsilon_i^j \qquad (4)$$

Consistent estimation of c^j and α is straightforward. However, we cannot from information on c^j and α deduce the coefficients η and μ^j. Thus, η and μ^j are not identifiable. This implies that if data are generated according to (2), then cross-section data can not be used to test the hypothesis of interdependent behavior.[5]

Suppose instead that we have panel data, or several cross-sections, and that μ^j does not vary over time. We assume that, for some reason, $\bar{W}^j$ varies over time.[6] This implies that the reduced form in (3) becomes

$$\tilde{h}_i^{jt} = \frac{\mu^j}{1-\eta} + \frac{\eta\alpha}{1-\eta}\bar{W}^{jt} + \alpha W_i^{jt} + \varepsilon_i^{jt}, \qquad t=1,...,T \qquad (5)$$

where *t* indices time. It is straightforward to consistently estimate the parameters of the function

$$\tilde{h}_i^{jt} = c^j + b\bar{W}^{jt} + \alpha W_i^{jt} + \varepsilon_i^{jt} \qquad (6)$$

Equations (5) and (6) imply the relations

5 The model that we estimate in this paper is nonlinear in both parameters and variables. (This is because the budget constraint is nonlinear.) We know that nonlinearities in general help to identify parameters. However, nonlinearities caused by nonlinear budget constraints do not help to identify μ^j and η if only cross-section data are used.

6 Applications which study the effect of taxes on labor supply typically have net rather than gross wage rates entering the supply function. Variations in the mean net wage rate over time can then, for example, occur due to changes in the tax system.

$$\frac{\mu^j}{1-\eta} = c^j \qquad (7a) \qquad \text{and} \qquad \frac{\eta}{1-\eta} = \frac{b}{\alpha} \qquad (7b)$$

From (7b) we can solve for η and obtain $\eta = b/(\alpha + b)$. We can then solve (7a) and obtain $\mu^j = \alpha c^j/(\alpha + b)$. Thus, if we have panel data or several cross-sections such that within each group $\overline{W}^j$ varies over time, then η and μ^j are identifiable.

We have just shown that μ^j and η are identifiable if panel data or several cross-sections are used. However, in our actual estimation two further problems arise. Firstly, we have cross-section data from only three points in time. Secondly, given the data available we must use a rather crude classification of individuals into groups. In section 3 we perform a set of Monte Carlo simulations to study if one, given these limitations, can identify interdependent behavior.

2.3 The econometric specification

As noted in section 2.1, we assume preferences are strictly convex and that desired hours of work for each individual are generated by utility maximization subject to a piecewise linear budget set. To describe the preferences we use the basic supply function as defined in Blomquist & Hansson-Brusewitz (1990). A basic supply function $b(w,y)$ shows desired hours of work generated by a linear budget constraint with slope w and intercept y. The budget sets we use in the empirical analysis and in the Monte Carlo simulations below are convex. Then, given that preferences are strictly convex, there is a unique solution to the individual's optimization problem. The basic supply function can be used in a simple search algorithm to find the desired hours of work (h^*) implied by the utility optimum.[7] Using the notation introduced in connection with figure 1 in section 2.1, if we calculate $b(w^1,y^1)$ and find that $H^0<b(w^1,y^1)\leq H^1$, then $b(w^1,y^1)$ is the unique global optimum. If $b(w^1,y^1)\leq H^0$ we have a corner solution at H^0. If $b(w^1,y^1)>H^1$ we move on to the second segment and calculate $b(w^2,y^2)$. If $b(w^2,y^2)<H^1$ we conclude that the unique optimum is at the first kink point, H^1. If instead $H^1<b(w^2,y^2)\leq H^2$, then $b(w^2,y^2)$ is the unique optimum. If $b(w^2,y^2)>H^2$ we continue to segment three e.t.c.

We assume preferences are such that, for individual i in group j at time t, the basic supply function is linear in the slope and intercept of the budget constraint, or,

$$b(w_i^{jt}, y_i^{jt}) = \psi^{jt} + \alpha w_i^{jt} + \beta y_i^{jt} + \delta X_i^{jt} \qquad (8)$$

7 This algorithm was first described in Hausman (1979).

where δ is a parameter vector and X_i^{jt} a vector of socio-economic variables representing observed preference heterogeneity within groups.[8] We will estimate a sequence of models that are special cases of (8):

i. In the first model, which is the most general, we allow the intercept of the estimated function, ψ^{jt}, to vary freely both across groups and over time.

ii. We impose the restriction $\psi^{jt} = \mu^j + \eta \bar{h}^{jt}$. That is, we have a specification with interdependent behavior as represented by $\eta \bar{h}^{jt}$ and group specific preferences as represented by the terms μ^j. This is the type of model discussed in the preceding subsection, where we concluded that preference variation across groups can be distinguished from interdependent behavior if the group effects are time independent.

iii. We impose the additional restriction $\psi^{jt} = \mu + \eta \bar{h}^{jt}$. That is, there is interdependent behavior but no group specific preferences.

iv. We impose the restriction $\psi^{jt} = \mu$. Hence, the fourth specification is a standard labor supply model which neither contains group specific effects nor interdependent behavior.

These hypotheses are nested, which implies that we may carry out a sequence of likelihood ratio misspecification tests in order to check the validity of the various models.

Our assumptions are such that the log-likelihood contribution from each group and time enters the likelihood function additively. It is therefore sufficient to present the likelihood contribution for one group at time *t*. The interdependence across observations implied by *ii.* and *iii.* will, in principle, make the likelihood function more complicated than the usual "Hausman" type likelihood function.[9] The likelihood function is derived in some detail in Appendix A. As explained in the Appendix, we can in practice drop the portion of the likelihood function that emanates from preference interdependence. In this section we therefore describe a likelihood function where the interdependence part is excluded.

8 The corresponding direct utility function was derived in Hausman (1980). We have chosen not to include unobserved preference variation within groups in our model. The reason for this is that the likelihood function we present here is a simplified version of a more complicated model described in Appendix A. To include random preferences in this underlying model would be very difficult. Hence, if data are generated with random preferences we can regard the likelihood function as a pseudo likelihood function. The fact that we do not allow for random preferences within groups might lead to a bias in our estimated parameters. To see if this is the case the Monte Carlo simulations in section 3 are performed for a data generating process where random preferences are an integral part. We find that this pseudo likelihood function does not lead to biased estimates.

9 Originally developed by Burtless and Hausman (1978), the so called Hausman method takes account of the complete form of individuals' budget constraints and uses maximum likelihood techniques to estimate the parameters of the labor supply function.

Observed hours of work are assumed to be given by $\tilde{h}_i^{jt} = h_i^{*jt} + \varepsilon_i^{jt}$, where $\varepsilon_i^{jt} \sim$ N.I.D$(0, \sigma_\varepsilon^2)$.[10] The likelihood contribution from a typical individual is then given by $(1/\sigma_\varepsilon)\phi\,[(\tilde{h}_i^{jt} - h_i^{*jt})/\sigma_\varepsilon]$, with ϕ being the standard normal density function. The log-likelihood function will therefore have the form:

$$\ln L = \sum_{j=1}^{K} \sum_{t=1}^{T} L^{jt} \tag{9}$$

where K is the number of social reference groups, T the number of cross-sections and

$$L^{jt} = \sum_{i=1}^{N^{jt}} \ln \frac{1}{\sigma_\varepsilon} \phi \left[\frac{\tilde{h}_i^{jt} - h_i^{*jt}}{\sigma_\varepsilon} \right] \tag{10}$$

As our model with interdependent behavior is formulated, it is the population means $\bar{h}^{jt}$ that should enter the likelihood function. However, due to the fact that we have not been able to obtain reliable information about the population means, we will instead turn to our data and calculate the sample means. Since the sample mean is a consistent estimator of the population mean, the maximum likelihood estimator is still consistent. There is however a possibility that the estimated asymptotic variances are biased. We investigate this in the Monte Carlo simulations.

3. MONTE CARLO SIMULATIONS

Suppose the true data generating process contains preference variation across social reference groups and interdependent behavior. What will be the result if we incorrectly estimate a function that fails to control for interdependent behavior and/or preference variation? What happens if we are unable to identify the social reference groups properly? The model and likelihood function developed in the previous section do not allow for unobserved heterogeneity within groups. What are the properties of the simplified estimation method if data are generated with random preferences? The estimation method described in the previous section is based on the assumption that population mean values of the hours of work in each reference group can be measured without error by using the sample mean values. Is this a harmless assumption or should we expect

10 This convenient form for the assumed data generating process follows from the fact that our sample does not include any corner observations at zero hours or at the upper physical limit (see section 4.1 below).

the estimated asymptotic variances to be seriously biased? In this section we study these issues with the aid of Monte Carlo simulations.

3.1 Design of experiments

Using the notation of section 2, the data generating process (DGP) is set up as follows: The basic labor supply function is of the form

$$b(w_i^{jt}, y_i^{jt}) = \psi_i^{jt} + \alpha w_i^{jt} + \beta y_i^{jt}, \quad i=1,\ldots,N^{jt};\ j=1,\ldots,K;\ t=1,\ldots,T \quad (11)$$

where $\psi_i^{jt} = \mu^j + \eta \bar{h}^{jt} + v_i^{jt}$. The DGP thus contains preference variation across social reference groups, interdependent behavior and unobserved preference variation within groups (the random term v_i^{jt} allows the intercept to vary across individuals). In order to mimic our empirical application, where we use data from three points in time and classify observations into eight different groups, we set K=8 and T=3.

The simulations employ a multivariate distribution of gross wage rates and nonlabor income from actual data. We use a sample of 500 married males with a positive labor supply from the 1981 Level of Living survey.[11] Coupled with a slightly stylized version of the Swedish 1980 income tax system, the wage and nonlabor income define a piecewise linear budget constraint for each individual with at most 19 linear segments. We then use multiples of the sample in order to generate a population of 80,000 individuals. The population is then split randomly into eight different social reference groups. We let the groups be of equal size, i.e. N^j= 10,000 for j=1,…,8.

To obtain variation in $\bar{h}^j$ not only between social reference groups, but also a variation over time for each reference group, we impose certain changes in the original tax schedule. Then, when combining the adjusted tax schedules with the wage rate and nonlabor income data, we generate data for two additional time periods. Data for t=2 are generated by imposing a 50% reduction of the marginal tax rates in the original tax schedule. Similarly, we obtain data for t=3 by setting all marginal tax rates to zero. The number of individuals in the reference groups are held constant across time. That is, N^{jt}=10,000, j=1,…,8 and t=1,2,3.

11 The data source is described more closely in section 4.1 below.

Individuals belonging to the first group (at any point in time) are assigned a constant term μ^1, those belonging to the second group μ^2, and so on. The random preference components v_i^{jt} are drawn from a normal distribution with mean zero, variance σ_v^2 and E($v_i^{jt} v_s^{k\tau}$)≠0 if and only if $i=s$, $t=\tau$ and $j=k$. Using the basic supply function we can then, for given values of the preference parameters, locate each individual's desired hours of work, h_i^{*jt}. The parameter values are given in table 1. Evaluated at mean values for desired hours, net wage rates and virtual income, these parameter values imply a mean wage rate elasticity of about 0.1 and an income elasticity of about -0.1. Finally, by adding a normally distributed measurement error (i.i.d) to the desired hours of work, we arrive at the number of hours that are actually observed, $\tilde{h}_i^{jt} = h_i^{*jt} + \varepsilon_i^{jt}$.

Table 1. Characteristics of the DGP

N^t = 80,000, t =1,…,3 ; N^{jt} = 10,000, j = 1,…,8 and t=1,…,3.
α = 0.01, β = -0.005, η = 0.5,
μ^1= 1.7, μ^2 = 1.5, μ^3 = 1.3,…,μ^8 = 0.3,
σ_ε = 0.1, σ_v = 0 in simulation 1 and σ_v = 0.2 in simulations 2, 3, 4 and 5.

Note: The gross wage rate is measured in SEK, nonlabor income in thousands of SEK and hours of work in thousands of hours.

To estimate the model we generate three cross-sections by random sampling of n^t observations from the full population N^t. The generated samples have size n^t=800 for t=1,2,3, which is the approximate size of the samples we will use in our empirical application. The experiments are listed in table 2, and will be described in greater detail below.

3.2 Simulation results[12]

Table 3 displays the results for the interdependence coefficient (η), the wage coefficient (α), the coefficient for nonlabor income (β) and the intercept of the first reference group (μ^1). The results for each experiment are based on 500 replications, with new samples generated for each replication. The percentage bias is calculated as $100 \cdot (\bar{\hat{\theta}} - \theta)/|\theta|$, where θ is the true parameter value and $\bar{\hat{\theta}}$ the mean estimate generated by 500 replications. The standard deviation (STD), which is a measure of the sampling variability, is measured by the square root of $(1/500)\sum_{k=1}^{500}(\hat{\theta}_k - \bar{\hat{\theta}})^2$. MESE denotes the mean of 500 estimated asymptotic standard errors. We

12 All simulations are performed with programmes written in the Fortran programming language. The GQOPT package of numerical optimization algorithms (version 6.12) was employed to find the optima for the likelihood functions. The variance-covariance matrix for the estimated parameter vector is calculated as the inverse of the Hessian of the log-likelihood function evaluated at the estimated parameter vector.

also report the ratio MESE/STD, which will be used as a measure of the bias in the variance-covariance matrix. For example, a ratio below unity suggests that the estimated asymptotic standard error underestimates, on average, the underlying standard deviation for the parameter. Finally, we should emphasize that conclusions and remarks made throughout this section are valid for (at least) the particular samples sizes used in the simulations. However, other sample sizes could possibly have generated different results.

For *simulation 1* data are generated without random preferences. At the estimation stage we correctly recognize the presence of both interdependent behavior and preference variation across groups. In addition, we are able to sort observations to the correct reference groups. The population moments $\bar{h}^{jt}$ are, however, unobserved in the estimation process. Instead, we let the sample counterparts $\hat{\bar{h}}^{jt} = (1/n^{jt})\sum_{i=1}^{n^{jt}} \tilde{h}_i^{jt}$ enter the likelihood function, where n^{jt} is the number of observations sampled from group *j* at time *t*. The results in table 3 show that we obtain very accurate estimates of the parameters in the model. All coefficients are virtually unbiased, and for η, α and μ^1 the asymptotic standard errors are, on average, close approximations of the actual standard deviations for the parameters.[13] The latter suggests that sample mean values of hours worked ($\hat{\bar{h}}^{jt}$) can be substituted for population values ($\bar{h}^{jt}$) without causing any significant bias in the estimated variance-covariance matrix.[14]

Recall that our likelihood function does not account for random preferences. *Simulation 2* is the same as simulation 1 except that data are now generated with random preferences. Table 3 shows that the coefficients are still estimated with a negligible bias. As expected, the sampling variability increases for all coefficients, but so do the estimated standard errors. However, the simplified estimation method tends to generate a moderately biased variance-covariance matrix. For the interdependence coefficient (η) the asymptotic standard error underestimates the underlying standard deviation by, on average, 10 percent. The nonlabor income coefficient (β) is

---

13 Simulations where we used larger samples (not shown in table 3) suggest that the bias in the variance-covariance matrix for the nonlabor income coefficient (β) is small sample bias that disappears in larger samples.

14 At this point one might ask whether these results carry over to a situation where some reference groups in the population are quite small in size. The problem is then that relatively few observations would be sampled from these particular groups. We have performed simulations to study this issue. The basic setup was the same as described in section 3.1, but for two groups we set the size to 1/50 of the population (rather than 1/8, which was the case for simulation 1). Consequently, a random sample of 800 observations contained, on average, 16 observations from each of these two groups. The results were quite similar to those reported for simulation 1. The main effect was a slight increase in the standard deviation for the constant terms corresponding to the two groups.

Table 2. Simulations

	DGP	Estimation	
Sim.	σ_v	*Estimated function*	*Misclassification*
1	0.0	$b(w_i^{jt}, y_i^{jt}) = \mu^j + \eta \hat{\hat{h}}^{jt} + \alpha w_i^{jt} + \beta y_i^{jt}$	No
2	0.2	"	"
3	"	"	50%
4	"	$b(w_i^{jt}, y_i^{jt}) = \mu + \eta \hat{\hat{h}}^{jt} + \alpha w_i^{jt} + \beta y_i^{jt}$	No
5	"	$b(w_i^{jt}, y_i^{jt}) = \mu^j + \alpha w_i^{jt} + \beta y_i^{jt}$	"

Table 3. Simulation results

Parameter	*Sim. 1*	*Sim. 2*	*Sim. 3*	*Sim. 4*	*Sim. 5*
η					
Bias %	0.0	0.5	10.6	102.6	
STD*	$1.07 \cdot 10^{-2}$	$2.20 \cdot 10^{-2}$	$8.04 \cdot 10^{-2}$	$2.98 \cdot 10^{-3}$	
MESE**	$1.08 \cdot 10^{-2}$	$1.98 \cdot 10^{-2}$	$4.43 \cdot 10^{-2}$	$4.87 \cdot 10^{-3}$	
MESE/STD	1.00	0.90	0.55	1.63	
α					
Bias %	0.0	-0.4	-5.2	-68.6	64.4
STD*	$2.08 \cdot 10^{-4}$	$4.64 \cdot 10^{-4}$	$1.92 \cdot 10^{-3}$	$2.85 \cdot 10^{-4}$	$3.93 \cdot 10^{-4}$
MESE**	$2.07 \cdot 10^{-4}$	$3.80 \cdot 10^{-4}$	$8.58 \cdot 10^{-4}$	$2.93 \cdot 10^{-4}$	$3.09 \cdot 10^{-4}$
MESE/STD	1.00	0.82	0.45	1.03	0.79
β					
Bias %	0.8	1.0	4.8	21.2	-18.7
STD*	$8.51 \cdot 10^{-5}$	$1.87 \cdot 10^{-4}$	$6.79 \cdot 10^{-4}$	$1.85 \cdot 10^{-4}$	$2.05 \cdot 10^{-4}$
MESE**	$7.33 \cdot 10^{-5}$	$1.23 \cdot 10^{-4}$	$2.18 \cdot 10^{-4}$	$1.45 \cdot 10^{-4}$	$1.40 \cdot 10^{-4}$
MESE/STD	0.86	0.66	0.32	0.78	0.68
μ^1					
Bias %	-0.1	-0.6	-31.4	-92.3	91.6
STD*	$3.41 \cdot 10^{-2}$	$6.85 \cdot 10^{-2}$	0.179	$1.43 \cdot 10^{-2}$	$2.20 \cdot 10^{-2}$
MESE**	$3.37 \cdot 10^{-2}$	$6.15 \cdot 10^{-2}$	0.101	$1.58 \cdot 10^{-2}$	$1.68 \cdot 10^{-2}$
MESE/STD	0.99	0.90	0.56	1.10	0.76

* Standard deviation of 500 estimates.
** Mean of 500 estimated standard errors.

the most troublesome coefficient; the estimated standard errors underestimate the actual standard deviation by almost 35 percent. We thus conclude that our simplified estimation method performs well in the sense that parameter estimates are virtually unbiased. t-values based on the estimated variance-covariance matrix will however be slightly overstated.[15]

Simulation 3 is the same as simulation 2 except that we are now unable to identify the social reference groups properly. We know (for some reason) that there are 8 different groups in the population, but we have imperfect information on the composition of groups. The mis-classification of observations is set up as follows: There is a fifty-fifty chance that an individual who truly belongs to group *j* is correctly classified as a group *j* observation. Consequently, 50 percent of the the sampled observations are assigned to an incorrect reference group. The destination group for a misclassified observation is selected purely at random. Table 3 suggests that this type of misclassification error might be a potential problem for empirical applications. The good news is that the coefficients for the net wage rate, nonlabor income and average hours of work are only moderately biased.[16] The coefficient for average hours of work (η) has a positive bias of approximately 10 percent. However, one might of course suspect this finding to be heavily influenced by the setup; alternative misclassification schemes could possibly cause more damage to the results. Simulation 3 also reveals that misclassification errors induce significantly larger standard deviations for all coefficients. The bad news is that the asymptotic variance-covariance matrix fails to capture this accurately. For the interdependence coefficient (η), the asymptotic standard error underestimates the underlying standard deviation by, on average, 45 percent.

Returning to the case of perfect information on the composition of groups, *simulation 4* investigates the consequences of estimating a function that fails to control for preference variation across groups. The coefficients for the net wage rate and nonlabor income are both heavily biased towards zero. The coefficient for average hours of work is picking up the effects of variations in

15 We have also considered two alternative ways of including random preferences in the DGP: 1) the wage rate coefficient (α) varies, and 2) the nonlabor income coefficient (β) varies across individuals. The results were similar to those reported in table 3.

16 The large negative bias for μ^1 is due to the inflow of misclassified observations for which hours of work are, on average, considerably fewer than for the true group 1 members (recall that group 1 is the group with the largest value for the constant term). Similarly, μ^8 has a large positive bias (not shown), since true members of group 8 are mixed with observations for whom data are generated with larger constant terms.

preferences across groups. Consequently, if we fail to control for preference variation, then we run the risk of exaggerating the importance of preference interdependence.[17]

What happens if we estimate a function that incorrectly neglects interdependent behavior? In *simulation 5* we recognize that preferences vary across groups, but we do not allow for interdependent preferences. Hence we may think of it as a conventional labor supply specification. Omitting the average hours of work variable, the coefficients for the net wage rate and nonlabor income are both heavily biased away from zero and the constant is picking up some of the effects of interdependent behavior. Thus, if the DGP is characterized by interdependent behavior and this is not accounted for in estimation, we encounter two major problems. First, we ignore the special dynamics that are associated with preference interdependencies. Second, we would expect the estimated parameters to be seriously biased.

4. EMPIRICAL ANALYSIS

4.1 Data

We use the Swedish "Level of Living Surveys" from 1974, 1981 and 1991. The surveys describe the respondents economic conditions in 1973, 1980 and 1990. The 1974 sample consists of 6593 randomly chosen individuals aged 15-75, the 1981 sample of 6813 individuals and the 1991 sample of 6710. The response rate was 85.2% in 1974, 82.4% in 1981 and 79.1% in 1991.

In our sample we only include married men in ages 20-60. Farmers, students, those with more than four weeks of sick leave, and those who were self-employed are excluded. The remaining sample consists of 777 observations for 1974, 863 for 1981 and 678 for 1991. For the three cross-sections combined there are 2318 observations. The data contains detailed information about taxable income, tax deductions etc. Thus, we are able to model each individual's nonlinear budget constraint in a detailed and precise way. We believe this kind of data is unique to the Scandinavian countries. We also take account of housing allowances, which is a quite important transfer system. The tax and transfer systems create a few minor nonconvexities in individuals'

17 Suppose the true data generating process contains preference variation across groups but no interdependent behavior, and we estimate a function that allows for interdependent behavior. Is there a risk that we falsely conclude that there is interdependent behavior? We have performed simulations to study this issue. The main results are the following: Controlling for both variations in preferences and, falsely, for interdependent behavior, we obtain small and insignificant estimates of η. On the other hand, if one does not control for preference variation, then the results resemble those reported for simulation 4.

budget sets. To simplify the computations we have convexified the budget sets by taking the convex hull of a given budget set. Results in Blomquist and Hansson-Brusewitz (1990) and MaCurdy et al. (1990) suggest that this is a harmless simplification.[18]

We use annual hours of work in 1973, 1980 and 1990 respectively as the dependent variable in the labor supply function. This variable is computed from the answers to a series of questions on average hours/week, and the number of weeks the person worked at full-time and part-time jobs and moonlighting. Total hours of work are obtained by multiplying the number of working weeks by hours worked per week for each type of job and then summing over jobs. All individuals included in our samples had at least some hours of paid employment.[19]

The gross wage rate is measured by a direct question about the person's wage rate at the time of the interview. We define nonlabor income on segment 1 of the budget constraint as $Y_1=MDI+CAP+BBF$, where *MDI* is the sum of the spouse's total income after tax including child allowances. We assume *MDI* affects the hours of work in the same way as exogenous nonlabor income, i.e. we assume the hours of work of the wife are given and independent of the hours of work of the man. *CAP* measures capital income net of the capital income tax, as it would have been at zero hours of work. It includes imputed income of owner occupied homes. *BBF* denotes housing allowances as they would have been at zero hours of work. All economic variables are adjusted to the 1980 price level. Detailed sample statistics are given in Appendix B.

We formulate the hypothesis of interdependent behavior in such a way that each individual belongs to a social reference group. The individual is influenced by the average hours of work in the social reference group to which he belongs. If such groups exist, they would be very difficult to establish correctly. The proper average would probably be a weighted average. However, given the data available we must use a rather crude classification of individuals into groups. Previous studies have usually classified individuals on the basis of observable sociodemographic

---

18 See Blomquist (1983) and Blomquist and Hansson-Brusewitz (1990) for details about the tax and transfer systems in 1973 and 1980, respectively. Table B3 in Appendix B shows the Swedish 1990 national income tax schedule.

19 For the time period in question the male unemployment rate in Sweden was very low. In the 1973 data set there was only one man that was unemployed the whole year. For 1990 the corresponding figure was two persons. In the 1980 sample none was unemployed the whole year. Those persons who were unemployed the whole year were excluded from the sample. Since only three persons are excluded we do not think this affects our results. Of course, some individuals were unemployed part of the year. We have not attempted to account for this, but regard these deviations from desired hours as optimization errors taken care of by the measurement/optimization error. Very few of the males in the sample had unemployment spells and for those that were unemployed the spells were usually quite short.

characteristics. We shall use a classification influenced by the one used in Woittiez (1990) and Woittiez and Kapteyn (1998).[20] We classify people into eight groups according to educational level, age and the presence of children under 18 years of age in the household. Low education is defined as education below secondary school. The classification is displayed in table 4.

Table 4. The social reference groups

Group	*Education*	*Age*	*Children*
1	Low	≤40	Yes
2	High	≤40	Yes
3	Low	≤40	No
4	High	≤40	No
5	Low	>40	Yes
6	High	>40	Yes
7	Low	>40	No
8	High	>40	No

In the period 1973 to 1990 the tax schedule changed considerably. Thus, if hours of work are sensitive to variations in economic variables, the change in the tax schedule would induce a variation in $\bar{h}^j$ over time. In our data we therefore obtain a variation in $\bar{h}^j$ not only between social reference groups, but also a variation over time for each reference group. As discussed in section 2.2 above, this should help to identify interdependent behavior. Table B2 in Appendix B provides some summary statistics for the social reference groups.

4.2 Estimation Results

Let us briefly recapitulate the model we estimate. Using the notation from section 2, observed hours of work are assumed to be generated as $\tilde{h}_i^{jt} = h_i^{*jt} + \varepsilon_i^{jt}$, where i=1,…, N^{jt}; j=1,…,8 ; t=1,…,3. ε_i^{jt} is a normally distributed error with mean zero, variance σ_ε^2 and E($\varepsilon_i^{jt}\varepsilon_s^{k\tau}$)≠0 if and only if $i=s$, $t=\tau$ and $j=k$. Desired hours of work, h_i^{*jt}, are assumed to be generated by utility maximization subject to the individual's nonlinear budget constraint. The preferences are defined by the basic supply function, $b(w_i^{jt}, y_i^{jt})$. We will estimate the following sequence of specifications:

$$\text{Model 1: } b(w_i^{jt}, y_i^{jt}) = \psi^{jt} \qquad + \alpha w_i^{jt} + \beta y_i^{jt} + \delta X_i^{jt} \qquad (12a)$$

20 Woittiez (1990) and Woittiez and Kapteyn (1998) classify observations according to education and age. Since Woittiez and Kapteyn have direct survey information on individuals' social environment, they can do a more sophisticated grouping. They estimate a latent variables model relating the direct information to the "true" but unobserved reference groups.

$$\text{Model 2: } b(w_i^{jt}, y_i^{jt}) = \mu^j + \eta \bar{h}^{jt} + \alpha w_i^{jt} + \beta y_i^{jt} + \delta X_i^{jt} \qquad (12b)$$
$$\text{Model 3: } b(w_i^{jt}, y_i^{jt}) = \mu + \eta \bar{h}^{jt} + \alpha w_i^{jt} + \beta y_i^{jt} + \delta X_i^{jt} \qquad (12c)$$
$$\text{Model 4: } b(w_i^{jt}, y_i^{jt}) = \mu + \alpha w_i^{jt} + \beta y_i^{jt} + \delta X_i^{jt} \qquad (12d)$$

X is a vector of observable personal attributes with corresponding parameters δ. We include in X the physical age of a person, given by the variable *AGE*, and the number of children under 18 that live at home, *NC*. The population moment $\bar{h}^{jt}$ is throughout this section estimated by its sample counterpart. The likelihood function used in the analysis is given by eqs. (9) and (10). We show the estimated functions in table 5.

Model 1 is the most general as we allow the intercept of the estimated function to vary freely across both groups and years. In model 2 we impose the restriction that ψ^{jt} takes the form $\mu^j + \eta \bar{h}^{jt}$. That is, we restrict the model in such a way that it allows for group specific preferences and interdependent behavior. We perform a likelihood ratio test of the restrictions imposed by model 2. Our null hypothesis is that model 2 (the restricted model) is correct. The alternative is given by model 1 (the unrestricted model). Let L^R and L^U denote the value of the likelihood for the restricted and unrestricted models. As test statistic we use -2(ln L^R-ln L^U), which has an asymptotic χ^2 distribution with degrees of freedom given by the number of linear restrictions imposed by the null hypothesis. The test statistic is in the present case 7.8. The null hypothesis imposes 24 linear restrictions and implies a critical value for the Chi-square test statisitic of 36.4 at the 5 % significance level. Hence, the null hypothesis is not rejected.

In model 3 we impose the restriction that there are no group specific preferences ($\mu^j = \mu$). Performing a test where we regard model 3 as the null (the restricted model) and model 2 as the alternative (the unrestricted model), we obtain a test statistic of 2.0. Since the critical value is 15.5, we cannot reject the null hypothesis. Finally, in model 4, we impose the additional restriction that $\eta = 0$, i.e. that there is no interdependent behavior. Model 4 is a traditional type of labor supply function, which neither incorporates interdependent behavior nor allows for variation in preferences across social reference groups. Performing a test where model 4 is the null and model 3 the alternative, we obtain a Chi-square test statisitic of 11.8. There is only one restriction imposed, so the critical value is 3.84. Hence, the null hypothesis is rejected and model 3 is our ”preferred” model.

Turning to the parameter estimates for model 3, it should first be noted that the linear basic supply function requires satisfaction of the Slutsky inequality α-$\beta h\geq 0$ in order for preferences to be convex. The results in table 5 suggest that the Slutsky condition does not represent a binding constraint for any observation. Hence, data could have been generated by utility maximization with convex preferences. The coefficient that captures interdependent behavior, η, is significant and large in the sense that its economic implications are important. We discuss the importance of this further below. The wage rate and nonlabor income coefficients are significant, and the point estimates are well in accordance with estimates presented in earlier Swedish studies (see e.g. Blomquist & Hansson-Brusewitz (1990) and Sacklén (1996)). The impact of socioeconomic characteristics is insignificant; the coefficients for age and children are small and estimated with poor precision.

Within the context of our model of interdependent behavior, wage and income elasticities can be defined in at least two different ways. We define the "partial" wage and income elasticities in terms of the reaction to a change in the own wage rate or nonlabor income, given that other individuals' behavior are unchanged. This is, of course, the way wage and income elasticities are usually defined. To evaluate the elasticities we merge our three cross-sections and use the mean of the variables at observed hours of work. Thus, the partial elasticities are formally calculated as $\alpha\overline{w}/\overline{h}$ and $\beta\overline{y}/\overline{h}$, where $\overline{w}$ is the mean net wage rate, $\overline{y}$ the mean virtual income and $\overline{h}$ the mean hours of work. For model 3 the implied partial wage and income elasticities are 0.061 and -0.017, respectively. The income elasticity is about the same as obtained in earlier studies for Sweden using the standard labor supply model, wheras the wage elasticity is slightly lower.

The partial elasticities might not be very useful if preferences are interdependent. In particular, the partial elasticities seem irrelevant for the evaluation of tax reform, since changes in the tax schedule affect almost all individuals' budget constraints and behavior. Our second type of wage and income elasticity therefore takes preference interdependencies into account. In table 5 these elasticities are referred to as the "total" wage and income elasticity, respectively. For model 3 we define the total elasticities in terms of the basic labor supply function (12c). Dropping the group and time superscripts, and averaging, we obtain $\overline{h} = \mu + \eta\overline{h} + \alpha\overline{w} + \beta\overline{y} + \delta\overline{X}$. The total wage elasticity is then calculated as

$$\frac{d\bar{h}}{d\bar{w}}\frac{\bar{w}}{\bar{h}} = \frac{a\bar{w}}{(1-\eta)\bar{h}} \tag{13}$$

and similarly for the income elasticity. Evaluating the elasticities at mean sample values for the variables, the total wage and income elasticities are 0.187 and -0.052, respectively. This implies that standard tax policy predictions, based on what we refer to as partial elasticities, will tend to seriously underestimate the effect of a tax reform on hours of work. Comparing the results for model 3 and model 4 (the standard model), and taking the estimate of η and the implied total elasticities at face value, we conclude that the standard analysis would capture only about a third of the actual change in hours of work.

Table 5. Estimation results

Parameter	*Model 1*	*Model 2*	*Model 3*	*Model 4*
$\bar{h}^{jt}$ (η)	-	0.67062 (3.32)	0.67391 (3.37)	-
Net wage rate (α)	0.00927 (6.47)	0.00807 (5.75)	0.00781 (5.59)	0.00889 (6.21)
Virtual income (β)	-0.00045 (-2.29)	-0.00058 (-2.29)	-0.00071 (-2.78)	-0.00077 (-2.93)
Constant	1.895 (52.0)	0.52613 (1.25)	0.57526 (1.38)	1.983 (54.21)
NC	0.00475 (0.85)	0.00509 (0.90)	0.00275 (0.49)	0.00373 (0.66)
AGE	0.00226 (3.70)	0.00171 (2.77)	0.00046 (0.73)	0.00053 (0.86)
σ_ε	0.271 (68.1)	0.271 (68.1)	0.272 (68.1)	0.272 (68.1)
lnL	-262.6	-266.5	-267.5	-273.4
LR-statistic	-	7.8	2.0	11.8
Partial wage elasticity	0.072	0.063	0.061	0.069
Total wage elasticity	-	0.192	0.187	-
Partial income elasticity	-0.011	-0.014	-0.017	-0.019
Total income elasticity	-	-0.043	-0.052	-

Note: Asymptotic t-values are given in parentheses. Hours of work are measured in thousands of hours. The marginal wage rate is measured in SEK and the virtual income in thousands of SEK, both at the 1980 price level. Detailed sample statistics are given in Appendix B. The various elasticities are defined and explained in the text. To conserve on space we do not report the full set of estimated constant terms. For model 1 we show the constant for group 1 in 1980, and for model 2 (in which the constant varies across groups, but not over time) the constant for group 1. The full set of estimates is available on request.

How reliable are the empirical results presented in this section? One reassuring result from the Monte Carlo experiments in section 3 concerns the identification issue. If data are generated with interdependent behavior, and this is correctly specified in the estimated function, then our simulations suggest that the labor supply parameters, including the coefficient for $\bar{h}^{jt}$, can be estimated without bias even if data from just three cross-sections are available. On the other hand, it should also be noted that in estimation we have not accounted for the following three complications: 1) there might be heterogeneity in preferences within social reference groups, 2) we use the sample mean of hours of work instead of the population mean, and 3) we do not have information on individuals' actual social reference groups. Our Monte Carlo simulations can shed some light on the extent to which these complications might bias our results.

In the simulations we studied a data generating process like model 2. However, we also included random preferences in the data generating process. According to simulation 2, the simplified estimation method (that does not take account of random preferences within groups) performs well in the sense that parameter estimates remain virtually unbiased. There is, however, a tendency for the asymptotic standard errors to slightly underestimate the true sampling variability. Hence, if data are generated with random preferences, we might suspect that the t-values in table 5 are overstated, but probably not to the extent that it affects the outcome of conventional significance tests. According to simulation 1, replacing the population variable $\bar{h}^{jt}$ by its sample counterpart appears to be a harmless operation; neither does it seem to bias the estimated coefficients, nor does it lead to any bias in the asymptotic variances for the interdependence and wage rate coefficients.

The complication that we feel might be important is the fact that we presumably construct reference groups in such a way that some individuals are assigned to an incorrect group. Simulation 3 suggests that misclassification errors will give rise to a positive bias in the estimate of η (the parameter corresponding to $\bar{h}^{jt}$), i.e. we may exaggerate the importance of interdependent behavior. The simulation also indicates that the misclassification of individuals causes significantly larger variances for all coefficients, and that this is not captured by the asymptotic variance-covariance matrix in an accurate way. Hence, we might suspect that the t-values in table 5 are upward biased. At the same time, we realize that these findings, to some extent, can be attributed to the particular setup; alternative misclassification schemes may give rise to very different results, which makes it very difficult to assess the effects of misclassification. However, given the problems to identify social reference groups, we feel that our estimate of the ”interdependence effect” should be interpreted carefully. Nevertheless, we do believe that our results support the idea that interdepedent behavior is an important economic phenomenon.

5. SUMMARY AND CONCLUSIONS

The predicted effect of tax reform on labor supply can be seriously biased if preferences are interdependent, but this is neglected when the effect of taxes is calculated. In the present study we set out to test empirically whether interdependent behavior is an important factor in the determination of individuals' labor supply. Taking full account of progressive taxation and income-related transfer programs, we test the hypothesis that individuals' choices of hours of work are influenced by the average hours of work in a social reference group.

There are several nontrivial estimation problems that have to be dealt with. First, to be able to identify the impact of interdependent behavior, it is important to control for preference variation across groups. In the context of the model set out in the paper, neglecting preference variation across groups means that the estimate of the coefficient for average hours of work may capture both preference interdependence and preference variation. In section 2.2 we show that it is not possible to disentangle preference variation across groups from preference interdependence if only cross-section data from one point in time are available. Instead, we need data such that there for a given reference group is a variation in average hours of work. Such data can, for example, be generated if we observe the same group at several points in time, and there are substantial tax reforms in between. In the present paper we use three Swedish cross-sections collected between 1974 and 1991, a time period during which the tax schedule changed considerably.

A second problem concerns the classification of individuals into social reference groups. Since our data contain no prior information that may help to classify observations into reference groups, we have to resort to a rather simple classification strategy. We classify observations into eight groups according to educational level, age and the presence of children under 18 years of age in the household. Hence, given this crude way of assigning individuals to groups, it seems quite likely that some individuals are assigned to an incorrect reference group. We have addressed this problem by performing a Monte Carlo simulation, where 50 percent of the sampled observations are assigned to an incorrect reference group. The results suggest that the coefficients (except for the constant terms) are only moderately biased, wheras standard errors might be substantially underestimated.

In the empirical analysis we estimate a sequence of nested labor supply models. The first model, which is the most general, allows the intercept term to vary freely both across groups and over time. In the second model we impose the restriction that the group effects are time independent, which makes it possible to distinguish these group effects from interdependent behavior. In the third model the group specific effects are excluded, while the fourth specification is a standard labor supply model which neither contains group specific effects nor interdependent behavior. A sequence of likelihood ratio tests imply (*i*) we cannot reject the second specification in favor of the first, which suggests that the group and time effects are interpretable as representing (time independent) group effects and interdependent behavior, (*ii*) we cannot reject that the group specific effects are absent, and (*iii*) we can reject the fourth specification in favor of the third, which implies that the model with interdependent behavior is preferred to the standard labor supply model. In our preferred specification, the coefficient that captures interdependent behavior turns out highly significant. This implies that conventional tax policy predictions, in which preference interdependencies are neglected, will tend to underestimate the effect of a tax reform on hours of work. Accepting the coefficient at face value, the implication is that conventional calculations would capture only about a third of the actual change in hours of work.

Several earlier studies have found support for the hypothesis of interdependent behavior. We believe that our empirical results strengthens the idea that interdependent behavior is an important economic phenomenon. However, the number of empirical studies is still quite small, and much research remains to be done in order to increase our understanding of how interdependent behavior should enter the econometric model. Given that the concept of social reference groups is valid, it would be particularly useful if future data sources could collect detailed survey information on individuals’ social environment and reference groups.

APPENDIX A

The purpose of this appendix is to derive the likelihood function when the complication that arises due to interdependent behavior is fully accounted for. We will then show that we in practice can neglect the interdependence term in the likelihood function. To simplify the exposition we ignore some minor difficulties that are associated with nonlinear taxes. (We take nonlinear taxes into account in the likelihood function used for actual estimation.) To simplify we also drop the income effect and the superscripts *j* and *t*.

We assume there are *N* individuals in the social reference group. In the following we think of this as the population. The data generating process is given by

$$\tilde{h}_i = \mu + \alpha w_i + \eta \bar{h} + \varepsilon_i \qquad i = 1,...,N \tag{A1}$$

where $\bar{h} = (1/N)\sum_{i=1}^{N}\tilde{h}_i$ and ε_i are i.i.d. and $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$. Let

$$\varepsilon_i = \tilde{h}_i - \mu - \alpha w_i - \eta \bar{h} = F_i(\tilde{h}_1, \tilde{h}_2, ..., \tilde{h}_N) \tag{A2}$$

Then the joint density $g(\tilde{h}_1,...,\tilde{h}_N)$ is given by

$$g(\tilde{h}_1,...,\tilde{h}_N) = \prod_{i=1}^{N} \frac{1}{\sigma_\varepsilon}\phi\left[\frac{1}{\sigma_\varepsilon}F_i(\tilde{h}_1,...,\tilde{h}_N)\right]\cdot |J| \tag{A3}$$

where *J* is the Jacobian with typical element $\partial F_i(\,)/\partial \tilde{h}_k$. For *i*=*k* we obtain $\partial F_i/\partial \tilde{h}_i$ = 1-(η/N) and for $i \neq k$ we have $\partial F_i/\partial \tilde{h}_k = -\eta/N$.

Introducing the notation $a = \eta/N$, we can write the Jacobian as:

$$|J| = \begin{vmatrix} 1-a & -a & . & . & -a \\ -a & 1-a & . & . & -a \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ -a & . & . & . & 1-a \end{vmatrix} \tag{A4}$$

This matrix has dimension $N \times N$. Using a result in Mardia et al. (1979) we find that the Jacobian is $|1 - N \cdot a|$, or

$$|J| = |1-\eta| \tag{A5}$$

(A5) will be referred to as the "interdependence" term. We can hence write the log-likelihood function as:

$$\ln \mathrm{L} = \ln|1-\eta| + \sum_{i=1}^{N} \ln\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{\tilde{h}_i - \mu - \alpha w_i - \eta \bar{h}}{\sigma_\varepsilon}\right)\right] = \ln|1-\eta| + \ln\tilde{\mathrm{L}} \qquad \text{(A6)}$$

where

$$\ln\tilde{\mathrm{L}} = \sum_{i=1}^{N} \ln\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{\tilde{h}_i - \mu - \alpha w_i - \eta \bar{h}}{\sigma_\varepsilon}\right)\right] \qquad \text{(A7)}$$

This likelihod function resembles the likelihood function obtained in the time series framework with autocorrelated errors (see Amemiya (1985) p. 176). As for that likelihood function we can define two estimators. One is defined as the estimates that maximize lnL and the other as the estimates that maximize $\ln\tilde{\mathrm{L}}$. For small N the two estimators will differ. Within the time series framework Beach and MacKinnon (1978) show that for small values of N it can be important to include the Jacobian term. However, the asymptotic properties of the two estimators are the same. We do not know the exact size of N in our application, but for the social reference groups we use N are in the order of magnitude of 50,000 or more. Hence, in our estimations we have dropped the "interdependence" term $\ln|1-\eta|$. However, we want to emphasize that for small values of N it might be important to include the "interdependence" term.

There is one further complication. We can not observe the population but only a random sample of *n* observations. The likelihood function we can use therefore takes the form:

$$\ln\hat{\tilde{\mathrm{L}}} = \sum_{i=1}^{n} \ln\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{\tilde{h}_i - \mu - \alpha w_i - \eta \hat{\bar{h}}}{\sigma_\varepsilon}\right)\right] \qquad \text{(A8)}$$

where the population mean $\bar{h}$ is replaced by the sample mean $\hat{\bar{h}}$. Since the sample mean is a consistent estimator of the population mean this should still yield consistent estimates. The fact that we use an estimate of the population mean might, however, lead to a bias in our estimate of the variance-covariance matrix. The sampe size we use is fairly large so this might not be a problem in practice. We investigate this in our Monte Carlo simulations.

APPENDIX B

Table B1. Summary statistics: sample means and standard deviations

	1973		1980		1990		
Variable	*Mean*	*St.dev.*	*Mean*	*St.dev.*	*Mean*	*St.dev.*	*Unit*
Hours of work	2.135	0.263	2.091	0.244	2.120	0.327	1000’s of hours
Gross wage rate	40.2	14.9	40.8	14.1	45.2	16.1	SEK
CAP	-6.998	9.318	-10.345	16.672	-17.141	17.660	1000’s of SEK
BBF	6.978	7.736	4.663	4.121	1.525	2.409	1000’s of SEK
MDI	23.889	17.040	42.089	24.049	39.012	13.575	1000’s of SEK
NC	1.30	1.07	1.21	1.06	1.43	1.10	No. of children
AGE	42.0	10.2	39.8	10.5	43.4	8.9	Years
No. of observations	777		863		678		Observations

Note: All economic variables are adjusted to the 1980 price level (CPI73=1.988, CPI90=0.482). The exchange rate SEK/$(U.S.) was, on average, 4.37 in 1973, 4.23 in 1980 and 5.92 in 1990.

Table B2. Reference groups: percent of sample, hours worked and net wage rates

	1973			1980			1990		
Group	*% of sample*	$\overline{w}$	$\overline{h}$	*% of sample*	$\overline{w}$	$\overline{h}$	*% of sample*	$\overline{w}$	$\overline{h}$
1	33.6	15.1	2.130	29.9	14.3	2.089	19.5	17.7	2.074
2	7.7	18.2	2.115	12.3	16.6	2.140	12.7	19.9	2.130
3	3.6	14.3	2.129	9.2	14.3	2.083	2.7	18.1	2.161
4	2.0	17.6	2.157	4.1	14.2	2.066	1.6	18.1	2.315
5	25.7	15.6	2.153	20.0	14.1	2.090	27.6	18.5	2.122
6	4.1	22.3	2.137	5.4	15.7	2.134	14.7	21.8	2.152
7	21.0	15.0	2.112	16.8	13.9	2.088	18.1	19.1	2.108
8	2.3	21.9	2.210	2.3	16.3	2.154	3.1	23.2	2.133

Note: The mean net wage rate, $\overline{w}$, is measured in SEK and hours of work, $\overline{h}$, in thousands of hours. Net wage rates are adjusted to the 1980 price level (CPI73=1.988, CPI90=0.482). We calculate net wage rates by linearizing individuals’ budget sets around observed hours of work.

Table B3. The 1990 national tax system

Basic tax		Added tax	
Taxable income (SEK)	*Marginal tax rate*	*Tax base (SEK)*	*Marginal tax rate*
-75,000	0.03	-140,000	0
75,000-	0.17	140,000-190,000	0.14
		190,000-	0.25

Note: The tax bases differed for the basic and the added tax, because certain deductions were not allowed when paying the added tax.

REFERENCES

Alessie R. and A. Kapteyn (1991), "Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System", *Economic Journal* 101, 404-419.

Amemiya, T. (1985), *Advanced Econometrics*, Cambridge, Mass. Harvard University Press.

Andreoni J. and J.K. Scholz (1990), "An Econometric Analysis of Charitable Giving with Interdependent Preferences", SSRI Workshop Series 9023, University of Wisconsin.

Beach, C. M. and J. G. MacKinnon (1978), "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors", *Econometrica* 46, 51-58.

Blomquist S. (1983), "The Effect of Income Taxation on the Labor Supply of Married Men in Sweden", *Journal of Public Economics* 22, 169-197.

Blomquist S. (1993), "Interdependent Behavior and the Effect of Taxes", *Journal of Public Economics* 51, 211-218.

Blomquist S. and U. Hansson-Brusewitz (1990), "The Effect of Taxes on Male and Female Labor Supply in Sweden", *Journal of Human Resources* 25, 317-357.

Burtless, G. and J. A. Hausman (1978), "The Effect of Taxation on Labor Supply", *Journal of Political Economy* 86, 1103-1130.

Hausman J.A. (1979), " The Econometrics of Labor Supply on Convex Budget Sets", *Economic Letters* 3, 171-174.

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

Kapteyn A. (1977), A Theory of Preference Formation, Ph.D. dissertation, Leiden University.

Kapteyn A., S. van de Geer, H. van de Stadt and T. Wansbeek (1997), "Interdependent Preferences: An Econometric Analysis", *Journal of Applied Econometrics* 12, 665-686.

MaCurdy T., D. Green and H. Paarsch (1990), "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply", *Journal of Human Resources* 25, 415-490.

Manski C. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies* 60, 531-542.

Mardia K.V., J.T. Kent and J.M. Bibby (1979), *Multivariate Analysis*, Academic Press, London.

Sacklén, H. (1996), Essays on Empirical Models of Labor Supply, Ph.D dissertation, *Economic Studies* 27, Department of Economics, Uppsala University.

Woittiez I. (1990), *Modelling and Empirical Evaluation of Labour Supply Behavior*, Berlin: Springer Verlag.

Woittiez I. and A. Kapteyn (1998), "Social Interactions and Habit Formation in a Model of Female Labour Supply", *Journal of Public Economics* 70, 185-205.