Piece-Wise or Differentiable Budget Constraint? Estimating Labour Supply Function for Finnish Females

Kuismanen, Mika

Piece-Wise or Differentiable Budget Constraint? Estimating Labour Supply Function for Finnish Females

Journal:	*Applied Economics*
Manuscript ID:	APE-05-0177.R1
Journal Selection:	Applied Economics
JEL Code:	C31 - Cross-Sectional Models|Spatial Models < , H24 - Personal Income and Other Nonbusiness Taxes and Subsidies < , J22 - Time Allocation and Labor Supply <
Keywords:	Labour supply, Progressive income taxation, Budget constraints, Maximum likelihood estimation

Piece–Wise or Differentiable Budget Constraint? Estimating Labour Supply Function for Finnish Females

Mika Kuismanen
European Central Bank
Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany.
Tel: 49-69-1344 7607
Fax: 49-69-1344 6575
Email: mika.kuismanen@ecb.int

March 2005

Abstract

Various estimation approaches have been used in recent literature to study the effect of non-linear income taxation on labour supply. Different techniques and data sets have produced wide range of income and substitution elasticities. In this study we utilise register data provided by tax authorities. This gives us good possibilities to construct detailed budget constraints for all individuals in our sample. We estimate labour supply function using the piece–wise linear budget constraint approach and the differentiable budget constraint approach suggested by MaCurdy et al. (1990) Our results support the view that if one is able to mimic the actual budget set closely and if the degree of progression is high then these two methods are likely to produce similar results. On the other hand, if the above mentioned factors are not present then the differentiable budget constraint approach is likely to be the safer choice.

1 Introduction

The purpose of this paper is to provide results conserning labour supply behaviour in Finland. It is known that the economic theory does not give much prediction power on how taxation affects labour supply because it leaves the signs of substitution and income elasticities open, see *e.g.* Blundell and MaCurdy (1999). In addition, introducing different kind of welfare systems may create non-convex budget sets, and thus, certain areas of the budget constraint cannot correspond the utility-maximising points. In many cases we do not even know if an increase in marginal tax rate will increase or decrease supplied hours. So, it seems that empirical research is needed to give us information concerning tax and benefit systems' behavioural effects.

Our paper compares two different empirical approaches to model labour supply behaviour when progressive income taxation is taken into account. We start by a conventional piece-wise linear method. In this approach we try to mimic the actual budget constraint as well as possible and then it is fully taken into account in the estimation procedure. Above approach means that the likelihood function takes into account the choice of hours over the entire exogenous tax schedule removing the endogeneity problem which is present in the simpler approaches, like in the linearised budget constraint approach. This method has been criticised for various reasons (see Heckman (1983)). First, to mimic the budget constraints requires a lot from the data. Secondly, it is questionable if econometrician can measure the constraint accurately and if individuals really know the actual shape of it.

To avoid the above mentioned problems but still allow the convex shape of the budget constraint we follow MaCurdy et al.(1990) to construct a differentiable budget constraint. The central idea is to approximate the actual piece–wise linear constraint by using continuous smooth polynomial function because it is unlikely that individual knows exact shape of her budget set. Technically approach means mimicing the tax schedule by fitting a polynomial function to the marginal tax rates. After integrating this function we get a differentiable relation which approximates the amount of total taxes. This method is much easier to estimate since a purely continuous distribution describes the hours of work decision.

Finnish Labour Force (LFS) data for the year 1989 is used in estimation. The sample consists of 2037 married females aged 25 to 60. Income data for these females and their partners is drawn from the Tax Register Data and then merged with LFS. This gives us a rich data set to build budget constraints. Subsection 4.1 includes more thorough discussion on the data. As non–linearities mainly arise from the income tax system a short description of it is given in the end of section two.

What we find is that a differentiable budget constraint is able to approximate the piece–wise linear one quite closely and the estimation results do not differ significantly between these two approaches. This result is in line with the study by MaCurdy and Flood (1992) but conflicts with the original MaCurdy et al. study. Our conclusion is that in the case of a high degree of progressivity these methods are likely to produce similar outcomes because then the polynomial function is simply capable of producing almost identical constraints compared to the piece–wise linear one.

The set up of this paper is as follows. In section 2 we shortly present the basic ideas behind the piece–wise linear budget constraints. After that we go through two estimation approaches for the labour supply function in the presence of non–linear budget constraint. Section 4 presents estimation results and section 5 concludes the paper.

2 Budget sets

Let us start by analysing a simplified income tax system with progressive elements. The income tax system consists of three tax brackets and thus three marginal tax rates; $t_1, t_2, and\, t_3$. Inside the tax brackets marginal tax rate is constant but it increases with income. The outcome is the familiar piece-wise linear budget constraint with four kink points.

Marginal tax rate t_1 leads to the net wage $w_1 = (1 - t_1)w$ and this corresponds to the first segment in tax schedule. Correspondingly the net wage rate in the second segment is $w_2 = (1 - t_2)w$ etc. $H_0, H_1 and\, H_2$ are kink points where the marginal tax rate changes and H_n stands for the upper limit of labour supply. y_1 is the exogenous income component and thus does not depend on hours supplied. Note that this component is directly observed from the data. y_2 and y_3 are called "virtual income" terms and must be calculated recursively in the following way

$$y_i = y_{i-1} + (w_{i-1} - w_i)H_{i-1}$$

Thus geometrically we just extend a given budget segment to the vertical axis. The crucial thing to realise is that these components cannot be observed directly from the data. Given this budget constraint the consumer makes his or her labour supply decision and the optimal labour supply can vary from zero hours to the maximum number of hours.

For simplicity we assume that non-labour income is not taxed and E (can also vary between individuals) stands for the tax deductions. In this case the taxable income I_i is earned income (wh) less the deductions and thus the kink points can be calculated as

$$H_i = (I_i + E)/w$$

When the budget set is convex and consumer preferences can be represented by the quasi-concave utility function, $u(c, h)$, which is non-decreasing with respect to consumption and non-increasing with respect to supplied hours, h.

But what are the pros and cons of this approach? First of all, this method allows us to construct[1] budget sets which recognise all the institutional features of the tax and social security system. The piece–wise linear method also treats the marginal tax rate endogenously in the estimation procedure. It also allows us to incorporate different stochastic assumptions, like allowing randomness in hours to arise from measurement (optimisation) and/or unobserved individual preferences. We can introduce fixed costs of working into the model quite easily and the treatment of unobserved wages can be done in different ways. MaCurdy et al.(1990) includes an excellent theoretical discussion about how unobserved wages should be calculated[2].

The fundamental assumption behind this model is that the observed market behaviour is the outcome of free rational choice subject to piece–wise linear budget constraint. In other words, we assume that on individual has perfect knowledge of the budget set and an econometrician is able to measure all the budget set variables without error. It is hard to imagine that the above mentioned criteria are actually met either by the econometrician or the decision making individual.

Probably the most serious claim against piece–wise linear modelling is the so–called MaCurdy-critique. MaCurdy et. al. claimed "that the this method requires the satisfaction of parametric restrictions that constraint the signs of estimated substitution and income effects" In his articles Blomquist (1995) (1996) has written that MaCurdy's claim is not generally true.

Is there then a way to overcome above mentioned problems? In their influential paper, MaCurdy et al. developed a new approach which utilises a differentiable function to approximate marginal tax function. The idea sounds complicated, but as we will show below, it turns out to be a simple and attractive alternative.

We can think that individuals do not know the exact shape of their budget set, but they do have an idea of its approximate shape. For example, most likely individuals do not exactly know the number of supplied hours when moving from one tax brackets to another or what their accepted tax deductions might be etc..

A differentiable budget constraint approach is also easier to apply in practise because purely continuous distribution describes the supplied hours for all individuals whose hours are strictly positive. If we assume that the constraint is convex and consumers' preferences are strictly quasi-concave then we can be sure that we can find a tangent point and this point is going to be a unique one. In addition, because we have "smoothed out" all the segments and kinks we are left with one continuous non-linear segment, we do not need to write the tax algorithm into the likelihood function as can be seen in the next section.

Finnish income tax system in 1989 basically had two parts. A progressive state income tax and a proportional local (municipal) income tax. In addition, individuals contribute to the National Pension Insurance scheme (1.55 percent from the taxable income) and National Health Insurance scheme (1.25 percent from the taxable income), which are proportional to income changes. Roughly speaking, the tax liability in state tax and municipal tax is the same excluding the tax deduction system. In 1989 the state income tax schedule was composed of seven marginal tax rates varying from 0 to 44 percent (thus having 7 tax brackets *i.e.* piece–wise linear segments) and the municipality tax rate varied from 14 percent to 19.5 percent. We also have developed a formula to calculate state tax deduction for all persons in the sample. Estimated tax deductions varied from 0 FIM to 29 500 FIM. See also subsection 3.2 for an example.

3 Two estimation approaches for the labour supply function in the presence of non–linear budget constraint

In the case of linear income tax system (or when income tax is ignored) estimation of labour supply function is straightforward. This is because the budget constraint individual faces is linear and there can only be one utility maximisation point (*i.e.* observed and optimal hours lie on the same linear segment). This changes when non–linearity is present (progressive income taxation). For example, when the consumer faces piece–wise linear budget constraint her observed and optimal labour supply may lie on different segments. This possibility has to be taken into account in estimation. We start by showing how this can be tackled by construction of tax algorithm which is then substituted to the likelihood function in final estimation stage. In subsection 3.2 below we introduce an alternative method where tax algorithm is not needed even when we allow non–linear budget constraints.

3.1 Piece–wise linear approach

In the case of piece–wise linear budget constraint economic theory predicts bunching of observations (hours of work) at kink points (where marginal tax rate changes) or just below of them. Reason is that if individuals increase their hours they will move to an upper tax bracket (or, for example, in social security system they move to the point where the credit is taxed away) where they will face higher marginal tax rate. Blundell, Duncan and Meghir (1998) provides an example where bunching was found. In Kuismanen (2002) it is shown that in Finnish data bunching was not an issue. Moffitt (1986) states that if observations are distributed evenly across the budget sets it provides a reason to introduce measurement error term into the model. Naturally, there are also other reasons to introduce measurement error, such as reporting errors in measured hours.

We start with a basic measurement error approach where the general labour supply function can be written as $h_i = h_i^*(w_i, y_i, z_i; \alpha, \beta, \gamma) + \varepsilon_i$. α, β and γ are parameters to be estimated. Vector z includes individual characteristics (*e.g.* socio-economical and demographic variables) and the variables w and y represent the marginal wage and virtual (non–labour) income variables correspondingly. ε represents the measurement/optimisation error. Subindex i denotes the individual.

In the statistical model we have to calculate the densities of h_i and this naturally requires evaluation of the maximum utilities received on each linear segment of the budget constraint. More formally, we now write the problem as

$$
\begin{aligned}
f(h_i) &= P[h_i = 0] + P[h_i > 0] * f(h_i \mid h_i > 0) + P[h_i = H_n] \\
&= P[\quad \text{at zero} \quad] \\
&+ P[\quad \text{below kink 1} \quad] * f(h_i \mid \quad \text{below kink 1} \quad) \\
&+ P[\quad \text{at kink 1} \quad] \\
&+ P[\quad \text{above kink 1} \quad] * f(h_i \mid \quad \text{above kink 1} \quad) \\
&\;\;. \\
&\;\;. \\
&\;\;. \\
&+ P[\quad \text{at maximum} \quad]
\end{aligned}
$$

The optimal supply of hours h^* can be found from the segment k ($k = 1, \ldots, n$), if

$$H_{k-1} < h^*(w_k, y_k, z; \alpha, \beta, \gamma) < H_k$$

Intuition behind this calculation rule is the following: after calculating the slope of the indifference curve from the direct utility function[3] we replace consumption c ($= w_k h + y_k$) by individual's income (calculated for all the segments) and then equate the slope of the indifference curve and the marginal wage w_k corresponding that segment. The algorithm iterate as long as this condition is satisfied. If in some cases we cannot find the solution we start to look it from the kink points.

Optimum h^* is found from the kink point H_k ($k = 1, \ldots, n-1$), if

$$h^*(w_k, y_k, z; \alpha, \beta, \gamma) \geq H_k \qquad \textit{and} \qquad h^*(w_{k+1}, y_{k+1}, z; \alpha, \beta, \gamma) \leq H_k.$$

Another way to express above condition is that the optimum can be found from the H_k when the slope of the indifference curve is bigger or equal than w_{k+1} and the slope is smaller or equal than w_k.[4]

The above formulation shows how optimal hours can be calculated under the progressive income taxation when the following aspects are known: tax schedule, hourly wage, exogenous (non–labour) income and the shape of the labour supply function (or correspondingly the form of the direct utility function).

Despite h^* can be calculated quite easily in the convex case the maximisation of the (log) likelihood function is not straightforward because h^* is not always well–behaving function respect to the parameters. First, the log–likelihood function is not differentiable everywhere (kink points) and secondly there can be parameter values where the function becomes flat. This can become a serious problem if there are not enough variation between the budget sets.

In the above discussion we did not make any specific assumptions about the stochastic specification *i.e.* we assumed that all variance in hours conditional on covariates is measurement/optimising error.[5] This means that preferences are assumed to be non–stochastic *i.e.* all variation in preferences is only due to observable personal characters (i.e. regressors). This is the usual procedure adapted by researchers, at least when using cross–section data sets. In the context of labour supply we might think that there exists different sources for stochastic disturbances to arise from. First, the usual measurement error interpretation implies that the parameters to be estimated are the same for all individuals, so there is only one utility maximising choice in the population (this is probably a questionable outcome). Second, there might exist randomness in preferences which is not captured by the variables we include in our regression function. As Moffitt (1986) argues it is reasonable to expect that at least some amount of the observed distribution of observations over the budget constraint is a result of heterogeneous preferences. It is natural to think that both aspects are relevant in the context of labour supply. Look Hausman (1985) for the other possible sources of stochastic disturbances.

The two different stochastic elements mentioned above have different implications for the data. In the standard measurement/optimisation error approach the observations should be distributed evenly over the whole budget constraint, as in the case of heterogeneity of preferences we should find clusters of observations at the kink points (in the case of convex budget constraint). In theory, the model which only includes heterogeneity term is possible, but not a very appropriate one for the most applications because it is unlikely that all observations are clustered to the kinks. Although empirical evidence shows some clustering (especially in the cases of big changes in marginal tax rates) it is not usually strong enough to leave measurement error term without modelling. One relatively easy way to proceed is to estimate the model with random preference term and then test if its variance is different from zero. One further motivation for including the heterogeneity term is that in the cross–section studies one finds a large amount of unexplained variance.

Let us now write the labour supply funcion as follows (note that we have dropped subindex i for notational simplicity)

(1) $$h = h^*(w, y, z, \eta; \alpha, \beta, \gamma) + \varepsilon,$$

where the expression for the desired/optimal hours h^* (below semi-logarithmic expression as in estimation) now includes the additive random variable η

(2) $$h^* = \alpha lnw + \beta y + \gamma z + \eta.$$

Maximum likelihood estimation requires the specification of these two stochastic terms (η, ε). We assume that they are independently normally distributed as

$$\varepsilon \sim N(0, \sigma_\varepsilon^2)$$

$$\eta \sim N(0, \sigma_\eta^2)$$

Reason for the independence assumption is following. If we interpret the terms as heterogeneity and as measurement (or optimisation) error, there are no reasons to expect them to be generated from a joint process.

In principle we can introduce the unobserved heterogeneity in many different ways. The most common solution in the literature has been to allow substitution or income elasticity to vary across individuals. We allow the constant term to vary between individuals because of the following reasons. We do not want a priory to restrict signs of substitution or income effects. As an example in the most used approach truncated normal distribution is used to force the substitution elasticity to be positive. Our strategy leads to a nonrestrictive specification without any theory based restrictions.

In this two random term model the algorithm to find optimal amount of labour supply can be constructed in the following way. Optimum h^* can be found from the segment k(k=1,...,n) if

$$\eta_{kl} < \eta < \eta_{ku}$$

where

$$\eta_{kl} = H_{k-1} - \alpha lnw_k - \beta y_k - \gamma z$$
$$\eta_{ku} = H_k - \alpha lnw_k - \beta y_k - \gamma z.$$

Above the subindex l indicates to the lower limit of the segment k and respectively the subindex u indicates to the upper limit of the segment k. To derive expressions for kinks is straightforward and not shown here.

We can now express the corresponding probabilities using the integrals. For example, the probability that the optimum is located on the second segment is

$$pr(h^* \, is \, on \, segment \, 2) = \int_{\eta_{2l}}^{\eta_{2u}} \left(\frac{1}{\sigma_\eta}\right) \Phi \left(\frac{\eta}{\sigma_\eta}\right) d\eta.$$

Next we proceed to likelihood functions used in estimation.[6] As above, we first start from the approach where we do not allow any individual heterogeneity and the error term is interpreted to be optimising or measurement error. Observed hours h may then deviate from the desired hours h^* by the amount of the optimising or measurement error ε, thus $h = h^* + \varepsilon$. We assume that $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and that $E(\varepsilon|h^*) = 0$.

It is natural to think that **observed** hours are generated by the following generalised Tobit–model

$h = 0$ if $h^* + \varepsilon = 0$
$h = h^* + \varepsilon$ if $0 < h^* + \varepsilon < H_n$
$h = H_n$ if $h^* + \varepsilon \geq H_n$

and the corresponding Likelihood Function can now be written as

$$(3) \quad L = \prod_{i \in A} \left[1 - \Phi\left(\frac{h^*}{\sigma_\varepsilon}\right)\right] \prod_{i \in B} \left[\frac{1}{\sigma_\varepsilon} \phi\left(\frac{h_i - h^*}{\sigma_\varepsilon}\right)\right] \prod_{i \in C} \left[1 - \Phi\left(\frac{H_n - h^*}{\sigma_\varepsilon}\right)\right].$$

where,

i belongs to index set A when $h = 0$

i belongs to index set B when $0 < h < H_n$
i belongs to index set C when $h \geq H_n$.
$\phi(\cdot)$ is Standardised Normal Density Function and $\Phi(\cdot)$ is Cumulative Normal.

The first part of the likelihood function corresponds individuals whose observed hours are zero. The second part corresponds those individuals whose observed hours are in one of the segments or kink points (in estimation the tax algorithm operates here) and the third part corresponds those whose observed hours are at maximum. Note that the second part without tax algorithm would only be used in the simple linear case.

Next step is to construct the likelihood function with two random terms. This is demanding because level of labour supply is an outcome from the two random terms *i.e.* we have to take into account all the combinations which can produce a certain level of hours.

As motivated earlier the stochastic specification is important when we face non–linear budget constraints and the error term has a more specific interpretation in these models. The most important drawback in measurement/optimisation error model is its restrictiveness to the labour supply responses. For example, according theory a change in the marginal tax rate in the case of convex budget set would have identical zero effect on the labour supply for all individuals not located on that segment (when we do not take income effects into account). In other words, change in slopes of the other segments do not have any behavioural effects.

We want to stress that η is not estimated for all individuals separately because this would require estimation of more parameters than we have observations. We estimate parameters of the distribution function of η. Each person's η is considered to be a random drawing from this distribution. In our case we assume that η follows normal distribution with $(0, \sigma_\eta^2)$ and when it is uncorrelated with ε we can derive the following likelihood function in the case of two additive random terms.

$$L(\alpha, \beta, \gamma, \sigma_\eta, \sigma_\varepsilon : w_k, y_k, z)$$

$$= \prod_{i=1}^{J} \Bigg\{ \left(\frac{1}{\sigma_\varepsilon}\right) \phi\left(\frac{h_i}{\sigma_\varepsilon}\right) \left[\Phi\left(\frac{(-\alpha lnw_{i1} - \beta y_{i1} - \gamma z_i)}{\sigma_\eta}\right)\right]$$

$$+ \sum_{k=1}^{n} \left(\frac{1}{\sigma}\right) \phi\left(\frac{e_{ik}}{\sigma}\right) \left[\Phi\left(\frac{\eta_{iku} - \left(\frac{\sigma_\eta^2 e_{ik}}{\sigma^2}\right)}{(\sigma_\varepsilon \sigma_\eta / \sigma)}\right) - \Phi\left(\frac{\eta_{ikl} - \left(\frac{\sigma_\eta^2 e_{ik}}{\sigma^2}\right)}{(\sigma_\varepsilon \sigma_\eta / \sigma)}\right)\right]$$

$$+ \sum_{k=1}^{n-1} \left(\frac{1}{\sigma_\varepsilon}\right) \phi\left(\frac{(h_i - H_{ik})}{\sigma_\varepsilon}\right) \left[\Phi\left(\frac{\eta_{i(k+1)l}}{\sigma_\eta}\right) - \Phi\left(\frac{\eta_{iku}}{\sigma_\eta}\right)\right]$$

$$+ \quad \left(\frac{1}{\sigma_\varepsilon}\right)\phi\left(\frac{h_i - H_n}{\sigma_\varepsilon}\right)\left[1 - \Phi\left(\frac{H_n - \alpha lnw_{in} - \beta y_{in} - \gamma z_i}{\sigma_\eta}\right)\right]\Bigg\},$$

where $e_{ik} = h_i - \alpha lnw_{ik} - \beta y_{ik}$ and $\sigma^2 = \sigma_\varepsilon^2 + \sigma_\eta^2$.

In the above likelihood the first term corresponds the probability that an individual's worked hours are zero. Second term corresponds that the optimum lies in some segment and the third term corresponds case where optimum can be found from some of the kink points. The last term corresponds the case that individual works maximum amount of hours. Note again that the corresponding tax algorithm will operate inside this likelihood function in estimation.

3.2 Differentiable Budget Constraint Approach

Differentiable budget constraint technique was first introduced by MaCurdy et al. (1990) and our presentation will follow it with suitable modifications to take account the Finnish tax system. Intuition behind this method is to approximate the tax schedule by fitting a function to the marginal tax rate. After integrating marginal tax function we get a differentiable relation approximating the amount of total taxes paid as a function of taxable income.

Let us introduce some new notation. Denote $I(h)$ for taxable income at h hours of work and $M[I(h)]$ for marginal tax rate function. Now, for example the simplified three bracket income tax system can be presented in the following way:

$$\begin{aligned} M[I(h)] &= t_1 \text{ [from } I(H_0)\, to\, I(h_1)] \\ &= t_2 \text{ [from } I(H_1)\, to\, I(h_2)] \\ &= t_3 \text{ [above } I(H_2)] \end{aligned}$$

where H_i denotes the kink points where marginal tax rates t_i changes. To approximate the marginal tax rate schedule the function must fit a step function presented above closely and it still should be differentiable at the step points (i.e. kink points). MaCurdy et al. suggested a following kind of approximation

$$(4) \qquad \overline{M[I(h)]} = \sum_{i=1}^{K}[\Phi_i(I(h)) - \Phi_{i+1}(I(h))] * p_i(I(h)),$$

where $\Phi_i(I(h))$ denote the cumulative normal distribution function evaluated at the income level I(h) with mean μ_i and variance σ_i^2. The idea is that the difference $\Phi_i(I(h)) - \Phi_{i+1}(I(h))$ takes value of one over the range where t_i is relevant and zero elsewhere. Now, we can control this by adjusting the mean and the variance. Adjusting the mean we can control the moment when the value of one begins and ends, and adjusting the variances we can control how quickly this happens. The trade–off here is the smoothness of transition against the precision. $p_i(I(h))$ are the polynomials in income. For example, in Finnish case in 1989 there were 7 tax brackets, so we can set $K = 7$ and p_i is the marginal tax rate t_i associated with the ith tax bracket.

To see how the above presented generalisation works let us go back to our simplified three bracket tax schedule discussed above. In this case we have three marginal tax rates $t_i < t_2 < t_3$, so we have "three segments to smooth out". We can now write our approximation function using the above notation for this problem as

$$\overline{M[I(h)]} = [\Phi_1(I(h)) - \Phi_2(I(h))] * t_1 + [\Phi_2(I(h)) - \Phi_3(I(h))] * t_2 + \Phi_3(I(h)) * t_3$$

So, the first segment has a height t_1(can be thought as a flat line with a height t_1.) and thus corresponding taxable income is from $I(H_0)$ to $I(H_1)$. This feature is captured by parameterising $\Phi_1(I(h))$ with mean $\mu_1 = I(H_0)$ and correspondingly $\Phi_2(I(h))$ with mean $\mu_2 = I(H_1)$. The first distribution function $I(H_0)$ takes value of one above the income level $I(H_0)$ and zero elsewhere and the second distribution function $I(H_1)$ takes value of zero below the income level $I(H_1)$ and then switches to one above it. So the difference of these functions is one between $I(H_0)$ and $I(H_1)$ and zero elsewhere and correspondingly for all other ranges. So, we can control the switch from zero to one(and vice versa) by adjusting the means. How quickly these switches will take place depends on the values given to the variances.

In 1989 Finnish marginal tax rates varied from zero to 44 percent including 7 tax brackets and the tax exemption level was 36 000 FIM. To put the income tax system into the described framework concerning differentiable budget constraints we get the following three parts. Part one is valid from zero income up to 36 000 FIM and the tax rate for this range is equal to the individuals local tax rate t_l. In the second part income ranges from 36 000 FIM up to 250 000 FIM where the tax rate is the local tax rate plus the monotonically increasing marginal tax rate $t_i, i = 1, ..., 7$. The third part is for the incomes over 250 000 FIM. In this case the tax rate is the local tax rate plus the federal tax rate of 44 percent.

So, according to given information we can write the approximation for the marginal tax rate function in the Finnish case as:

$$\overline{M[I(h)]} = \left[\Phi_1\left(\frac{(I(h)) - \mu_1}{\sigma_1}\right) - \Phi_2\left(\frac{(I(h)) - \mu_2}{\sigma_2}\right)\right] * t_l$$

$$
\begin{aligned}
& + \quad \left[\Phi_2\left(\frac{(I(h))-\mu_2}{\sigma_2}\right) - \Phi_3\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)\right] * F(I) \\
(5) \qquad & + \quad \Phi_3\left[\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)\right] * (t_l + 0.44),
\end{aligned}
$$

where $F(I)$ is a polynomial in taxable income which approximate the increasing tax rates from 36000 FIM to 250000 FIM.

Approximation of the $F(I)$ is done by running the following Ordinary Least Squares regression

$$
(6) \qquad t_i = \xi_0 + \xi_1 * I_i + \xi_2 * I_i^2 + \xi_3 * I_i^3,
$$

where t_i is the marginal *federal* tax rate. What we do is that we create an variable which increases by 100 FIM starting at 36 000 FIM and ending at 250 000 FIM. In other words we create 2140 equally spaced combinations (t_i, I_i). After estimation, we can use the estimates from this third-degree polynomial approximation to specify the above equation to be

$$
F(I) = -1219 + t_l + 7.46 * 10^{-6} * I - 3.80 * 10^{-11} * I^2 + 6.46 * 10^{-17} * I^3,
$$

where t_l is the local tax rate. Our model explains 95 percent of the variance in marginal tax rates.[7] Plugging the above estimated formula into the marginal tax rate function and integrating it with respect to the income we can derive formula which approximates the amount of total taxes paid.

$$
\begin{aligned}
(7)\ T(I(h)) \quad & = \quad \int \overline{M[I(h)]dI} \\
& = \quad \left[t_l\Phi_1\left(\frac{(I(h))-\mu_1}{\sigma_1}\right) - t_l\Phi_2\left(\frac{(I(h))-\mu_2}{\sigma_2}\right)\right] \\
& + \quad \left[\xi_0 \int \Phi_2\left(\frac{(I(h))-\mu_2}{\sigma_2}\right)dI + \xi_1 \int \Phi_2\left(\frac{(I(h))-\mu_2}{\sigma_2}\right)IdI\right. \\
& + \quad \left.\xi_2 \int \Phi_2\left(\frac{(I(h))-\mu_2}{\sigma_2}\right)I^2dI + \xi_3 \int \Phi_2\left(\frac{(I(h))-\mu_2}{\sigma_2}\right)I^3dI\right] \\
& - \quad \left[\xi_0 \int \Phi_3\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)dI + \xi_1 \int \Phi_3\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)IdI\right. \\
& + \quad \left.\xi_2 \int \Phi_3\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)I^2dI + \xi_3 \int \Phi_3\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)I^3dI\right] \\
& + \quad \left[(t_l + 0.44) \int \Phi_3\left[\left(\frac{(I(h))-\mu_3}{\sigma_3}\right)dI\right]\right].
\end{aligned}
$$

From this expression it is relatively straightforward to derive the final form used in estimation. We just need to find analytical solutions to it and this can be done by calculating the above integrals.[8]

This differentiable approach is more straightforward than the piece–wise linear one since a purely continuous distribution describes the hours of work and no tax algorithm is needed. Intuition behind this differentiable approach follows the idea presented by Hall (1973). We can think marginal wage rate and virtual income as a function of working hours. Hall's idea was to linearise the actual non-linear budget constraint at the observed hours. The implied slope of this linearised constraint is the marginal wage rate and the intercept of the vertical axis is the virtual income. In other words, utility maximisation implies a solution for hours of work which can be written in the form of implicit equation, $h = f[w(h), y(h)]$. By applying the Implicit Function Theorem to it we can solve this equation for h and hence derive the labour supply. Here the same idea is used but still allowing non–linearities.

4 Estimation results

4.1 Data

Our study utilises the Finnish Labour Force (LFS) survey data. It is a cross–section data including individuals of age between 15 to 64. In the first stage the sample is drawn from the Finnish Population Census using geographical weights. After that the LFS sample is drawn randomly by age and gender. In 1989 data, the sample size is 7820 individuals. From this we selected females and were left with 4124 observations. For the empirical analysis we selected married women aged 25–60. We also deleted some groups like farmers and self–employed mainly because different tax and social security legislation. The final sample size used in this study is 2037 observations.

Income data of the corresponding individuals in the LFS is drawn from the Tax Register Data and then merged with the LFS. The income information is not based on the survey[9] data and it includes approximately 70 variables on individual's earnings. Of course, it is very unlikely that someone's earnings are composed from all these components. However, the data shows that individuals' earnings come from very different sources. Actually, for some individuals traditionally used income variables do not play any role at all. The income data also includes the same 70 variables for the spouse, so all in all we have approximately 140 variables for married individuals to construct the budget sets. Detailed information how crucial variables like working hours, wages and exogenous incomes are calculated can be found from Kuismanen (2002).

Next we will shortly comment the main features of the data (see also appendix). The participation rate in the selected sample is 72 percent. Unemployment rates vary geographically and the figures are the lowest in the Helsinki metropolitan area and the highest in the eastern and northern part of the country. The average unemployment rate is 3.2 percent. The Blue–collar workers are more likely to have zero hours observations than white–collar ones and the women with two young children have the highest probability to be out of work. In data participants are slightly younger than the non–participants and they are also better educated. It also seems that the likelihood of being a non–participant increases if the spouse is also a non–participant.

4.2 Results

Our main goal in this paper is to estimate similarly specified labour supply functions using the piece–wise linear and differentiable approach. This is because we are interested that will these two methods produce similar results and thus are we able to conclude that, at least using our data, the piece–wise linear method will produce sensible results.[10] We also estimated the unobserved heterogeneity model but the term was not statistically significant and the rest of the covariates were very much in line with those reported here.[11] This finding is in line with the fact that no bunching was found.

Our data set does not have direct information on individuals' hourly wages, thus we have to construct it using the income and the hours of work variables. This means that the possible bias in the hours of work variable also shifts into the marginal hourly wage rate. For example, if worked hours are smaller than the right value then hourly wage rate becomes too high. To get rid of this bias we estimate the log–wage equation using Heckman's selection method and the predicted values are used in the final analysis as an instrument for the hourly wage rate. Results are presented in Table 1.

Table 1. -About Here-

Results show the familiar age–wage effect, i.e. that up to a certain age wage increases and then decreases. Education (Educ10-Educ15 are dummy variables indicating the number of completed years of education. Reference group is individuals with less than 10 years of education.) has a positive effect on wage rate and the effect gets stronger with an increase in the number of years of education. Work experience (measured in years) increases the wage rate up to a certain experience level and then starts to decrease it. Tenure variable (number of years with the same employer) shows the similar quadratic shaped effect. Individuals who have a permanent job get a higher wage, as is to be expected as do individuals who are white–collar workers. Individuals in managerial positions seem to earn more than others. It is also evident that individuals living in the south of Finland earn more.

For working hours we use regular reported weekly working hours also taking into account regular hours in the second job. When calculating the exogeneous income term we took into account the following components: Interest (both taxable and nontaxable), dividend payments, sales profits, regular untaxable pensions, other regular subsidies etc. From all the components which are taxable we have subtracted the corresponding amount taxes paid i.e. our constructed variable measures net exogenous incomes.

As mentioned earlier our aim is to study whether the choice of how differently constructed budget sets affect the estimation results. For this reason we have estimated exactly the same labour supply specifications using the same data in both models. We have chosen the semi–logarithmic labour supply function with a measurement/optimisation approach to our representative model. This is due to the fact that we have previously estimated labour supply functions using the same data as here with different functional specifications and with different unobserved heterogeneity assumptions and our chosen specification has turned out to be the most robust one.

Table 2. -About Here-

As can be seen from Table 2, results between two estimated models are almost identical. Estimates for the net wage term satisfy theoretical expectations and it is also precisely estimated in both cases. The exogenous income variable has a negative sign in both models and it is statistically insignificant. Indicator variables indicate that the age of the youngest child in home affects the desired labour supply. The presence of 0–3 year old children reduces the desired labour supply and the effect is significant in both cases. If the youngest child is older than ten years, then there is a tendency to want more work (measured in hours). Age increases the labour supply up to a saturation point and the number of children in the household reduces the desired labour supply.

Results are consistent between both models. Both models above show a negligible income effect[12] but a reasonably large uncompensated wage effect and thus also the level of compensated wage elasticities are reasonably large. It is also worth mentioning that no violation of the Slutsky condition was found.

In the piece–wise linear model compensated wage elasticity is 0.21 and in the differentiable budget constraint model it is 0.215. These are smaller than the result got by Ilmakunnas (1992) for Finnish married females using 1987 data. She estimated labour supply function only for participants using linear labour supply function and her estimate for the compensated wage elasticity was 0.29. Kuismanen (1995) estimated labour supply function for married males also using only participants and linear labour supply function and his estimate for the compensated labour supply was 0.08. In case of Sweden, Blomquist and Hansson–Brusewitz (1990) obtain wage elasticities for females that vary from 0.34 to 0.75 using roughly the similar kind of statistical approach.

It is interesting to comment on how compensated elastisities vary across different demographic groups. We use a piece–wise linear model as our reference model. When we divide our sample by age we see that the age group 25–35 have highest elasticity, 0.26 per cent and it monotonically decreases with age and it is only 0.14 per cent among the individuals above 55 years. Elasticities also vary between industries. Using Statistics Finland's classification we divided individuals into eight categories. From these categories individuals who worked in private sector services or in education and research have the highest mean elasticities, 0.25 and 0.27 respectively. All the other groups' (see appendix for definitions) elasticities are near the overall mean except in the case of manufacturing workers, whose elasticity is 0.17. It is a bit surprising that the number of children do not have a great effect on elasticities. Mean elasticity for females with no children is 0.20 and it increases monotonically with the number of children.

What is the evidence from the other two similar type of studies? In the MaCurdy et al.(1990) paper they found that in the case of the piece–wise linear approach, estimates implied larger responses than differentiable approach. They used U.S. data from 1976 (PSID; 1017 prime–aged males). This data set is almost the same as in Hausman's (1980) seminal paper but the findings differ. Even in the piece–wise linear case MaCurdy et al. find much more modest labour supply responses than in Hausman.

The other similar study by MaCurdy and Flood (1992) is more likely to ser as a better comparison to our paper due to the similar tax and other institutional systems in Sweden. Their analysis deals with male labour supply and the data set is drawn from the Swedish Household Market and Non–market Activities Survey (HUS) for the year 1984. Results indicated that differentiable and piece–wise linear approaches produced identical results. Authors concluded that results might depend on the degree of progressivity in a way that in the case of high degree of progressivity methods are likely to produce similar results and when the tax system consists of only few tax brackets, then a differentiable approach might function better. Their results are very similar to results by Blomquist (1983) and Blomquist and Hansson–Brusewitz (1990) who also used Swedish data.

It is clear that our results are not a proof that these methods work similarly in the presence of high degree of progressivity, but together with the above mentioned studies it gives a strong indication that this might be the case. If we go back to Section 2 it can be seen that introducing a very high degree of progressivity to the piece–wise linear case, then the shape of budget constraint actually becomes more and more like differentiable constraint.

5 Conclusions

Sophisticated Maximum-likelihood approaches have become standard tools for analysing the labour supply disincentive effects of income tax systems. On the other hand, these ML approaches have generated a discussion among researcher of these method's robustness. To analyse the Finnish tax system we have estimated labour supply function with nonlinear income taxation using two different ML approaches. First we consider the so–called piece–wise linear approach, which has been the most popular procedure in recent years. Our second approach is the so–called differentiable function method, which approximates the actual piece–wise linear function as closely as possible using polynomial function.

One advantage of the piece–wise linear approach is that it allows us to carefully model all the instutional chracteristics which affect the shape of the budget sets. Obviously, this approach is a data intensive one. It has been argued that individuals do not know the exact shape of their budget constraints and thus the differentiable approach could be a suitable substitute.

Our results indicate that proxying tax schedules by smooth continuous functions produces similar results with the piece–wise linear approach when the degree of progressivity is high. For example, the mean compensated wage elasticities are almost the same in both models (approximately 0.21) and also other covariates behaves similarily. Our tentavive conclusion is that because of the high degree of progressivity these two approaches actually generate very similar budget constraints and because our functional specification and data set is exactly the same one in both approaches, then it is no surprise that similar results appear. For example, MaCurdy and Flood (1992) found that when using Swedish data, these two methods produced almost identical results. Their conclusion was also that this result might depend on the degree of progressivity. This finding is not necessary valid when the tax system includes only few tax brackets. Results by MaCurdy at al. (1990) show that in the case of the U.S. data, the piece–wise linear approach yields larger labour supply responses than the piece–wise linear approach.

Acknowledgements
I am especially indebted to Richard Blundell, Richard Dickens, Ian Walker and Ilpo Suoniemi for very helpful and detailed comments and suggestions. Also discussions with Heikki A. Loikkanen, Matti Tuomala and Matti Viren have been very useful. Kathleen McAleenan deserves many thanks for checking the language. The usual disclaimer applies. Financial aid from the Yrjo Jahnsson Foundation is gratefully acknowledged.

References

Blomquist, S. (1983) The Effect of Income Taxation on the Labour Supply of

Married Men in Sweden, *Journal of Public Economics* **22**, 169–197.

Blomquist, S. (1995) Restrictions in Labour Supply Estimation: Is The MaCurdy Critique Correct?, *Economic Letters* **47**, 229–235.

Blomquist, S. (1996) Estimation Methods for Male Labour Supply Functions: How to Take Account of Nonlinear Taxes, *Journal of Econometrics* **70**, 383–405.

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

Blundell, R., Duncan, A. and Meghir, C. (1998) Estimating Labour Supply Responses Using Tax Policy Reforms, *Econometrica* **66**, 827-861.

Blundell, R. and MaCurdy, T. (1999) Labour Supply: A Review of Alternative Approaches, *Handbook of Labour Economics.* Edited by O. Ashenfelter and D. Card. North-Holland.

Dudewicz, E. and Mishra, S. (1988) *Modern Mathematical Statistics.* John Wiley and Sons.

Flood, L. and MaCurdy, T. (1992) Work Disincentive Effects of Taxes: an Empirical Analysis of Swedish Men, *Carnegie–Rochester Conference Series on Public Policy*, **37**, 239–278.

Hall, R. (1973) *Wages, Income and Hours of Work in the U.S. Labour Force.* Markham Chicago.

Hausman, J.(1980) The Effect of Wages, Taxes and Fixed Costs on Women's Labour Force Participation, *Journal of Public Economics*, **14**, 161–194.

Hausman, J. (1985) The Econometrics of Nonlinear Budget Sets, *Econometrica*, **53**, 1255–1282.

Heckman, J. (1979) Sample Selection Bias as a Spesification Error, *Econometrica*, **42**, 153–162.

Heckman, J. (1983) Comment, *Behavioural Simulation Methods in Tax Policy Analysis*, Edited by M. Feldstein. University of Chicago Press.

Ilmakunnas, S. (1992) Income Taxation, Hours Restriction and Labour Suply, Working Paper 113, Labour Institute for Economic Research, Helsinki.

Kuismanen, M. (2002) Empirical Essays on Non–Linear Income Taxation and Labour Supply, Unpublished Ph.D Thesis. University College London.

MaCurdy, T., Green, D. and Paarch, H. (1990) Assessing Empirical Approaches for Analyzing Taxes and Labour Supply, *Journal of Human Resources*, **25**, 415–490.

Moffit, R. (1986) The Econometrics of Piecewise–Linear Budget Constraints, *Journal of Business and Economic Statistics*, **4**, 317–328.

Pudney, S. (1989) *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*, Basil Blackwell, Oxford.

Footnotes

[1]Pudney (1989) shows how difficult it is to actually construct the budget sets accurately
[2]As far as we know, their method has not been used in empirical studies.
[3]Note that one can always derive indirect and direct utility functions from the labour supply function or vice versa.
[4]For the completeness we can show that optimum can be found from the zero hours $h = 0$ if

$$h^*(w_1, y_1, z; \alpha, \beta, \gamma) \leq 0$$

or correspondingly from the maximum hours $h = H_n$, if

$$h^*(w_n, y_n, z; \alpha, \beta, \gamma) \geq H_n.$$

[5]It is important to realise that in the literature measurement error is interpreted in two different ways. The older interpretation is that the *positive* observed hours is measured with error. In this case one must choose the density function which ensures that reported hours of work are always positive with a feasible ε. The second interpretation is the optimisation error which reflects to the degree to which individuals' actual hours of work deviate from their desired hours. Thus, it is possible to observe that some individuals are not working even their desired hours are strictly positive because a realisation of ε causes measured hours to be non-positive. Most studies made are consistent with this latter interpretation.
[6] We do not derive here the likelihood function in the case of unobserved heterogeneity due to space limitations. It is available from the author upon request. Similar technique is also used in Blomquist (1983) and Pudney (1989).
[7] We experimented using other combinations of (t_i, I_i), but the above specification produced the best approximation.
[8] As an example $\int \Phi dI = I\Phi + \phi$ etc. For the technique see Dudewicz and Mishra (1988).
[9]LFS data set also includes some information about individuals' financial situation.
[10]Obviously this is not to say that the differentiable method is the right one when estimating labour supply functions in the presence of non–linear income taxation. In the differentiable budget constraint approach we need fewer restrictive assumptions in the background than in the piece–wise linear approach and thus we can "test" if these assumptions are important or not.
[11]Results are available from the author upon request.
[12]All elasticities are calculated using mean values.

Appendix 1: Definitions of the variables

union=1, if the respondent is a member of a union
age=Age of the respondent
age2= Age squared
educ10=1, if the respondent has 10 years of education. Otherwise zero.
educ12=1, if the respondent has 11-12 years of education. Otherwise zero.
educ14=1, if the respondent has 13–14 years of education. Otherwise zero.
educ15=1, if the respondent has 15+ years of education. Otherwise zero.
ueduc=1, if the respondent has a university degree in the following fields: Technology, business, law, natural science and social sciences
nchild=Number of dependent children.
cdum1,...,cdum4= Dummy variables for the youngest child. Age groups are 0–3,4–6,7–9 and 10+.
schild=Number of children aged 0–3.
cchild=Number of children aged 4–6.
bchild=Number of children aged 7–9.
exp= Working experience
exp2= Experience squared
tenure= Duration of the current job
tenure2= Square of tenure
pjob=1, if respondent has a permanent job
phusb=1, if respondent's husband is working
stat=1, if the respondent is a white–collar worker and 0 if a blue–collar worker.
socio=1, if the respondent is a upper white–collar worker. Otherwise zero
hwage= Hourly wage rate.
shwage= Subjective hourly wage rate.
exo= Unearned income.
exo+hnet= Unearned income plus husband's net incomes.
south=South Finland.
west=West Finland.
east=East Finland.
middle=Middle Finland.
north=North Finland.
lapl=Lapland.

Appendix 2: Descriptive Statistics

Descriptive statistics: participants and Non–paricipants

Variables	*Participants*	*non–participants*
Hours	1855.58(560.10)	
union	0.71(0.45)	0.21(0.41)
age	41.29(8.47)	43.13(11.61)
educ10	0.30(0.46)	0.26(0.41)
educ12	0.18(0.39)	0.16(0.36)
educ14	0.05(0.21)	0.06(0.24)
educ15	0.07(0.26)	0.06(0.24)
cdum1	0.13(0.34)	0.32(0.46)
cdum2	0.12(0.33)	0.06(0.24)
cdum3	0.12(0.32)	0.05(0.21)
cdum4	0.25(0.44)	0.10(0.30)
workexp	19.50(9.32)	17.30(11.60)
jobdur	8.60(8.34)	
permjob	0.77(0.41)	
phusb	0.86(0.33)	0.68(0.46)
hwage	48.81(23.99)	
shwage	44.28(19.35)	
exo	5525.66(14028.9)	6935.43(11183.8)
exo+hnet	84590.67(61197.5)	77666.32(46447.6)
south	0.25(0.44)	0.21(0.41)
west	0.16(0.36)	0.13(0.33)
east	0.19(0.39)	0.24(0.43)
middle	0.14(0.34)	0.15(0.36)
north	0.18(0.39)	0.18(0.38)
lapl	0.08(0.27)	0.08(0.27)

For definitions of the variables see Appendix 1.

Table 1. Wage Equation

Wage Equation. Dependent variable: ln hwage.		
Variables	*Coefficient*	*Standard Error*
Constant	2.83342	0.2511
Age	0.01753	0.0135
Age2	-0.00017	0.0001
Educ10	0.06881	0.0024
Educ12	0.19534	0.0297
Educ14	0.27270	0.0469
Educ15	0.51690	0.0469
Exp	0.01659	0.0053
Exp2	-0.00027	0.0001
Tenure	0.02410	0.0038
Tenure2	-0.00045	0.0001
Pjob	0.04720	0.0299
Husb	0.00760	0.0290
Stat	0.10338	0.0241
Socio	0.23919	0.0366
Nchild	-0.03065	0.0104
South	0.15898	0.0222
Exo+hnet	3.95e-07	1.67e-07
Occ. dummies	Yes	
Ln L	-1221.91	

NOTE: The selection index is a function of the individual, geographical and demand side variables. The selectivity effect was statistically significant. Reference group for occupation is manufacturing workers.

For definitions of the variables see Appendix 1.

Table 2. Results for the labour supply functions

Maximum Likelihood Estimates Asymptotic Standard Errors in Parenthesis		
Variables	*piece–wise linear budget constraint*	*Differentiable budget constraint*
Constant	-2.57905	-2.70210
	(0.55650)	(0.55474)
Ln W	0.37046	0.39121
	(0.12135)	(0.12444)
Exog. inc	-0.00045	-0.00045
	(0.00022)	(0.00020)
Cdum1	-0.33917	-0.33823
	(0.09948)	(0.0990)
Cdum2	-0.00487	-0.00482
	(0.10437)	(0.10431)
Cdum3	0.09616	0.09601
	(0.10050)	(0.10049)
Cdum4	0.14310	0.14310
	(0.07690)	(0.07688)
Age	0.16118	0.16144
	(0.02484)	(0.02488)
Age*Age	-0.00227	-0.00229
	(0.00028)	(0.00028)
Sosio	0.19945	0.19912
	(0.09521)	(0.09567)
Nkids	-0.08419	-0.09011
	(0.03235)	(0.03239)
σ_ϵ^2	0.98208	0.96998
	(0.01907)	(0.01918)
Ln L	-2669.61	-2675.56

Note: In both of the above models, the dependent variable (yearly hours) is divided by 1000. The exogenous income variable contains only person's own exogenous income components (net) and it is divided by 100. For definitions of variables see Appendix 1.

Piece–Wise or Differentiable Budget Constraint? Estimating Labour Supply Function for Finnish Females

Mika Kuismanen
European Central Bank
Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany.
Tel: 49-69-1344 7607
Fax: 49-69-1344 6575
Email: mika.kuismanen@ecb.int

September 2006

Abstract

Various estimation approaches have been used in recent literature to study the effect of non-linear income taxation on labour supply. Different techniques and data sets have produced a wide range of income and substitution elasticities. In this study we utilise register data provided by the tax authorities. This gives us good possibilities to construct detailed budget constraints for each individual in our sample. We estimate labour supply function using the piece–wise linear budget constraint approach and the differentiable budget constraint approach suggested by MaCurdy et al. (1990) Our results support the view that if one is able to mimic the actual budget set closely and if the degree of progression is high then these two methods are likely to produce similar results. Some sensitivity analysis is also carried out using alternative asumptions concerning the budget sets.

1 Introduction

The purpose of this paper is to provide results conserning labour supply behaviour in Finland. It is known that the economic theory does not give much prediction power on how taxation affects labour supply because it leaves the signs of substitution and income elasticities open, see *e.g.* Blundell and MaCurdy (1999). In addition, introducing different kind of welfare systems may create non-convex budget sets, and thus, certain areas of the budget constraint cannot correspond the utility-maximising points. In many cases we do not even know if an increase in marginal tax rate will increase or decrease supplied hours. So, empirical research is needed to give us information concerning tax and benefit systems' behavioural effects.

Our paper compares two different empirical approaches to model labour supply behaviour when progressive income taxation is taken into account. We start with a conventional piece-wise linear method. In this approach we try to mimic the actual budget constraint as well as possible and then it is fully taken into account in the estimation procedure. The above mentioned approach means that the likelihood function takes into account the choice of hours over the entire exogenous tax schedule, removing the endogeneity problem which is present in simpler approaches, like in the linearised budget constraint approach. Piece–wise linear method has been criticised for various reasons (see Heckman (1983)). First, mimicing the budget constraints requires a lot from the data. Secondly, it is questionable if an econometrician can measure the constraint accurately and if individuals really know the actual shape of it.

To avoid the above mentioned problems but still allow a convex shape of the budget constraint we follow MaCurdy et al.(1990) to construct a differentiable budget constraint. The central idea is to approximate the actual piece–wise linear constraint by using continuous smooth polynomial function because it is unlikely that an individual knows exact shape of her budget set. Technically, this approach means mimicing the tax schedule by fitting a polynomial function to the marginal tax rates. After integrating this function we get a differentiable relation which approximates the amount of total taxes. This method is much easier to estimate since a purely continuous distribution describes the hours of work decision.

Finnish Labour Force Survey (LFS) data for the year 1989 is used in the estimation. The sample consists of 2037 married females aged 25 to 60. Income data for these females and their partners is drawn from the Tax Register Data and then merged with the LFS. This gives us a rich data set to build budget constraints. Subsection 4.1 includes more thorough discussion on the data. As non–linearities mainly arise from the income tax system a short description of it is given in the end of section two. An example of a labour supply study using the longitudinal data can be found from Blundell, Duncan and Meghir (1998) and an example of a study using time series data in Fraser and Paton

(2003).

What we find is that a differentiable budget constraint is able to approximate the piece–wise linear one quite closely and that the estimation results do not differ significantly between these two approaches. The result is in line with the study by MaCurdy and Flood (1992) but conflicts with the original MaCurdy et al. (1990) study. Our conclusion is that in the case of a high degree of progressivity these methods are likely to produce similar outcomes because then the polynomial function is simply capable of producing almost identical constraints compared to the piece–wise linear one. We also provide some empirical evidence that results might be sensitive to the way how budget sets are calculated.

The set up of this paper is as follows. In section 2 we shortly present the basic ideas behind the piece–wise linear budget constraints. After that we go through two estimation approaches for the labour supply function in the presence of non–linear budget constraint. Section 4 presents estimation results and section 5 concludes the paper.

2 Budget sets

Let us start by analysing a simplified income tax system with progressive elements. The income tax system consists of three tax brackets and thus three marginal tax rates; $t_1, t_2, and\, t_3$. Inside the tax brackets marginal tax rate is constant but it increases with income. The outcome is the familiar piece-wise linear budget constraint with four kink points.

Marginal tax rate t_1 leads to the net wage $w_1 = (1 - t_1)w$ and this corresponds to the first segment in tax schedule. Correspondingly the net wage rate in the second segment is $w_2 = (1 - t_2)w$ etc. $H_0, H_1 and\, H_2$ are kink points where the marginal tax rate changes and H_n stands for the upper limit of labour supply. y_1 is the exogenous income component and thus does not depend on hours supplied. Note that this component is directly observed from the data. y_2 and y_3 are called "virtual income" terms and must be calculated recursively in the following way

$$y_i = y_{i-1} + (w_{i-1} - w_i)H_{i-1}$$

Thus geometrically we just extend a given budget segment to the vertical axis. The crucial thing to realise is that these components cannot be observed directly from the data. Given this budget constraint the consumer makes his or her labour supply decision and the optimal labour supply can vary from zero hours to the maximum number of hours.

For simplicity we assume that non-labour income is not taxed and E (can also vary between individuals) stands for the tax deductions. In this case the taxable income I_i is earned income (wh) less the deductions and thus the kink points can be calculated as

$$H_i = (I_i + E)/w$$

When the budget set is convex and consumer preferences can be represented by the quasi-concave utility function, $u(c, h)$, which is non-decreasing with respect to consumption and non-increasing with respect to supplied hours, h.

But what are the pros and cons of this approach? First of all, this method allows us to construct[1] budget sets which recognise all the institutional features of the tax and social security system. The piece–wise linear method also treats the marginal tax rate endogenously in the estimation procedure. It also allows us to incorporate different stochastic assumptions, like allowing randomness in hours to arise from measurement (optimisation) and/or unobserved individual preferences. We can introduce fixed costs of working into the model quite easily and the treatment of unobserved wages can be done in different ways. MaCurdy et al.(1990) includes an excellent theoretical discussion about how unobserved wages should be calculated[2].

The fundamental assumption behind this model is that the observed market behaviour is the outcome of free rational choice subject to piece–wise linear budget constraint. In other words, we assume that on individual has perfect knowledge of the budget set and an econometrician is able to measure all the budget set variables without error. It is hard to imagine that the above mentioned criteria are actually met either by the econometrician or the decision making individual.

Probably the most serious claim against piece–wise linear modelling is the so–called MaCurdy-critique. MaCurdy et. al. claimed "that the this method requires the satisfaction of parametric restrictions that constraint the signs of estimated substitution and income effects" In his articles Blomquist (1995) (1996) has written that MaCurdy's claim is not generally true.

Is there then a way to overcome above mentioned problems? In their influential paper, MaCurdy et al. developed a new approach which utilises a differentiable function to approximate marginal tax function. The idea sounds complicated, but as we will show below, it turns out to be a simple and attractive alternative.

We can think that individuals do not know the exact shape of their budget set, but they do have an idea of its approximate shape. For example, most likely individuals do not exactly know the number of supplied hours when moving from one tax brackets to another or what their accepted tax deductions might be etc..

A differentiable budget constraint approach is also easier to apply in practise because purely continuous distribution describes the supplied hours for all individuals whose hours are strictly positive. If we assume that the constraint is convex and consumers' preferences are strictly quasi-concave then we can be sure that we can find a tangent point and this point is going to be a unique one. In addition, because we have "smoothed out" all the segments and kinks we are left with one continuous non-linear segment, we do not need to write the tax algorithm into the likelihood function as can be seen in the next section.

Finnish income tax system in 1989 basically had two parts. A progressive state income tax and a proportional local (municipal) income tax. In addition, individuals contribute to the National Pension Insurance scheme (1.55 percent from the taxable income) and National Health Insurance scheme (1.25 percent from the taxable income), which are proportional to income changes. Roughly speaking, the tax liability in state tax and municipal tax is the same excluding the tax deduction system. In 1989 the state income tax schedule was composed of seven marginal tax rates varying from 0 to 44 percent (thus having 7 tax brackets *i.e.* piece–wise linear segments) and the municipality tax rate varied from 14 percent to 19.5 percent. We also have developed a formula to calculate state tax deduction for all persons in the sample. Estimated tax deductions varied from 0 FIM to 29 500 FIM (approx. 5000 EUR). See also subsection 3.2 for an example.

3 Two estimation approaches for the labour supply function in the presence of non–linear budget constraint

In the case of linear income tax system (or when income tax is ignored) estimation of labour supply function is straightforward. This is because the budget constraint individual faces is linear and there can only be one utility maximisation point (*i.e.* observed and optimal hours lie on the same linear segment). This changes when non–linearity is present (progressive income taxation). For example, when the consumer faces piece–wise linear budget constraint her observed and optimal labour supply may lie on different segments. This possibility has to be taken into account in estimation. We start by showing how this can be tackled by construction of tax algorithm which is then substituted to the likelihood function in final estimation stage. In subsection 3.2 below we introduce an alternative method where tax algorithm is not needed even when we allow non–linear budget constraints.

3.1 Piece–wise linear approach

In the case of piece–wise linear budget constraint economic theory predicts bunching of observations (hours of work) at kink points (where marginal tax rate changes) or just below of them. Reason is that if individuals increase their hours they will move to an upper tax bracket (or, for example, in social security system they move to the point where the credit is taxed away) where they will face higher marginal tax rate. Blundell, Duncan and Meghir (1998) provides an example where bunching was found. In Kuismanen (2004) it is shown that in Finnish data bunching was not an issue. Moffitt (1986) states that if observations are distributed evenly across the budget sets it provides a reason to introduce measurement error term into the model. Naturally, there are also other reasons to introduce measurement error, such as reporting errors in measured hours.

We start with a basic measurement error approach where the general labour supply function can be written as $h_i = h_i^*(w_i, y_i, z_i; \alpha, \beta, \gamma) + \varepsilon_i$. α, β and γ are parameters to be estimated. Vector z includes individual characteristics (*e.g.* socio-economical and demographic variables) and the variables w and y represent the marginal wage and virtual (non–labour) income variables correspondingly. ε represents the measurement/optimisation error. Subindex i denotes the individual.

In the statistical model we have to calculate the densities of h_i and this naturally requires evaluation of the maximum utilities received on each linear segment of the budget constraint. More formally, we now write the problem as

$$
\begin{aligned}
f(h_i) &= P[h_i = 0] + P[h_i > 0] * f(h_i \mid h_i > 0) + P[h_i = H_n] \\
&= P[\ \text{at zero}\ ] \\
&+ P[\ \text{below kink 1}\ ] * f(h_i \mid\ \text{below kink 1}\ ) \\
&+ P[\ \text{at kink 1}\ ] \\
&+ P[\ \text{above kink 1}\ ] * f(h_i \mid\ \text{above kink 1}\ ) \\
&. \\
&. \\
&. \\
&+ P[\ \text{at maximum}\ ]
\end{aligned}
$$

So, we can think that **observed** hours are generated by the following generalised Tobit–model (note that we have dropped the subscript i)

$h = 0$ if $h^* + \varepsilon = 0$
$h = h^* + \varepsilon$ if $0 < h^* + \varepsilon < H_n$

$h = H_n \qquad\qquad \text{if } h^* + \varepsilon \geq H_n$

and the corresponding Likelihood Function can now be written as

$$(1) \quad L = \prod_{i \in A} \left[1 - \Phi\left(\frac{h^*}{\sigma_\varepsilon}\right)\right] \prod_{i \in B} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{h_i - h^*}{\sigma_\varepsilon}\right)\right] \prod_{i \in C} \left[1 - \Phi\left(\frac{H_n - h^*}{\sigma_\varepsilon}\right)\right].$$

Where,

A is index set when $h = 0$
B is index set when $0 < h < H_n$
C is index set when $h \geq H_n$.
$\phi(\cdot)$ is Standardised Normal Density Function and $\Phi(\cdot)$ is Cumulative Normal.

The first part of the likelihood function correspond individuals whose observed hours are zero. The second part corresponds those individuals whose observed hours are in some of the segments or kink points and the third part corresponds those whose observed hours are at maximum.

At this stage we need to show how to determine the optimal supply of hours in the presence of a kinked (convex) budget constraint (i.e. to build a search algorithm to operate inside the second component of the above presented Likelihood function).

The optimal supply of hours h^* can be found from the segment k ($k = 1, \ldots, n$), if

$$H_{k-1} < h^*(w_k, y_k, z; \alpha, \beta, \gamma) < H_k$$

Intuition behind this calculation rule is the following: after calculating the slope of the indifference curve from the direct utility function[3] we replace consumption c ($= w_k h + y_k$) by individual's income (calculated for all the segments) and then equate the slope of the indifference curve and the marginal wage w_k corresponding that segment. The algorithm iterate as long as this condition is satisfied. If in some cases we cannot find the solution we start to look it from the kink points.

Optimum h^* is found from the kink point H_k ($k = 1, \ldots, n-1$), if

$$h^*(w_k, y_k, z; \alpha, \beta, \gamma) \geq H_k \qquad \textit{and} \qquad h^*(w_{k+1}, y_{k+1}, z; \alpha, \beta, \gamma) \leq H_k.$$

Another way to express above condition is that the optimum can be found from the H_k when the slope of the indifference curve is bigger or equal than w_{k+1} and the slope is smaller or equal than w_k.[4]

The above formulation shows how optimal hours can be calculated under the progressive income taxation when the following aspects are known: tax schedule, hourly wage, exogenous (non–labour) income and the shape of the labour supply function (or correspondingly the form of the direct utility function).

Despite h^* can be calculated quite easily in the convex case the maximisation of the (log) likelihood function is not straightforward because h^* is not always well–behaving function respect to the parameters. First, the log–likelihood function is not differentiable everywhere (kink points) and secondly there can be parameter values where the function becomes flat. This can become a serious problem if there are not enough variation between the budget sets.

In the above discussion we did not make any specific assumptions about the stochastic specification *i.e.* we assumed that all variance in hours conditional on covariates is measurement/optimising error.[5] This means that preferences are assumed to be non–stochastic *i.e.* all variation in preferences is only due to observable personal characters (i.e. regressors). This is the usual procedure adapted by researchers, at least when using cross–section data sets. In the context of labour supply we might think that there exists different sources for stochastic disturbances to arise from. First, the usual measurement error interpretation implies that the parameters to be estimated are the same for all individuals, so there is only one utility maximising choice in the population (this is probably a questionable outcome). Second, there might exist randomness in preferences which is not captured by the variables we include in our regression function. As Moffitt (1986) argues it is reasonable to expect that at least some amount of the observed distribution of observations over the budget constraint is a result of heterogeneous preferences. It is natural to think that both aspects are relevant in the context of labour supply. Look Hausman (1985) for the other possible sources of stochastic disturbances. For non-parametric estimation of labour supply responses in the case of piece–wise linear budget constraint see Blomquist and Newey (2002).

The two different stochastic elements mentioned above have different implications for the data. In the standard measurement/optimisation error approach the observations should be distributed evenly over the whole budget constraint, as in the case of heterogeneity of preferences we should find clusters of observations at the kink points (in the case of convex budget constraint). In theory, the model which only includes heterogeneity term is possible, but not a very appropriate one for the most applications because it is unlikely that all observations are clustered to the kinks. Although empirical evidence shows some clustering (especially in the cases of big changes in marginal tax rates) it is not usually strong enough to leave measurement error term without modelling. One relatively easy way to proceed is to estimate the model with random preference term and then test if its variance is different from zero. One further motivation for including the heterogeneity term is that in the cross–section studies one finds a large amount of unexplained variance.

Pudney (1989) provides an excellent introduction how random preference model can be operationalised. In our empirical part we found that the unobserved heterogeneity component was statistically insignificant and that also the rest of the covariates were in line with the model when only measurement error term was included.

3.2 Differentiable Budget Constraint Approach

Differentiable budget constraint technique was first introduced by MaCurdy et al. (1990) and our presentation will follow it with suitable modifications to take account the Finnish tax system. Intuition behind this method is to approximate the tax schedule by fitting a function to the marginal tax rate. After integrating marginal tax function we get a differentiable relation approximating the amount of total taxes paid as a function of taxable income.

Let us introduce some new notation. Denote $I(h)$ for taxable income at h hours of work and $M[I(h)]$ for marginal tax rate function. Now, for example the simplified three bracket income tax system can be presented in the following way:

$$
\begin{aligned}
M[I(h)] &= t_1 \,[\text{from } I(H_0)\, to\, I(h_1)] \\
&= t_2 \,[\text{from } I(H_1)\, to\, I(h_2)] \\
&= t_3 \,[\text{above } I(H_2)]
\end{aligned}
$$

where H_i denotes the kink points where marginal tax rates t_i changes. To approximate the marginal tax rate schedule the function must fit a step function presented above closely and it still should be differentiable at the step points (i.e. kink points). MaCurdy et al. suggested a following kind of approximation

$$
(2) \qquad \overline{M[I(h)]} = \sum_{i=1}^{K} [\Phi_i(I(h)) - \Phi_{i+1}(I(h))] * p_i(I(h)),
$$

where $\Phi_i(I(h))$ denote the cumulative normal distribution function evaluated at the income level I(h) with mean μ_i and variance σ_i^2. The idea is that the difference $\Phi_i(I(h)) - \Phi_{i+1}(I(h))$ takes value of one over the range where t_i is relevant and zero elsewhere. Now, we can control this by adjusting the mean and the variance. Adjusting the mean we can control the moment when the value of one begins and ends, and adjusting the variances we can control how quickly this happens. The trade–off here is the smoothness of transition against the precision. $p_i(I(h))$ are the polynomials in income. For example, in Finnish case in 1989 there were 7 tax brackets, so we can set $K = 7$ and p_i is the marginal tax rate t_i associated with the ith tax bracket.

To see how the above presented generalisation works let us go back to our simplified three bracket tax schedule discussed above. In this case we have three marginal tax rates $t_i < t_2 < t_3$, so we have "three segments to smooth out". We can now write our approximation function using the above notation for this problem as

$$\overline{M[I(h)]} = [\Phi_1(I(h)) - \Phi_2(I(h))] * t_1 + [\Phi_2(I(h)) - \Phi_3(I(h))] * t_2 + \Phi_3(I(h)) * t_3$$

So, the first segment has a height t_1(can be thought as a flat line with a height t_1.) and thus corresponding taxable income is from $I(H_0)$ to $I(H_1)$. This feature is captured by parameterising $\Phi_1(I(h))$ with mean $\mu_1 = I(H_0)$ and correspondingly $\Phi_2(I(h))$ with mean $\mu_2 = I(H_1)$. The first distribution function $I(H_0)$ takes value of one above the income level $I(H_0)$ and zero elsewhere and the second distribution function $I(H_1)$ takes value of zero below the income level $I(H_1)$ and then switches to one above it. So the difference of these functions is one between $I(H_0)$ and $I(H_1)$ and zero elsewhere and correspondingly for all other ranges. So, we can control the switch from zero to one(and vice versa) by adjusting the means. How quickly these switches will take place depends on the values given to the variances.

In 1989 Finnish marginal tax rates varied from zero to 44 percent including 7 tax brackets and the tax exemption level was 36 000 FIM. To put the income tax system into the described framework concerning differentiable budget constraints we get the following three parts. Part one is valid from zero income up to 36 000 FIM (approx. 6030 EUR) and the tax rate for this range is equal to the individuals local tax rate t_l. In the second part income ranges from 36 000 FIM up to 250 000 FIM (approx. 42 050 EUR) where the tax rate is the local tax rate plus the monotonically increasing marginal tax rate $t_i, i = 1, ..., 7$. The third part is for the incomes over 250 000 FIM. In this case the tax rate is the local tax rate plus the federal tax rate of 44 percent.

So, according to given information we can write the approximation for the marginal tax rate function in the Finnish case as:

$$\begin{aligned}\overline{M[I(h)]} &= \left[\Phi_1\left(\frac{(I(h)) - \mu_1}{\sigma_1}\right) - \Phi_2\left(\frac{(I(h)) - \mu_2}{\sigma_2}\right)\right] * t_l \\ &+ \left[\Phi_2\left(\frac{(I(h)) - \mu_2}{\sigma_2}\right) - \Phi_3\left(\frac{(I(h)) - \mu_3}{\sigma_3}\right)\right] * F(I)\end{aligned}$$

$$(3) \qquad + \quad \Phi_3\Big[\Big(\frac{(I(h)) - \mu_3}{\sigma_3}\Big)\Big] * (t_l + 0.44),$$

where $F(I)$ is a polynomial in taxable income which approximate the increasing tax rates from 36000 FIM to 250000 FIM.

Approximation of the $F(I)$ is done by running the following Ordinary Least Squares regression

$$(4) \qquad t_i = \xi_0 + \xi_1 * I_i + \xi_2 * I_i^2 + \xi_3 * I_i^3,$$

where t_i is the marginal *federal* tax rate. What we do is that we create an variable which increases by 100 FIM starting at 36 000 FIM and ending at 250 000 FIM. In other words we create 2140 equally spaced combinations (t_i, I_i). After estimation, we can use the estimates from this third-degree polynomial approximation to specify the above equation to be

$$F(I) = -1219 + t_l + 7.46 * 10^{-6} * I - 3.80 * 10^{-11} * I^2 + 6.46 * 10^{-17} * I^3,$$

where t_l is the local tax rate. Our model explains 95 percent of the variance in marginal tax rates.[6] Plugging the above estimated formula into the marginal tax rate function and integrating it with respect to the income we can derive formula which approximates the amount of total taxes paid.

$$\begin{aligned}
(5)\ T(I(h)) &= \int \overline{M[I(h)]dI} \\
&= \Big[t_l\Phi_1\Big(\frac{(I(h)) - \mu_1}{\sigma_1}\Big) - t_l\Phi_2\Big(\frac{(I(h)) - \mu_2}{\sigma_2}\Big)\Big] \\
&+ \Big[\xi_0 \int \Phi_2\Big(\frac{(I(h)) - \mu_2}{\sigma_2}\Big)dI + \xi_1 \int \Phi_2\Big(\frac{(I(h)) - \mu_2}{\sigma_2}\Big)IdI \\
&+ \xi_2 \int \Phi_2\Big(\frac{(I(h)) - \mu_2}{\sigma_2}\Big)I^2dI + \xi_3 \int \Phi_2\Big(\frac{(I(h)) - \mu_2}{\sigma_2}\Big)I^3dI\Big] \\
&- \Big[\xi_0 \int \Phi_3\Big(\frac{(I(h)) - \mu_3}{\sigma_3}\Big)dI + \xi_1 \int \Phi_3\Big(\frac{(I(h)) - \mu_3}{\sigma_3}\Big)IdI \\
&+ \xi_2 \int \Phi_3\Big(\frac{(I(h)) - \mu_3}{\sigma_3}\Big)I^2dI + \xi_3 \int \Phi_3\Big(\frac{(I(h)) - \mu_3}{\sigma_3}\Big)I^3dI\Big] \\
&+ \Big[(t_l + 0.44) \int \Phi_3\Big[\Big(\frac{(I(h)) - \mu_3}{\sigma_3}\Big)dI\Big].
\end{aligned}$$

From this expression it is relatively straightforward to derive the final form used in estimation. We just need to find analytical solutions to it and this can be done by calculating the above integrals.[7]

This differentiable approach is more straightforward than the piece–wise linear one since a purely continuous distribution describes the hours of work and no tax algorithm is needed. Intuition behind this differentiable approach follows the idea presented by Hall (1973). We can think marginal wage rate and virtual income as a function of working hours. Hall's idea was to linearise the actual non-linear budget constraint at the observed hours. The implied slope of this linearised constraint is the marginal wage rate and the intercept of the vertical axis is the virtual income. In other words, utility maximisation implies a solution for hours of work which can be written in the form of implicit equation, $h = f[w(h), y(h)]$. By applying the Implicit Function Theorem to it we can solve this equation for h and hence derive the labour supply. Here the same idea is used but still allowing non–linearities.

4 Estimation results

4.1 Data

Our study utilises the Finnish Labour Force Survey (LFS) data. It is a cross–section data including individuals of age between 15 to 64. In the first stage the sample is drawn from the Finnish Population Census using geographical weights. After that the LFS sample is drawn randomly by age and gender. In 1989 data, the sample size is 7820 individuals. From this we selected females and were left with 4124 observations. For the empirical analysis we selected married women aged 25–60. We also deleted some groups like farmers and self–employed mainly because of different tax and social security legislation. The final sample size used in this study is 2037 observations. Examples of empirical household labour supply studies are wrote by Barmby and Smith (2001), and by Garcia and Marcuello (2002).

Income data of the corresponding individuals in the LFS is drawn from the Tax Register Data and then merged with the LFS. The income information is not based on the survey[8] data and it includes approximately 70 variables on individual's earnings. Of course, it is very unlikely that someone's earnings are composed from all these components. However, the data shows that individuals' earnings come from very different sources. Actually, for some individuals traditionally used income variables do not play any role at all. The income data also includes the same 70 variables for the spouse, so all in all, we have approximately 140 variables for married individuals to construct the budget sets. Detailed information how crucial variables like working hours, wages and exogenous incomes are calculated can be found in Kuismanen (2004).

Next, we will shortly comment the main features of the data (see also appendix). The participation rate in the selected sample is 72 percent. Unemployment rates vary geographically and the figures are the lowest in the Helsinki metropolitan area and the highest in the eastern and northern parts of the country. The average unemployment rate is 3.2 percent. The blue–collar workers are more likely to have zero hours observations than white–collar ones, and the women with two young children have the highest probability to be out of work. In our sample labour force participants are slightly younger than the non–participants and they are also better educated. It also seems that the likelihood of being a non–participant increases if the spouse is also a non–participant.

4.2 Results

Our goal in this paper is to estimate similarly specified labour supply functions using the piece–wise linear and differentiable approaches. We are interested whether these two methods produce similar results, and thus, are we able to conclude, at least using our data, that the piece–wise linear method will produce sensible results.[9] We also estimated the unobserved heterogeneity model but the term was not statistically significant and the rest of the covariates were very much in line with those reported here.[10] This finding is in line with the fact that no bunching was found. In addition, we also carry out some sensitivity analysis by comparing results from two alternative data sets for wages and income. Our baseline results are from the tax register data and as an alternative source we use wage and income information from the survey data (the LFS survey data also has information on earnings). We also estimate our labour supply with different assumptions concerning the tax deductions and municipality tax rate.

Our data set does not have direct information on individuals' hourly wages (neither in the tax register data nor in the LFS data). Thus, we have to construct this variable using the income and the hours of work variables. This means that the possible bias in the hours of work variable also shifts into the marginal hourly wage rate. For example, if worked hours are smaller than the right value then hourly wage rate becomes too high. To get rid of this bias we estimate the log–wage equation using Heckman's selection method and the predicted values are used in the final analysis as an instrument for the hourly wage rate. Results for our baseline case are presented in Table 1.

Table 1. -About Here-

Results show the familiar age–wage effect, i.e. that up to a certain age wage increases and then decreases. Education (Educ10-Educ15 are dummy variables indicating the number of completed years of education) has a positive effect on wage rate and the effect gets stronger with an increase in the number of years of education. Work experience (measured in years) increases the wage rate up to a certain experience level and then starts to decrease it. Tenure variable (number of years with the same employer) shows the similar quadratic shaped effect. Individuals who have a permanent job get a higher wage, as is to be expected. Same is also true for individuals who are white–collar workers. It is also evident that individuals living in the south of Finland earn more.

For working hours we use regular reported weekly working hours also taking into account regular hours in the second job. When calculating the exogeneous income term we took into account the following components: Interest (both taxable and nontaxable), dividend payments, sales profits, regular untaxable pensions, other regular subsidies etc. From all the components which are taxable we have subtracted the corresponding amount taxes paid i.e. our constructed variable measures net exogenous incomes.

As mentioned earlier our aim is to study whether the choice of differently constructed budget sets affect the estimation results. For this reason we have estimated exactly the same labour supply specifications using the same data in both models. We have chosen the semi–logarithmic labour supply function with a measurement/optimisation approach to our representative model. This is due to the fact that we have previously estimated labour supply functions using the same data as here with different functional specifications and with different unobserved heterogeneity assumptions and our chosen specification has turned out to be the most robust one.

Table 2. -About Here-

As presented in Table 2, results between two estimated models are almost identical. Estimates for the net wage term satisfy theoretical expectations and it is also precisely estimated in both cases. The exogenous income variable has a negative sign in both models and it is statistically insignificant. Indicator variables indicate that the age of the youngest child at home affects the desired labour supply. The presence of 0–3 year–old children reduces the desired labour supply and the effect is significant in both models. If the youngest child is older than ten years, then there is a tendency to want more work (measured in hours). Age increases the labour supply up to a saturation point and the number of children in the household reduces the desired labour supply.

Results are consistent in both models. Both models show a negligible income effect[11] but a reasonably large uncompensated wage effect and, thus, also the level of compensated wage elasticities are reasonably large. It is also worth mentioning that no violation of the Slutsky condition was found.

In the piece–wise linear model compensated wage elasticity is 0.21 and in the differentiable budget constraint model it is 0.215. These are smaller than the result got Ilmakunnas (1992) for Finnish married females using the dat for year 1987. She estimated labour supply function only for participants using linear labour supply function and her estimate for the compensated wage elasticity was 0.29. In Kuismanen (1995), we estimated labour supply function for married males also using only participants and linear labour supply function and our estimate for the compensated labour supply was 0.08. In case of Sweden, Blomquist and Hansson–Brusewitz (1990) obtain wage elasticities for females that vary from 0.34 to 0.75 using roughly the similar kind of statistical approach we have used.

It is interesting to comment on how compensated elastisities vary across different demographic groups. We use a piece–wise linear model as our reference model. When we divide our sample by age we see that the age group 25–35 has highest elasticity, 0.26 per cent and elasticity monotonically decreases with age and it is only 0.14 per cent among the individuals older than 55 years. Elasticities also vary between industries. Using Statistic Finland's classification we divided individuals into eight categories. From these categories individuals who worked in private sector services or in education and research have the highest mean elasticities, 0.25 and 0.27 respectively. All the other groups' (see appendix for definitions) elasticities are near the overall mean except in the case of manufacturing workers, whose elasticity is 0.17. It is a bit surprising that the number of children does not have a great effect on elasticities. The mean elasticity for females with no children is 0.20 and it increases monotonically with the number of children.

What is the evidence from the other two similar type of studies? In the MaCurdy et al.(1990) paper they found that in the case of the piece–wise linear approach, estimates implied larger responses than differentiable approach. They used U.S. data from 1976 (PSID; 1017 prime–aged males). This data set is almost the same as in Hausman's (1980) seminal paper but the findings differ. Even in the piece–wise linear case MaCurdy et al. find much more modest labour supply responses than Hausman.

The other similar study by MaCurdy and Flood (1992) is more likely to serve as a better comparison to our paper due to the similar tax and other institutional systems in Sweden. Their analysis deals with male labour supply and the data set is drawn from the Swedish Household Market and Non–market Activities Survey (HUS) for the year 1984. Results indicated that differentiable and piece–wise linear approaches produced identical results. Authors concluded that results might depend on the degree of progressivity in a way that in the case of high degree of progressivity methods are likely to produce similar results and when the tax system consists of only few tax brackets, then a differentiable approach might function better. Their results are very similar to results by Blomquist (1983) and Blomquist and Hansson–Brusewitz (1990) who also used Swedish data.

However, do we still get similar kinds of results after we have changed some crucial features of the budget sets, namely using subjective wage measure (wage rate is calculated from the survey data, not from the register data), not taking tax deductions into account and using constant local tax rate? Below we will only comment results from the differentiable budget constraint approach and discuss the value of compensated wage elasticity due to its importance for policy makers.

The results are quite interesting. When using predictions for subjective wage measure we get a higher compensated wage elasticity (0.277) than using wage measure from the archive (baseline case) data (0.215). This might be because there is less variation in subjective wage measure, but it is very hard to say the exact reason for this result. The results also show that, consistently, the compensated elasticity is lower (0.159) when individually calculated tax deductions are neglected. First of all, tax deductions affect the location of kink points i.e. point where marginal tax rate changes. Our approach assumes that individuals rationally choose their labour supply behaviour given the budget set they face and this is directly related to Heckman's (1983) criticisms that the approach in question to model taxes is insufficient if we are not able to measure the budget sets accurately enough. The calculation of individual tax deductions indicate that it may be a significant factor when defining individual's yearly earnings. For example, the maximum amount of deduction in our case is 5000 euros, which is a one fifth of that person's yearly incomes. It is not a very difficult task for an individual to calculate her own possible deductions, so neglecting them from our analysis we contradict our earlier made assumptions. Taking deductions into account means higher (compared to no deductions case) net yearly earnings which might also be a one possible explanation for our result.

When using a constant municipal tax rate (17%), compensated elasticity is slightly higher (0.254) than when using calculated municipal (baseline case) tax rates. Using common local tax rate to all individuals means that there is less variation in the measure of net hourly wage and thus we might violate the assumptions made about individuals' behaviour.

In his well cited paper Mroz (1987) got interesting results using different economical and statistical assumptions when estimating the labour supply function for for US females. One of his results, broadly speaking, was that Tobit–specification exaggerate both wage and income elasticities. Our results show that this is also true in our case. When estimating the model only for workers we get much lower compensated wage elasticity (0.1)

5 Conclusions

Sophisticated ML–approaches have become standard tools for analysing the labour supply disincentive effects of income tax systems. On the other hand, ML–approaches have generated discussion among researchers about the robustness of these methods. To analyse the Finnish tax system we have estimated labour supply function with nonlinear income taxation using two different ML–approaches. First we consider the so–called piece–wise linear approach, which has been the most popular procedure in recent years. Our second approach is the so–called differentiable function method, which approximates the actual piece–wise linear function as closely as possible using polynomial function.

One advantage of the piece–wise linear approach is that it allows us to carefully model all the chracteristics which affect the shape of the budget sets. Obviously, this approach is a data intensive one. It has been argued that individuals do not know the exact shape of their budget constraints and thus the differentiable approach could be a suitable substitute.

Our results indicate that proxying tax schedules by smooth continuous functions produces similar results as with the piece–wise linear approach when the degree of progressivity is high. For example, the mean compensated wage elasticities are almost the same in both models (approximately 0.21) and also other covariates behaves similarily.

Our tentavive conclusion is that because of the high degree of progressivity these two approaches actually generate very similar budget constraints and because our functional specification and data set is exactly the same one in both approaches, then it is no surprise that similar results appear. For example, MaCurdy and Flood (1992) found that when using Swedish data, these two methods produced almost identical results also in the case when only survey data was used. Their conclusion was also that this result might depend on the degree of progressivity. This finding is not necessary valid when the tax system includes only few tax brackets. Results by MaCurdy at al. (1990) show that in the case of the U.S. data, the piece–wise linear approach yields larger labour supply responses than the piece–wise linear approach.

Our results show that one should take the construction of the budget sets seriously. Results vary with differently defined budget sets. Especially the way how wage and exogenous income term is defined might give different answers to the key policy elasticities. These elasticities are almost without an exception used in microsimulation models to study the behavioural aspects of the tax reforms.

Acknowledgements
I am especially indebted to Richard Blundell, Richard Dickens, Ian Walker and Ilpo Suoniemi for very helpful and detailed comments and suggestions. Also discussions with Heikki A. Loikkanen, Matti Tuomala and Matti Viren have been very useful. The usual disclaimer applies. Financial aid from the Yrjo Jahnsson Foundation is gratefully acknowledged.

References

Barmby, T. and Smith, N. (2001) Household Labour Supply in Britain and Denmark: Some Interpretations Using a Model of Pareto Optimal Behaviour, *Applied Economics* **33**, 1109–1116.

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

Blomquist, S. (1995) Restrictions in Labour Supply Estimation: Is The MaCurdy Critique Correct?, *Economic Letters* **47**, 229–235.

Blomquist, S. (1996) Estimation Methods for Male Labour Supply Functions: How to Take Account of Nonlinear Taxes, *Journal of Econometrics* **70**, 383–405.

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

Blomquist, S. and Newey, W. (2002) Nonparametric Estimation with Nonlinear Budget Sets, *Econometrica* **70**, 2455-2480.

Blundell, R., Duncan, A. and Meghir, C. (1998) Estimating Labour Supply Responses Using Tax Policy Reforms, *Econometrica* **66**, 827-861.

Blundell, R. and MaCurdy, T. (1999) Labour Supply: A Review of Alternative Approaches, *Handbook of Labour Economics.* Edited by O. Ashenfelter and D. Card. North-Holland.

Dudewicz, E. and Mishra, S. (1988) *Modern Mathematical Statistics.* John Wiley and Sons.

Flood, L. and MaCurdy, T. (1992) Work Disincentive Effects of Taxes: an Empirical Analysis of Swedish Men, *Carnegie–Rochester Conference Series on Public Policy*, **37**, 239–278.

Fraser, S. and Paton, D. (2003) Does Advertising Increase Labour Supply? Time Series Evidence from UK, *Applied Economics*, **35**, 1357–1368.

Garcia, I. and Marcuello, C. (2002) Family Model of Contributions to Non-Profit Organizations and Labour Supply, *Applied Economics* **34**, 259–265.

Hall, R. (1973) *Wages, Income and Hours of Work in the U.S. Labour Force.* Markham Chicago.

Hausman, J.(1980) The Effect of Wages, Taxes and Fixed Costs on Women's Labour Force Participation, *Journal of Public Economics*, **14**, 161–194.

Hausman, J. (1985) The Econometrics of Nonlinear Budget Sets, *Econometrica*, **53**, 1255–1282.

Heckman, J. (1979) Sample Selection Bias as a Spesification Error, *Economet-*

rica, **42**, 153–162.

Heckman, J. (1983) Comment, *Behavioural Simulation Methods in Tax Policy Analysis*, Edited by M. Feldstein. University of Chicago Press.

Ilmakunnas, S. (1992) Income Taxation, Hours Restriction and Labour Suply, Working Paper 113, Labour Institute for Economic Research, Helsinki.

Kuismanen, M. (1995) Labour Supply of Finnish Males, *Government Institut for Economic Research*, Research Report n. 14.

Kuismanen, M. (2004) Labour Supply and Income Taxation: Estimation and Simulation Exercise for Finland, *Finnish Economic Papers*, **1/2005**, 16–30.

MaCurdy, T., Green, D. and Paarch, H. (1990) Assessing Empirical Approaches for Analyzing Taxes and Labour Supply, *Journal of Human Resources*, **25**, 415–490.

Moffit, R. (1986) The Econometrics of Piecewise–Linear Budget Constraints, *Journal of Business and Economic Statistics*, **4**, 317–328.

Mroz, T, (1987) Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions, *Econometrica*, **55**, 765–800.

Pudney, S. (1989) *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*, Basil Blackwell, Oxford.

Footnotes

[1]Pudney (1989) shows how difficult it is to actually construct the budget sets accurately
[2]As far as we know, their method has not been used in empirical studies.
[3]Note that one can always derive indirect and direct utility functions from the labour supply function or vice versa.
[4]For the completeness we can show that optimum can be found from the zero hours $h = 0$ if

$$h^*(w_1, y_1, z; \alpha, \beta, \gamma) \leq 0$$

or correspondingly from the maximum hours $h = H_n$, if

$$h^*(w_n, y_n, z; \alpha, \beta, \gamma) \geq H_n.$$

[5]It is important to realise that in the literature measurement error is interpreted in two different ways. The older interpretation is that the *positive* observed hours is measured with error. In this case one must choose the density function which ensures that reported hours of work are always positive with a feasible ε. The second interpretation is the optimisation error which reflects to the degree to which individuals' actual hours of work deviate from their desired hours. Thus, it is possible to observe that some individuals are not working even their desired hours are strictly positive because a realisation of ε causes measured hours to be non-positive. Most studies made are consistent with this latter interpretation.
[6] We experimented using other combinations of (t_i, I_i), but the above specification produced the best approximation.
[7] As an example $\int \Phi dI = I\Phi + \phi$ etc. For the technique see Dudewicz and Mishra (1988).
[8]LFS data set also includes some information about individuals' financial situation.
[9]Obviously this is not to say that the differentiable method is the right one when estimating labour supply functions in the presence of non–linear income taxation. In the differentiable budget constraint approach we need fewer restrictive assumptions in the background than in the piece–wise linear approach and thus we can "test" if these assumptions are important or not.
[10]Results are available from the author upon request.
[11]All elasticities are calculated using mean values.

Appendix 1: Definitions of the variables

union=1, if the respondent is a member of a union
age=Age of the respondent
age2= Age squared
educ10=1, if the respondent has 10 years of education. Otherwise zero.
educ12=1, if the respondent has 11-12 years of education. Otherwise zero.
educ14=1, if the respondent has 13–14 years of education. Otherwise zero.
educ15=1, if the respondent has 15+ years of education. Otherwise zero.
ueduc=1, if the respondent has a university degree in the following fields: Technology, business, law, natural science and social sciences
nchild=Number of dependent children.
cdum1,...,cdum4= Dummy variables for the youngest child. Age groups are 0–3,4–6,7–9 and 10+.
schild=Number of children aged 0–3.
cchild=Number of children aged 4–6.
bchild=Number of children aged 7–9.
exp= Working experience
exp2= Experience squared
tenure= Duration of the current job
tenure2= Square of tenure
pjob=1, if respondent has a permanent job
phusb=1, if respondent's husband is working
stat=1, if the respondent is a white–collar worker and 0 if a blue–collar worker.
socio=1, if the respondent is a upper white–collar worker. Otherwise zero
hwage= Hourly wage rate.
shwage= Subjective hourly wage rate.
exo= Unearned income.
exo+hnet= Unearned income plus husband's net incomes.
south=South Finland.
west=West Finland.
east=East Finland.
middle=Middle Finland.
north=North Finland.
lapl=Lapland.

Appendix 2: Descriptive Statistics

Descriptive statistics: participants and non–paricipants

Variables	*Participants*	*non–participants*
Hours	1855.58(560.10)	
union	0.71(0.45)	0.21(0.41)
age	41.29(8.47)	43.13(11.61)
educ10	0.30(0.46)	0.26(0.41)
educ12	0.18(0.39)	0.16(0.36)
educ14	0.05(0.21)	0.06(0.24)
educ15	0.07(0.26)	0.06(0.24)
cdum1	0.13(0.34)	0.32(0.46)
cdum2	0.12(0.33)	0.06(0.24)
cdum3	0.12(0.32)	0.05(0.21)
cdum4	0.25(0.44)	0.10(0.30)
workexp	19.50(9.32)	17.30(11.60)
jobdur	8.60(8.34)	
permjob	0.77(0.41)	
phusb	0.86(0.33)	0.68(0.46)
hwage	48.81(23.99)	
shwage	44.28(19.35)	
exo	5525.66(14028.9)	6935.43(11183.8)
exo+hnet	84590.67(61197.5)	77666.32(46447.6)
south	0.25(0.44)	0.21(0.41)
west	0.16(0.36)	0.13(0.33)
east	0.19(0.39)	0.24(0.43)
middle	0.14(0.34)	0.15(0.36)
north	0.18(0.39)	0.18(0.38)
lapl	0.08(0.27)	0.08(0.27)

For definitions of the variables see Appendix 1.

Table 1. Wage Equation

Wage Equation. Dependent variable: ln hwage.		
Variables	*Coefficient*	*Standard Error*
Constant	2.83342	0.2511
Age	0.01753	0.0135
Age2	-0.00017	0.0001
Educ10	0.06881	0.0024
Educ12	0.19534	0.0297
Educ14	0.27270	0.0469
Educ15	0.51690	0.0469
Exp	0.01659	0.0053
Exp2	-0.00027	0.0001
Tenure	0.02410	0.0038
Tenure2	-0.00045	0.0001
Pjob	0.04720	0.0299
Husb	0.00760	0.0290
Stat	0.10338	0.0241
Socio	0.23919	0.0366
Nchild	-0.03065	0.0104
South	0.15898	0.0222
Exo+hnet	3.95e-07	1.67e-07
Occ. dummies	Yes	
Ln L	-1221.91	

NOTE: The selection index is a function of the individual, geographical and demand side variables. The selectivity effect was statistically significant. Reference group for occupation is manufacturing workers.

For definitions of the variables see Appendix 1.

Table 2. Results for the labour supply functions

Maximum Likelihood Estimates

Asymptotic Standard Errors in Parenthesis

Variables	*piece–wise linear budget constraint*	*Differentiable budget constraint*
Constant	-2.57905	-2.70210
	(0.55650)	(0.55474)
Ln W	0.37046	0.39121
	(0.12135)	(0.12444)
Exog. inc	-0.00045	-0.00045
	(0.00022)	(0.00020)
Cdum1	-0.33917	-0.33823
	(0.09948)	(0.0990)
Cdum2	-0.00487	-0.00482
	(0.10437)	(0.10431)
Cdum3	0.09616	0.09601
	(0.10050)	(0.10049)
Cdum4	0.14310	0.14310
	(0.07690)	(0.07688)
Age	0.16118	0.16144
	(0.02484)	(0.02488)
Age*Age	-0.00227	-0.00229
	(0.00028)	(0.00028)
Sosio	0.19945	0.19912
	(0.09521)	(0.09567)
Nkids	-0.08419	-0.09011
	(0.03235)	(0.03239)
σ_ϵ^2	0.98208	0.96998
	(0.01907)	(0.01918)
Ln L	-2669.61	-2675.56

Note: In both of the above models, the dependent variable (yearly hours) is divided by 1000. The exogenous income variable contains only person's own exogenous income components (net) and it is divided by 100. For definitions of variables see Appendix 1.