Labour Supply, Unemployment and Income Taxation: An Empirical Application

Mika Kuismanen[1]
Government Institute for Economic Research, Helsinki[2]
and
University College London[3]

August 1997

[1]The author is indebted to the seminar participants at Tillburg University and University College London. Author is particularly indebted to Richard Blundell, Heikki A. Loikkanen and Luigi Pistaferri for their helpful comments. All errors are and will remain author's responsibility.

[2]I thank Yrjö Jahnsson Foundation and Government Institute for Economic Research for financial help. I especially like to thank Government Institute for Economic Research for allowing me to visit in the University College London.

[3]**Address for correspondence:** Department of Economics. Gower Street, London, WC1E 6BT. FAX:44–(0)171–916 2775. e-mail: m.kuismanen@ucl.ac.uk

Abstract

In most cases estimation of the labour supply function is based on the assumption that individuals can work as many hours as they desire and/or that all individuals who report zero hours do not want to work. A closer look to micro data shows that this is not always the case. Following the work of Blundell, Ham and Meghir[7][8] we implement a model of female labour supply, participation and employment which incorporates unemployment and discouraged workers. In addition, we incorporate nonlinear income taxation into the sample likelihood function. Data is taken from the Finnish Labour Force Survey (LFS) and from the Finnish Tax Register. We use repeated cross-section data sets for the years 1987, 1989, 1991 and 1993. Unlike in some previous studies we do not pool these data sets. Our results indicates that the wage and income elasticities for females estimated in previous studies (using Finnish data) may be overestimated due to the simple methods used.

Keywords: Labour supply, Unemployment, Discouraged workers, Non-linear taxation, Index functions, ML-method.

Tiivistelmä

Työn tarjontafunktion estimointi on useimmissa tapauksissa perustunut olettamukseen, että yksilöt voivat valita haluamansa työtuntimäärän. Toisen vahvan olettamuksen mukaan yksilöt, joiden raportoidut työtunnit ovat ei-positiiviset, eivät halua työskennellä. Molemmat oletukset ovat kumottavissa tarkastelemalla tarkemmin mikrotason aineistoa. Tässä tutkimuksessa sovelletaan Blundell, Ham ja Meghirin kehittämää empiiristä mallia, jossa otetaan työn tarjontafunktiota estimoitaessa huomioon työttömät ja työtä etsivät yksilöt. Työssä yleistetään ym. mallia ottamalla huomioon myös ei-lineaarinen tuloverojärjestelmä. Aineistona käytetään työvoimatutkimuksen vuosihaastatteluja vuosilta 1987, 1989, 1991 ja 1993. Saadut estimointitulokset indikoivat, että aikaisemmissa tutkimuksissa saadut palkka–ja tulojoustojen arvot voivat olla ylöspäin harhaisia johtuen käytetystä ekonometrisestä lähestymistavasta.

Asiasanat: Työn tarjonta, työttömyys, työn etsintä, ei–lineaarinen tuloverotus, Ml-estimointimenetelmä, Indeksi funktiot.

1 Introduction

In the last few decades the estimation of labour supply functions has been one of the most active research areas in the labour economics. A vast majority of the empirical work is based on the neoclassical labour supply model which isolates the wage rate (depends on hours of work) and exogenous or unearned incomes (does not depend on hours of work) as a main economic factors that determine the individual's allocation between labour supply and leisure time. It is well documented that estimated income and substitution effects vary considerable from one study to another and they are sensitive to economical and statistical assumptions. See e.g. Mroz[29]. There are no generally accepted robust ways to estimate labour supply functions and the problem gets even more complicated when we introduce some real life phenomena to empirical analysis, like nonlinear taxes or fixed costs, just to mention few.

Earlier empirical work has in many cases been based on the crucial assumption that individuals can work as many hours as they desire and if they are not working that is because they do not want to work. These assumptions are clearly debatable. For example, demand side restrictions can prevent individuals from choosing their desired amount of hours and increased unemployment in many countries is not probably due to individuals increased unwillingness to work. In fact, empirical evidence do not support the latter claim. Many micro–data sets shows that a significant number of individuals out of work are actually seeking employment.

In this paper our aim is to relax some of these assumptions and estimate labour supply functions taking into account three different groups in the labour markets. These groups are i) employed ii) non–participants and iii) job–seekers. We follow the method developed by Blundell, Ham and Meghir[7] [8]. This method is actually a generalisation of the Cragg's[12] double–hurdle model. In addition, we explicitly model non–linear income tax system and take it into account in our econometrical application. We argue that our data set is more suitable to this kind of analysis than some other data sets which have been used in similar kinds of studies.

Among the first theoretical studies which relaxed the assumption that unemployment represents leisure time were Malinvaud[27] and Grossman and Hart[15]. In his model Malinvaud presents a keynesian disequilibrium model where unemployment represents a constraint on labour supply behaviour and Grossman and Hart developed a theoretical implicit contract model where the unemployed persons are willing to work at the value of their marginal product. In the empirical work the above mentioned problems and other generalisations, like taxation, also started to get attention in the 80's. Ham[16] develops an econometrical model for labour supply which takes unemployed and also underemployed into account. Blundell, Ham and Meghir[7] use Cragg's double–hurdle idea to extend the model to unemployed and discouraged workers. Taxation and labour supply literature is vast, see e.g. Hum and Simpson[22]. These extensions to the basic labour supply model are important because when the workers are truly constrained, the standard approach to estimate labour supply functions is inappropriate.

The familiar Tobit specification, which is usually used in labour supply studies, is inadequate to deal with above mentioned extensions. The basic problem is that the (simple) Tobit spesification do not take into account the different reasons for an individual being out of work i.e. it characterises those who report zero hours of work as not wanting to work. As already mentioned above, this is not a plausible assumption in all situations.

In an empirical application we study Finnish labour markets in years 1987, 1989, 1991 and 1993 and particularily we concentrate married females aged 25–60 years. This period is exceptional in Finnish economic history. Finland experienced a steady economic growth during the early 1980's and then in late 1980's the economy boomed. Recession came in early 1990's and it was unseeingly severe, as we will see later. During this period, for example, the annual growth rate of GDP turned from 5 percent to negative and unemployment rate rose to a highest level ever; in 1993 it was 17.9 percent. As Blundell, Ham and Meghir[8] argue, particularly if economy is experiencing such an extreme business cycle, it does not seem sensible to simply assume, a priori, away unemployed workers and discouraged workers. The other aspects which makes this period interesting is that in 1989 finnish income tax legislation went through a major changes. Roughly saying, before 1989 the tax system was quite complicated and marginal tax rates were high (even in Scandinavia). For details, see appendix 2.

Incorporating unemployed and discouraged individuals means that we have to deal with zero hours observations, i.e. we do take the labour force participation into account in empirical spesification. Mroz[29] shows that incorporating labour force participation to estimation may change the results. The statistical procedure to deal with non–participants and employment seekers is to introduce two latent index functions. The other, namely employment index, tries to model the probability of getting a job and the participation index just simply tries to model the participation process. So, our likelihood function consists at three different parts respective to contribution of these three different groups.

During the recent decades it has been increasingly apparent that progressive income taxation, which causes discontinuities to the budget constraints individuals face, may have effects on labour supply behaviour. In some tax schedules marginal wages fall subtancially as workers cross points of discontinuity in their budget sets (in other words, when workers move from one tax bracket to another) and it is important to examine the effects of such discontinuities on work incentives. This topic has been studied heavily during the last few decades and the results and methods vary substiantially. Probably one of the few things on which all researchers agree is that estimation of the labour supply function in a robust and generally accepted way is a very difficult task to do and one of the sources which creates difficulties is nonlinear income taxation.

In this paper we use the method developed by Hausman[17][18]. In this approach we model the whole budget constraint the individual faces. This approach has been criticised for different reasons. One of the main problems is that this approach requires a lot from the data. Heckman[21] argues that in the most cases the budget constraint cannot be measured accurately enough. We argue that our data set overcomes some of these problems and we are able to construct the budget sets accurately. In addition, we take into account different municipality income tax rates, which is crucial as we show in the later sections. We also develop a method to calculate the tax deductions which is usually ignored in previous studies.

Changing macroeconomic environment affects the amount of unemployed in the economy and thus affects the labour market participation rates. So, it is quite natural to think that business–cycle variables have some explanatory power in empirical labour supply spesifications, particularily in the participation index function and in the job availability index. Unfortunately, many micro–data sets do not have macro level information or the amount of it is limited. This is also the case in our data sets. We have collected some business–cycle variables and the merged them into our data and the empirical part we will test their significanse. Blundell, Ham and Meghir[8] find that these business–cycle variables are statistically significant in the participation and in the job availability indices.

Our data comes from two different sourches. The main data set is the Finnish Labour Force (LFS) survey for years 1987,1989,1991 and 1993. These are independent cross-section data sets. We merge this data set to the Tax Registration Data which includes all the income information concerning corresponding individual in the LFS plus their possible spouse's income information. This data is taken from the Finnish Tax Register, so it is an archive data. Income data sets includes all possible income sources and thus it gives us a good opportunity to model individual's budget set accurately enough. The reasons why we do not pool these data sets are following. First, modelling different tax schedules into the same likelihood is extremely complicated, or even impossible. Secondly, the method of collecting the data sets has changed. Thirdly, some definitions of the variables have changed and fourthly, the survey is collected only every second year.

The remaining sections of the paper are as follows. In section 2 we shortly discuss some points from the labour supply theory which give backround to our empirical work. Following section describe the data used in this study and in section 4 we describe the Finnish tax system and discuss shortly how economic situation changed during our sample period. In the next section we introduce our econometrical approach we also discuss our choice of the labour supply function. Section 6 presents the results and section 7 concludes.

2 Some remarks from the labour supply theory

Since the standard labour supply theory is quite well known and well documented, we only shortly comment those points of it which we think are relevant for our empirical purposes.

It is usually assumed that individuals maximize the value of their utility function with respect to two constraints, namely budget and time constraints. One way to express the utility function is to assume that individual's utility depends on consumption c and leisure l. The Budget constraint can be written as $wh + y = pc$. w refers to (market)wage rate and y refers to exogenous (unearned) income. h is the hours worked and p is the price of consumption.[1] Time constraint can be written as $h + l = T$ with the obvious notations. Assuming that our utility function is strictly quasi–concave and because our constraints are linear, optimal allocation between leisure time and consumption can be found from the point where the ratio of marginal utility of the consumption and the marginal utility of the leisure equals the ratio of corresponding prices, i.e. $[(\partial u/\partial l)/(\partial u/\partial c)] = [w/p]$. In other words, the solution means that we have a unique tangency point between indifference curve and budget constraint for each wage rate. Obviously this is not the case under the nonlinear budget constraint and below we will return to this issue.

But how about those individuals who are not participants in the labour markets? We usually assume that individual is participant if her reservation wage is lower than the wage she gets from the markets. In our data individuals can be found in the three different labour market positions: 1) working, 2) seeking work and 3) not working. So, why some people seek to work and some not? One reason is that expected benefits from seeking employment for some individuals exceeds the search costs. It can also be the case that individual would like to work at her market wage but she does not partici-

[1] Usually c denotes Hicksian composite commodity and thus p is set as unity.

pate because the search costs exceeds the expected benefits. Blundell, Ham and Meghir[7] develops a theoretical model following the work of Burdett and Mortensen[10]. The outcome is familiar and the individual chooces that state in time t where her value function is greatest. For technical details see above mentioned articles. In section 5 we show how this can be incorporated into the statistical model.

As we mentioned above the basic neoclassical labour supply analysis gets complicated when the budget constraint becomes nonlinear. In the case of linear income tax it is straightforward to derive Slutsky condition and see that linear tax has only effect of a scale factor which dampens the substitution and income effects of a change in the wage rate compared to case of no taxation. Already this simply exersice shows, that estimating labour supply function without taking income taxation into account may lead to biased (at this case upward) wage estimates.

In the presence of progressive income taxation the budget constraint becomes nonlinear, or strictly speaking piecewice linear with discontinuity points where the marginal tax rates change.[2] To the best of our knowledge, deriving comparative statics results in this case is impossible. Let us illustrate this point by a simple example. Let us assume that our tax schedule generates convex piecewice linear budget constraint and our tax parameters are lump sum tax, exemption level, change of tax bracket limit, change of marginal tax rate and change of gross wage rate. In real life tax reforms usually more than one of these change simultaneusly, but let us change them one at a time.

A change in a lump sum tax shifts our convex piecewise linear budget constraint upwards or downwards in a parallel way, so the only effect is the income effect. When we change the exemption level, it will affect exogenous income and all the virtual income components and all the points where the tax brackets changes. It will not affect marginal wages. When we change the limit of some tax bracket j, it will have the following effects: it will leave all lower segments unaffected. It will change the tax bracket limit j and it will increase all the virtual income components above the segment j. A change of the marginal tax rate (say it increases) for the bracket j will leave all the lower segments unaffected. It will change the slope of the segment j (decrease) and change the corresponding virtual income (increase). Finally, a change of the gross wage rate will change all the slopes and it will change all the tax bracket limits, but it has no effect on virtual income components.

Now, it is clear that changes in the tax parameters have different effects to

[2]In this context word progressivity means that the marginal tax rate increases with income, but it is constant within the tax bracket.

different individuals depending their initial location in the budget constraint, i.e. reaction of different individuals with a different initial locations may even be opposite. We can also see that small changes in wage rate and virtual income may keep individual in the same segment (when predictions from the 'basic model' are valid)but there is always the likelihood that individual will change the segment. And finally we know that utility maximization in the presence of the kink is compatible with many different marginal rates of substitution, so individual may as well stay at the kink.

Above example gives us an idea how we should model the income tax schedule in our empirical work and in fact, we will do so following a method developed by Hausman[17][18]. Because a lot of critique towards this method has been araised, and one of the resons have been data sets used, we will next turn to describe our data set which we think is rich enough to model tax schedules.

3 The data

A sample of married women of age between 25–60 is drawn from the Finnish Labour Force Survey (LFS) for the years 1987,1989, 1991 and 1993. These are independent cross-section data sets and they include 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 final sample is drawn randomly by age and gender. In years 1987,1989 and 1991 sample sizes were around 8000 individuals and in 1993 the size was only 5500 due to a change in data collecting procedure. Before 1993 the LFS were collected in every second year and in the 1987 data set an substitute interviewees were allowed. Comparing the data sets between these years should be done by some caution, because the method of collecting the data and some definitions have changed during this period. As an example, in 1989 an substitute interviewees were not allowed which means that the loss is bigger than in previous years but the number of 'not known' answers is respectively smaller and answers should also be more precise than before.

Income data corresponding those 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 data and it includes approx. 70 variables on individuals earnings. Of course, it is very unlikely that someone's earnings are composed from all of these components. However, the data shows that individuals' earnings comes from the very different sourches. 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 possible spouse, so all in all we have approx 140 variables (if married) to construct the budget sets individuals face.[3]

3.1 Hours of work, wages and exogenous incomes

In the empirical labour supply analysis the following three variables are under special attention; namely hours of work, wage rate and unearned incomes. That is because the classical question usually asked in the labour supply studies is the following: what happens to the hours of work when wages and/or unearned incomes change? In order to answer this question we need to estimate the substitution and income elasticities i.e. we need (ideally accurate) information on wages and unearned incomes in our deterministic part of the regression function. ¿From the statistical point of view it is crucial that the hours of work variable (endogenous variable) varies enough around its mean. When studying a labour supply responses in Finland (or any other Scandinavian country) we should especially examine this carefully.[4]

Our data includes information both from the regular weekly hours and from the hours worked in the survey week[5]. In this study we use the regular weekly hours. In addition, we take into account regular hours in the second job. Unfortunately, we do not have any information for the overtime working hours. This is a clear drawback, because in many cases it is reasonable to assume that overtime hours are the flexible part in the labour supply decision. But, on the other hand, wage rate from the overtime work usually differs from the regular one, so incorporating that into a statistical analysis is not necessary a straightforward task to do.

The data set does not have the direct information on individuals' hourly wage, thus we have to construct it using the income and hours of work variables. This procedure means that the possible bias in the hours of work variable shifts also into the marginal hourly wage rate. Statistically it means that the dependent and the independent variable are negatively correlated.

[3]We would like to stress that the income data used in this study is exceptionally rich compared to some other data sets, for example to Panel Study Income Dynamics (PSID, based on survey). Most of the variables have gone through several checkings by tax authoroties.

[4]A widely shared view is that Skandinavian societies are highly unionized and the labour markets are not very flexible. This is partly true. If we look hours of work by gender, we see that male hours are much more concentrated than female hours. This partly reflect the fact that the male dominated unions may have different objects in negotiations that the female dominated unions.

[5]In case of many individuals, worked hours in the survey week deviates from the regular hours, so calculating the yearly hours from the survey hours lead to unrealistic, much too high or low, values.

For example, if worked hours are smaller than their right value then the value of hourly wage becomes too high. We estimate the log wage equation using Heckman's method and the predicted values are used in the final analysis.

By definition unearned (exogenous) income does not depend on worked hours. In empirical work this definition is not always valid because, at least in the long run, it is quite difficult to say which income components depend on worked hours and which do not. The data set in hand gives us a good opportunity to construct different kinds of exogenous income variables. Unlike most studies made e.g. in UK, we do not have to use consumption information to evaluate this variable. We use the following income components when calculating the unearned incomes: interests (both taxable and non taxable),dividend payments, property incomes, sales profits, regular untaxable pensions, other regular subsidies etc. From all the components which are taxable we have subtracted the corresponding amount of taxes paid, so our variable measures net exogenous incomes. We have also taken into account spouse's net incomes. We argue that our data gives us enough detailed information to construct the budget sets which correspond the real ones.

3.2 Sample selection and some descriptive statistics

The original sample sizes are:

- 1987: 8155 observations $\Longrightarrow$ 4124 females
- 1989: 7825 observations $\Longrightarrow$ 3976 females
- 1991: 8301 observations $\Longrightarrow$ 4265 females
- 1993: 5831 observations $\Longrightarrow$ 2939 females

For the empirical analysis we select married women aged 25–60. There are many reasons for this choice. First of all, looking the data shows that female labour supply is more flexible than male labour supply. It is still the case that womens are secondary workers in most families. This means, a priori, tax reforms are more likely to affect their labour supply behaviour. Secondly, we select above age group because of the following facts which characterize the society. Compared to, for example UK or US, Finnish young people start to work[6] at much elderly age. The average graduation age from the university is around 26 and individuals graduating from vocational, trade schools etc. are usually over 20 years old. In addition, the data shows us

[6] We do not take into account summer jobs or other short term jobs students usually have between semesters.

that the proportion of females under 25 with regular incomes is low. The reason for the choice of the upper bound also reflects the features of Finnish labour markets. Retirement age is quite low compared to the some other European countries. It is common that women above 60 years old are on pension. As one of our objectives in this paper is to study how income taxation affect hours of work, we want that the majority of our final sample are (at least potentially) active in the labour markets.

Thirdly, we chose married[7] women because this allows us to study the role of spouse's incomes in the labour supply decision.[8] It is also the case that these individuals represent more or less the 'basic cases' and constructing their budget sets are much easier than for example, single mothers with children.[9]

In the second step we deleted some special groups like farmers and self-employed mainly because of the different tax system. After these selections the final sample sizes used in the analysis are: 2023 observations in 1987, 2037 observations 1989, 2094 observations in 1991 and 1466 observations in 1993.

Next, we shortly comment basic features of our data. A comprehensive data appendix can be found from the appendices 3 and 4. Lets's first look at labour supply behaviour. The participation rates in our sample in 1987,1989,1991 and 1993 are 69,72,67 and 61 percent respectively. When we compare these figures to the unemployment rates given earlier, the changes in participation rates more or less follows these figures, as expected. Difference in participation rates between 1989(peak year) and 1993 (recession year) is 11 percent and the difference between the women's unemployment rate in the corresponding years was 12.4 percent. Unemployment/participation rates vary geographically and the figures are consistently lowest in the Helsinki metropolitan area and highest in the East and North part of the country.

Finland, like the other Scandinavian countries, is traditionally an highly unionized country and in 1987 75(among participants) per cent belonged to some union. In 1989 the share declined to 71 percent, due to the well booming

[7]In principle, we should have also chosen the cohabiting cases but this leads to the following problem. First, we do not have any information on their partners incomes as we do in case of married couples. The size of cohabiting cases is less than a half a percentage point.

[8]Unfortunately we do not have information on spouse's hours, so we cannot study intra-household labour supply desicions using this data set.

[9]This is not to say that this group is not worth of studying. Actually, it would be very useful exersice to do that because, a priori, this group's labour supply behaviour is probably the most sensitive one for the tax and benefit system reforms. The practical difficulty also arises if one likes to use LFS, because the number of single mothers with children is relatively small.

economy. In the early stage of the recession one of the main arguments was that due to the high unionization the country is not flexible enough to respond exogenous shocks and that the share should be lower. Probably so, but individuals seemed to think differently and they joined to unions (in 1993 80 per cent of the females belonged to some union) or then employers fired the non-union ones during the recession. As a member of the union your unemployment benefit is usually higher than non–members' benefit.

When we divide the data to three different parts, namely for participants, unemployed seekers and non-participants, we can see differences between the groups. Unemployed seekers are slightly elder and they have a lower work experience than among participants, which can be due to the long unemployment spells. Also their husbands' participation rates are clearly lower than in the other two groups. Among the participants their spouses participation rates varied from 88 percent (1987) to 77 percent (1989) and corresponding figures among the unemployed seekers were 78 and 65 percent. In 1989 the mean of the spouses participation rate was only 50 percent. The husbands' participation figures among the non–participant females are also clearly lower than among participant females. These numbers indicate that women are more likely to be out of the work if their husbands' are not participating to labour markets.

Geographical location also seems to differ between these groups. If you are unemployed seeker or non–participant you more likely live away from the Helsinki metropolitan area. Non-participants have more young children, as expected, (0–2 years) than participants have. Non–participants have higher exogenous (when we do not take into account husbands income)incomes than participants have, but this is not true after taking account husbands incomes.

Let's take a closer look at the hours of work variable. As from the cross-tabulations can be seen blue-collar workers are more likely to have zero hours and also naturally union members are more likely to have positive number of hours. Among women who have children those who have two child have the highest probability to be out of work. For closer look see appendix 4.

4 Finnish income tax system and economy in 1987–1993: an overview

The income tax system consists two parts: a progressive state income tax and a proportional local (municipal) income tax. In addition, individuals contribute to the National Pension Insurance (NP) scheme and National Health Insurance (NH) scheme, which are proportional to income changes. Roughly speaking, the tax liability in state tax and municipal (or local) tax is the same excluding the tax deduction system. A Further distinctive feature in Finnish tax system compared to some other European countries is that all individual are separate tax units. Husband's marginal tax rate does not affect wife's tax contribution or vice versa.

In 1987 the state income tax schedule was composed of 11 marginal tax rates varying from 6 per cent to 51 per cent. The tax deduction system was complicated and it was favourable for high income earners. Tax reform took place in 1989. Following the other tax reforms made previously in other countries the goal was to decrease marginal tax rates and to simplify the tax system considerably. So, after the income tax reform the tax schedule included 6 bracktets and the marginal tax rates varied from 11 to 44 per cent. In 1991 and in 1993 the tax schedules were only slightly different from the 1989 one. See appendix 2 for further details.

To the best of our knowledge all previous studies have assumed fixed local tax rate in empirical spesifications. We argue that this is an questionable assumption. For example, in 1989 the local tax varied from 14 per cent to 19.5 percent and in 1993 the variance was even higher due to the recession.[10] We also have to keep in mind, that the local tax is paid from all incomes unlike in the case of state income tax where the exemption level varies from year to year. In this study we have calculated, using the additional information,[11] local tax rates to all individuals. For example, in 1989 the mean of the local tax rate was 16.38 per cent and the mean of calculated tax rate is 16.21 per cent. It is obvious that taking the variation in local tax rates into account when constructing the budget constraints lead to more precise procedure.

One other small innovation used in this study is that we have developed a formula to calculate the accepted tax deductions.[12] Previous studies have assumed a constant tax deductions (or no tax deductions at all, which is the same thing) for all individuals which means that authors have assumed independence between the tax deductions and the income level. This is clearly untrue, at least in Finnish tax systems. We use the tax function and the tax parameters to calculate the tax deductions individually. Of course, this procedure is not a perfect[13], but it confirms the result that the high earners have also higher deductions and thus makes the actual tax system less progressive than it is usually thought to be. As an example, in 1989 estimated tax deductions varied for females from 0 FMK to 29 500 FMK. If

[10]Poor municipalities rose their tax rates considerable during the recession years.

[11]Statistic Finland deletes the municipality code from the data sets due to the legislation, but we can trace it using the income data sets.

[12]We owe thanks to Ilpo Suoniemi who asked us to think this possibility.

[13]For example, the tax legislation allows to shift deductions between couple in some special cases and tracing this using the given information is impossible.

we compare this information (by income quintiles) to one calculated by tax authoroties, we get almost similar results. Again, taking tax deductions into account we get more precise picture about individuals' budget constraints.

4.1 Finnish economy between 1987 and 1993

As we previously mentioned, incorporating unemployed individuals and job seekers into the labour supply analysis can be useful when economy goes through changes in macroeconomic environment. ¿From 1987 to 1993 Finland experienced a boom and recession which was (is) a severe one as we will see. In the following we analyse shortly what has happend in Finnish economy during the last ten years or so.

The 1980's were time for steady economic growth and especially the latter part of it was a period when the economy overheated. During the 80's the growth rate (annual change in the GDP) was above its historical trend and the growth rate in period 1986 to 1988 was over 5 per cent and in 1989 it still was 4.3 per cent. Naturally, investment rate was also high in that period and unemployment rate fell below its historical trend and it reached the lowest value in the 1989.

During the first half of the 1990, state of the economy deteriorated quickly. Collapse of the ex–Soviet Union was one reason. Before the collapse Soviet Union's share in Finland's foreign trade was a significant. Secondly, the international recession's effects started to affect in that time. Thirdly, the Finnish economy was clearly overheated in late 80's and fourthly, it has been argued that the liberalization of the monetary policy was done in the wrong moment (1987). In 1990 the annual change in GDP was 0 per cent and from 1991 to 1993 it was negative. 1994 was the first positive growth year. We get better picture from the depth of the recession when we look at the annual change of the Private Gross Fixed Capital Formation. It started to sunk already in 1990 and the highest negative years were from 1991 to 1993 when the figures were -23.1,-19.6 and -19.5 correspondingly.

In the same time period labour markets changed drastically. Unemployment rate in the 1987 was 5.1 per cent and in the 1990 it was 3.4 per cent which is historically very low figure. Unemployment rate rose sharply and in the 1993 it was as high as 17.9 per cent. Recession treated different groups unequally. It was worst for young males and e.g. the unenployment rate in case of 20 to 24 years old males in 1993 was 36.1 percent. Women's unemployment rates never rose as sharply as males. The main reasons for this is that there are proportionally more women working in the public sector (which did not fire employees as much as the private sector) and many women moved volontarily out from the labour markets. In the late 80's there was a reform concerning the child benefit system and it basically made it easier (in money terms) for women (or men) to start to take care of children at home by themselves. When the recession hit many women took advantage of this reform.

In the empirical application we study the significancy of some business-cycle variables. More spesifically, we have collected unenmployment information by the gender and location. Of course, there are many other possible business–cycle variables which affect individuals' labour supply behaviour, but the availability of these variables is limited.

5 Econometric spesification to incorporate unemployed and discouraged workers

We start with the following general labour supply spesification,

$$h_i = h^*(w_i, y_i, z_i; \alpha, \beta, \gamma) + \varepsilon_i \tag{1}$$

where w_i represents the marginal wage rate, y_i represents the unearned (exogenous) income and z_i represents a vector of demographic and sosio-economic variables. α, β, γ are parameters to be estimated. These elements form our deterministic part of the regression function. The stochastic error term is assumed to be normally distributed, such that $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

Mostly used model in labour supply analysis is the simple Tobit spesification which includes two parts; a discrete and a continious ones. Using our above labour supply spesification we can formulate the model in the following way.

$$h_i = \begin{cases} h_i^*(.) + \varepsilon_i & \text{if } h_i^*(.) + \varepsilon_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

Now, the probability of observing zero hours of work can be easily shown to be

$$\begin{aligned} Pr[h_i = 0] &= Pr[\varepsilon_i \leq -h_i^*(.)] \\ &= F[-h_i^*(.)] \\ &= 1 - F_i \end{aligned}$$

Where F_i is the distribution function of h_i evaluated at $h^*(.)$.

The corresponding log-likelihood function can be written as

$$\begin{aligned} ln(L) &= \sum_{i=0} ln[Pr(h_i \leq 0)] + \sum_{i>0} ln[Pr(h_i > 0) * f(h_i \mid h_i > 0)] \\ &= \sum_{i=0} ln(1 - F_i) + \sum_{i>0} ln f(h_i), \end{aligned} \tag{2}$$

where summations are over the non-workers (zero reported hours) and workers (positive reported hours). It is crucial to realize that the above sample likelihood is not fully specified without the wage equation (or some other (statistical) method which produce wage rates for nonworkers) because we usually do not have wage information for nonworkers. We estimate the wage equation using Heckman's[20] correction to control selection bias.[14] Wage equations are estimated separately for all sample years and results are shown in Section 6. The predicted wages are used in the final analysis.

The basic situation in above model is following. There is an event which at each observation may or may not occur (in our case, labour market participation). If it does occur, then it will be associated with positive continious random variable. If it does not occur, then this variable has a zero value. The tobit model presented above assumes that whether to participate (or in the case of consumer demand problem, acquire) and how much to supply hours (or how much to acquire) are the same in this model, in the sense that the same variables and parameters occur in discrete and continious part of the likelihood. This is crucial assumption. An other way to say this is that the probability of individual reporting zero hours is equivalent to the probability of her/his not having positive *desired* hours of work. Now, it is obvious that at least job seekers want to work and so they have positive desired hours of work.

In what follows we will utilise Cragg's[12] seminal idea that we allow the determination of the size of our labour supply variable when it is not zero to depend on different parameters or variables from those determining the probability of its being zero. In the context of labour supply Blundell, Ham and Meghir[7] were the first who used the Cragg's double–hurdle idea. But in that study they did not explicitly model job seekers.

So, our data clearly shows that there are individuals who are willing to work but cannot find it because of search costs, economic conditions or who are unable to secure employment given their own demographic characters

[14] A potential observation i is observed (in this case wage) if

$$A_i\zeta + \mu_{1i} > 0$$

where μ_i has a standard normal distribution. Simultaneously, there is another regression function

$$W_i = X_i\kappa + \sigma\mu_{2i}$$

where μ_{2i} also has a standard normal distribution, but μ_{2i} is potentially correlated with μ_{1i} with correlation ρ. When $\rho \neq 0$, standard regression techniques applied to the second equation yield biased estimates. We estimate our wage equation using the maximum–likelihood method. The other possibility is to use Heckman's two step method.

(Duncan[13]). So we observe in our data three different groups of individuals, namely those who are employed, non–participants or job–seekers.

If we want to model all these three groups statistically, we need some additional information. FLS includes, at the first stage, information from the individuals labour force participation. If one does not work then more detailed questions are asked and we can separate unemployed individuals who are actively seeking work, are not seeking employment, are in pension etc. Using this information, we can proceed first by defining an index function which models an individual's desire to participate to the labour market. This can be done in the following way.

$$P_i = \begin{cases} 1 & \text{if } P_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

with,

$$P_i^* = X_i\gamma + \nu_i, \tag{3}$$

where $\nu_i \sim N(0,1)$ and X represents a matrix of independent variables. Like usually, the probability of participation can be defined to be equivalent to the probability that $P_i^* > 0$. As Duncan[13] mentions there may exist a case where an individual wants to work ($h_i^* > 0$) but chooses not to seek employment ($P_i^* \leq 0$) due to, for example, positive search costs. From this we can derive that the probability to find a *discouraged worker* from the data is $Pr(h_i^* > 0, P_i^* \leq 0)$.

An other group which Tobit spesification does not take into account is those individuals who are seeking employment. Statistical modelling of this group follows the same strategy as above. We define another index function to model those individuals who actively seek employment. For the subset of seekers and workers in the data we can define an employment probability index E_i such that

$$E_i = \begin{cases} 1 & \text{if } E_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

with,

$$E_i^* = Z_i\varsigma + \varepsilon_i, \tag{4}$$

where Z reresents a vector of regional, demographic and economic characteristics and the stochastic term is again assumed to follow standard normal distribution.

Now, with zero search costs, and for the subset of seekers and workers in the sample, the probability of observing a working individual is given by $Pr(E_i^* > 0)$. This means that we can find from the data individuals for whom desired working hours are positive, who are not discouraged from working, but who have not find a job ($E_i^* \leq 0$). The probability of observing an unemployed seeker is then $Pr[h_i^* > 0. P_i^* > 0, E_i^* \leq 0]$

Participation and Employment indices can be incorporated into a general labour supply specification which we can now write as

$$h_i = \begin{cases} h_i^* & \text{for } h_i^* > 0, E_i^* > 0, P_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

Where h_i^*,P_i^* and E_i^* are defined as above and error terms are assumed to be mutually independent (of course, this can be tested).

The sample likelihood for the above model can now be written under the assumption of mutual independence of equation errors as a contribution of the the three groups.

The contribution of employed individuals to the likelihood function is

$$L_i^w = f(h_i^*) * Pr[E_i^* > 0, P_i^* > 0 \mid h_i^*], \tag{5}$$

where the first part term models individuals working hours and the latter term in brackets indicates that individual satisfies the equations 3 and 4 respectively.

The contribution of employment seeker is

$$L_i^{ES} = Pr[E_i^* \leq 0, P_i^* > 0], \tag{6}$$

because we know that she fullfills ($P_i^* > 0$ but she has not still found the job ($E_i^* \leq 0$).

Finally, the non–participants contribution is simply

$$L_i^{NP} = 1 - Pr[P_i^* > 0] = Pr[P_i^* \leq 0]. \tag{7}$$

So, our likelihood function which is to be maximized has the following form.

$$L = \prod_W f(h_i^*) * Pr[E_i^* > 0, P_i^* > 0 \mid h_i^*] \prod_{ES} Pr[E_i^* \leq 0, P_i^* > 0] \prod_{NP} Pr[P_i^* \leq 0], \tag{8}$$

where NP, ES anw W stands for the non–participants, seekers and workers respectively.[15]

[15] Alternative, and probably more intuitive way to define components in the sample likelihood function is following. The probability of observing a non–participant is

$$Pr_{NP} = Pr(h_i^* \leq 0) + Pr(h_i^* > 0, P_i^* \leq 0)$$

In the contex of labour supply, above likelihood spesification is valid under the linear budget constraint i.e. in the case when the individual faces only a constant marginal wage rate (or constant marginal tax rate). When the income tax schedule is progressive this is not the case. As Hausman[18] mentions there exists a potential for bias in parameter estimates if we do not take into account the income taxation. In the next subsection we discuss our choice of the labour supply function and after that we describe how nonlinear taxes can be incorporated to the likelihood function presented above.

5.1 Labour supply spesification

In many cases the micro–level data shows that there are quite a substantial amount of variety in individuals' behaviour. When this is the case we need to try to take this into account in the empirical work. But this is not an easy task to do and there are (almost) always a trade–off between statistical modelling and economic theory. For example, in this context sometimes our empirical models are used for policy analysis and this naturally requires simplicity and theory consistence. On the other hand data may require that the statistical model should be a flexible one. Stern[32] lists following general aspects which should be kept in mind when considering the choice of the functional form: Consistency with the utility theory, convenience in estimation, facility for incorporating in theoretical studies, ease of use in applied problems and flexibility in the type response it permits. In the older studies linear labour supply function (strictly speaking, linear respect to variables) was popular in empirical analysis, but it is quite obvious that this is an arguable choice as Blundell[5] mentions.[16]

where the first component identifies those who have no desire to work and the second component defines those individuals who want to work but are discouraged. The probability of observing an unemployment seeker is defined as

$$Pr_{ES} = Pr(h_i^* > 0, E_i^* \leq 0, P_i^* > 0)$$

The intuition behind this formulation is following: Individual wants to work and tries to find work but have not yeat find it. The probability of observing a worker is

$$Pr_w = Pr(h_i^* > 0, E_i^* > 0, P_i^* > 0)$$

[16] He gives the following clarifying example. Consider an increase in the tax rate in a linear tax system. This reduces the hourly wage rate among the employed individuals and reduces the pay–off to every extra hours worked. Individual who are free to enter to the labour markets are less likely to do so after the change in tax rate. Those in work are expected to reduce their hours. This latter argument is only a prediction from the

It would be desirable if the labour supply function allows a positive wage effect at zero hours and it would also allow wage effect to become negative for higher hours. One usually used functional form which fulfills these properties has the following form.

$$h^* = \alpha lnw + \beta(\frac{y}{w}) + \gamma z. \tag{9}$$

where w is the marginal hourly wage rate, y is the unearned income and z is the vector of demographic variables.[17] The precise form of the labour supply function is left as an empirical choice.[18]

5.2 Derivation of the likelihood function under non–linear taxes

Practically in all countries majority of governments tax and transfer schemes, like progressive income taxation and mean–tested income programs, create non–linear budget sets and discontinuities in the labour supply schedules. Crucial feature in the traditional empirical consumer demand analysis is that the consumer is assumed to purchase any desired quantity at a constant price subject to a budget constraint, ie. budget constraint is assumed to be linear. For example, in the case of progressive income tax system the "price" (i.e. wage) is not a constant, it varies with hours of work. This means that consumers face many different marginal wage rates.[19]

Hausman[18] argues that ignoring nonlinearities in empirical work is potential source of a misspesification. How non–linearities then should be taken into account? The simpliest solution is to use net wages as a regressor. But, it is apparent that net wage is correlated with the hours through the nonlin-

economic theory if those already in work are compensated for the loss in utility generated by the loss in leisure time. In the absence of this kind of compensation, the income effect generated by the loss in earned income may increase desired work effort. Result is the so–called backward bending labour supply curve. Blundell[5]

[17]Above functional form can be generalized allowing demographic variables (*demographic translation*) enter into the hours equation through the parameters α and β as suggested by Pollack and Wales[30]. See also Blundell and Meghir[9] for model spesification in the context of labour supply. Above funtional form or its genaralizations have been widely used in recent labour supply studies.

[18]One can show that the corresponding **indirect utility function** has the following form

$$V = \frac{w^{1+\beta}}{1+\beta}\left(\frac{y}{w}(1+\beta)^2 + \alpha lnw + \gamma - \frac{\alpha}{1+\beta}\right). \tag{10}$$

Direct utility function can be found from the indirect utility function using the standard procedure.

[19]Kinked budget constraints are also present in many other demand applications, like in the case of rationing and in a case of buying a durable goods.

ear tax (and possible some transfer schedule) system causing the endogeneity problem.[20]

The Maximum Likelihood method proposed by Burtless and Hausman[11] is where the whole budget constraint is taken into account in the estimation procedure. The general principle in this approach is that the consumer chooses her most preferred labour supply point on each budget segment, determine the corresponding utility of that choice and then chooses the one that yields *maximum maximorum* of utility across all segments (Hausman[18]). In labour supply context above means that the likelihood function takes into account the choice of hours over the entire exogenous tax schedule removing the endogeneity problem mentioned earlier.

Before going into the technical presentation, a few words about the so called Hausman–methodology is in place. MaCurdy, Green and Paarch[26] argued that a proper spesification of the log–likelihood over the region of a convex interior kink(s) is only possible if the estimated coefficients fulfill the condition that the compensated sustitution effect is positive. In other words, if the compensated substitution effect is not positive for some observations then they will have a negative probability of locating at the kink. In their study the data had a problems to meet that condition which means that the income coefficient must be constrained to be negative. As Blomquist[3] mentions, if the condition is legitimately met by the data then the estimation procedure imposes nothing on the data. There are no clear evidence that the data always fail to meet the condition mentioned above. Blundell, Duncan and Meghir[6] avoid the problem by deleting the observations near the convex kink[21] and then adding a selection term in a IV-regression to take the exclusion into account. This procedure is questionable if there are many kink points. First of all it means that we had to exclude considerable amount of observations from the analysis. Secondly, the definition of ”near the kink” is unclear.

Relating to above discussion it is evident that this approach reguires a lot from the data. Heckman[21] shows that in most of the cases the budget constraint is not accurately enough measured, but to estimate the model requires that the all kinks in the constraint for all individuals can be accu-

[20]This procedure actually means that we linearize the budget constraint around the observed hours. The linearized and original budget constraints yield the same optimum if the data are generated by utility maximization with globally convex preferences. Ordinary Least Squares can now be applied to the linearized data. However, using OLS can lead to inconsistent estimates, See e.g. Pudney[31], and therefore many researchers have applied Instrumental Variable method. Choice of instrument sets varies among the authors due to different exogeneity assumptions. See e.g. Blomquist[4] and MaCurdy et al[26].

[21]In their application there is only one kink point.

rately determined. Moffit[28] writes following;*"It is difficult given the data available to accurately determine the exact location of the kinks... because insufficient data are available on deductions, filing status, tax avoidance and so on. As a result, the location of the kinks assumed for the analysis may be incorrect".* As mentioned above, we argue that in this paper we can overcome these problems because of the data in hand and how we construct the budget constraints.

First assumption we will make is that preferences are non–stochastic i.e. any variation in preferences comes entirely from the observable personal attributes.[22] Let the tax system be the following. There are n linear tax segments and $n-1$ kink points where the marginal tax rates changes $(H_1, \ldots, H_{n-1})$. Budget set is assumed to be convex. Zero hours and maximal hours are denoted H_0 and H_n correspondingly. In the case of linear budget constraint it is easy to find the optimal labour supply because there are only one marginal tax rate, but in the case of progressive income taxation there are many different marginal rates, so we have to develop an search algorithm to find the optimal solution.[23]

Third, and in some sense most important assumption, is that we believe that observed labour market behaviour is the outcome of free rational choice subject to the constrains imposed by the income tax schedule. As Pudney[31] mentions, if one believes to the above mentioned assumption, there are good reasons to use statistical techniques that take complexity of the budget sets properly into account.

In presence of a convex budget set a quasiconcave utility function implies the existence of unique optimum given the maximization. Now, we start

[22] A model with random preferences means that individuals with same measured characteristics can make different choices. We derived the likelihood which takes this heterogeneity into account and also incorporates participation and job seeking desiscions. We assumed that the two randon terms were normally distributed. Our experience is that this model is extremely difficult and time consuming to estimate.

[23] Linearization method is sensitive for the two sources of bias. First is the simultaneity problem. Tax rate affects individuals work decision, but tax rate is also affected by the hours decision and the simple regression approach does not distinguish between these two causal links. At least one column of the matrix of observable characteristics is endogenous because it depends on the tax rate which depends on individuals's gross wage and this is in turn function of hours decision. The result is non-zero correlation between the stochastic term and the deterministic term (one or more components of it), thus the term $N^{-1}\sum \beta\epsilon$ has a non–zero probability limit.
The other source of the bias comes from the possible misspesification of the regression function. Linearization method assumes that if the individual is observed to be in one of the segments, then his or her optimal choice must also be in the **interior** point of that segment. But, the precence of the stochastic term means that observed and optimal choices can differ and thus lead to biased estimates.

by deriving desired labour supply on the first segment by substituting the corresponding net wage w_1 and unearned income y_i into the labour supply function. If desired hours h^* are less than equal to zero, then the individual is in the corner (zero hours). If desired hours lies between zero and H_1(second kink point. Note that the zero hours is the first kink point) we have the unique optimal desired hours for the individual. If desired hours are above the second kink point we move to the second segment. Net wage is now w_2 and virtual income[24] is y_2. Now, desired hours are determined using these figures and if the desired hours is less than H_1, then H_1 is the unique optimum. This can be seen applying the revealed preference argument. If $H_1 < h^* < H_2$, then unique optimum can be found from the second segment etc.

Like earlier we can write our general labour supply function as $h_i = h_i^*(w_i, y_i, z_i; \alpha, \beta, \gamma) + \epsilon_i$. In the statistical model we have to calculate the densities of of h_i and this requires evaluation of the maximum utilities received on each linear segment of the budget set just like above verbally described. More formally, we can now write the problem as

$$\begin{aligned} f(h_i) &= P[h_i = 0] + P[h_i > 0] * f(h_i \mid h_i > 0) + P[h_i = H_n] \\ &= P[\text{ at zero }] \\ &+ P[\text{ below kink 1 }] * f(h_i \mid \text{ below kink 1 }) \\ &+ P[\text{ at kink 1 }] \\ &+ P[\text{ above kink 1 }] * f(h_i \mid \text{ above kink 1 }) \\ &. \\ &. \\ &. \\ &+ P[\text{ at maximum }] \end{aligned} \tag{11}$$

So, we can think that **observed** hours are generated by the following generalized Tobit–model[25]

[24]Note that the y_1 should be called unearned or exogenous incomes and not virtual income. y_1 is the first vertical interception in the two dimensional space of labour supply and consumption of the composite good. This can be constructed directly from the data i.e. all components of it are observable. $y_i, (i > 1)$, should be called virtual incomes, because we do not find these figures from the data. These figures have to be calculated recursively.

[25]Note that we have dropped the subscript i.

$h = 0 \qquad\qquad\qquad \text{if } h^* + \varepsilon = 0$
$h = h^* + \varepsilon \qquad\quad \text{if } 0 < h^* + \varepsilon < H_n$
$h = H_n \qquad\qquad\quad \text{if } h^* + \varepsilon \geq H_n$

and the corresponding Likelihood Function can now be written as

$$L = \prod_{i \in A} \left[1 - \Phi \left(\frac{h^*}{\sigma_\varepsilon} \right) \right] \prod_{i \in B} \left[\frac{1}{\sigma_\varepsilon} \phi \left(\frac{h_i - h^*}{\sigma_\varepsilon} \right) \right] \prod_{i \in C} \left[1 - \Phi \left(\frac{H_n - h^*}{\sigma_\varepsilon} \right) \right]. \quad (12)$$

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 Standardized Normal Density Function and $\Phi(\cdot)$ is Cumulative Normal.

The first part of the likelihood function corresponce 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 correspons those whose observed hours are at maximum. At this stage we have to show how to find out the way to calculate the optimal supply of hours in the presence of kinked budget constraint.

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

$$H_{k-1} < h^*(w_k, y_k, z; \alpha, \beta, \gamma) < H_k \quad (13)$$

Intuition behind this calculation rule is following: after we have calculated the slope of the indiffrence curve from the direct utility function we replace consumption c ($= w_k h + y_k$) by individuals income (calculated for all the segments) and after that we equate the slope of the indefference curve and the marginal wage w_k corresponding that segment. The algorithm iterate as long as this condition is satisfied. If, for some individul we cannot find the solution we start to look if we can find it from some of the kink points.

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

$$h^*(w_k, y_k, z; \alpha, \beta, \gamma) \geq H_k \qquad \textit{and} \qquad h^*(w_{k+1}, y_{k+1}, z; \alpha, \beta, \gamma) \leq H_k. \quad (14)$$

An other way to express this condition is that the optimum can be found from the H_K, if the slope of the indifference curve is bigger or equal than w_{k+1} and the the slope is smaller or equal than w_k.

For the completenes we can show that optimum can be found from the zero hours $h = 0$ if

$$h^*(w_1, y_1, z; \alpha, \beta, \gamma) \leq 0 \tag{15}$$

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

$$h^*(w_n, y_n, z; \alpha, \beta, \gamma) \geq H_n. \tag{16}$$

The above formulation[26] shows how we can calculate the optimal hours under the progressive income taxation when we know the following aspects: tax schedule, hourly wage, exogenous incomes and the shape of the labour supply function (or correspondingly the shape of the utility function).

Despite that h^* can be calculated guite easily in the convex case, the maximization of the (log) likelihood function is not straightforward, because h^* is not a well behaved function respect to parameters. First of all, the log–likelihood function is not diffentiable everywhere (kink points) and secondly there can be parameter values where the function becomes very flat. The latter can become serious problem if there are not enough variation between the budget sets. In fact, it is very likely that the likelihood function is not diffentiable everywhere. Kendall and Stuart[24] have shown that ml-estimator is asymptotically consistent even if the likelihood function is not differentiable everywhere.

6 Results

Before our empirical results it is worth to say some words about the identification. We estimate for all years four equations: 1) wage equation, 2) participation index, 3) job availability index and 4) labour supply equation.

The reduced form equation for participation is non–parametrically identified from the sample split between non–participants on the one hand and job seekers and and workers on the other. Structural labour supply equation is conditional on both participation ($P_i^* > 0$) and employment ($E_i^* > 0$). In addition, it includes one endogenous variable: the net wage. This means that we need at least three variables to be excluded from the labour supply equations which enter into the participation equation, the employment equation and the wage equation. Wage equation is identified by excluding variables from it which the probit part includes. Finally, employment probability index includes variables which are not present in the labour supply function and the participation index.

[26] In practice the algorithm can be constructed in quite straightforwardly in the case of convex budget set. We estimated these models using Gauss and the programs are available from the author upon request.

Table 1: Group sizes in years 1987, 1989, 1991 and 1993

Sample separation.

Year	*Workers*	*Non–participants*	*Job seekers*	*N*
1987	1387(0.68)	585(0.29)	60(0.03)	2032
1989	1460(0.72)	543(0.26)	34(0.02)	2037
1991	1392(0.66)	590(0.28)	112(0.06)	2094
1993	894(0.61)	422(0.29)	150(0.10)	1466

The other relevant factor is group sizes. Following table shows group sizes to workers, non–participants and job seekers in all years.

As from the table can be seen, in year 1989 the amount of job seekers is very low and one should remember this when interpreting the results shown below. As expected, the proportion of job seekers increases sharply in the last two years.

The presentation of our results follows the sequental nature of the estimation strategy. The reduced–form wage equations are presented in the table 2. Wage equations for all years are estimated using Heckman correction. The wage equation is identified by including individual (more detailed information about the family composition), geographical and demand side variables in the probit equation. The selectivity effect is statistically significant in all other cases than in 1987.

Results confirm that 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 years of education. In years 1987 and 1989 age–wage profile differs from the years 1991 and 1993(recession years). Age effect is not statistically significant throughout the years, and so we do not speculate what might be behind this result. As it is expected, individuals who live in Helsinki metropolitan area and who are white–collar workers tend to have a higher wage rates. Interesting finding is that business–cycle variables (Funp/reg and Munp/reg) do not have any effect at all in years 1987, 1989 and 1991. In 1993 Female unemployment rate per region seems to have positive effect on wage rates and Male unemployment rate per region seems to have a reverse effect. We use the predicted wages from above equations in the sample likelihood functions and let us now turn to these results.

First, in table 3 we present the results from the employment probability indeces. We have to remember now that these equations are estimated (using

Table 2: Wage Equations

Wage Equations. Dependent variable: ln hwage.

Variables	*1987*	*1989*	*1991*	*1993*
Constant	3.40134	3.87261	3.34029	3.03162
	(0.0635)	(0.2215)	(0.2399)	(0.3432)
Age		-0.02702	0.01382	0.02503
		(0.0120)	(0.0127)	(0.0176)
Age2		0.00029	-0.00022	-0.00037
		(0.0001)	(0.0001)	(0.0002)
Age2534	-0.05020			
	(0.0348)			
Age3549	0.00643			
	(0.0298)			
Educ10	0.03643	0.05388	0.04302	0.02490
	(0.0215)	(0.0217)	(0.0230)	(0.0341)
Educ12	0.14532	0.13353	0.16770	0.11719
	(0.0265)	(0.0262)	(0.0230)	(0.0396)
Educ14	0.17418	0.21843	0.23726	0.18891
	(0.0409)	(0.0425)	(0.0443)	(0.0556)
Educ15	0.40510	0.43908	0.40439	0.32469
	(0.1748)	(0.0431)	(0.0418)	(0.0615)
Exp	0.00822	0.01283	0.01041	-0.00521
	(0.0041)	(0.0047)	(0.0052)	(0.0070)
Exp2	-0.00019	-0.00024	-0.00017	0.00016
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Tenure	0.15781	0.02840	0.01622	0.01265
	(0.0036)	(0.0030)	(0.0036)	(0.0049)
Tenure2	-0.00016	-0.00059	-0.00024	-0.00015
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Pjob	0.06379	0.09487	0.07614	0.11159
	(0.0273)	(0.0256)	(0.0270)	(0.0390)
Husb	-0.00431	0.04117	-0.09703	-0.03655
	(0.0270)	(0.0261)	(0.0228)	(0.0325)
Stat	0.07205	0.10338	0.13826	0.14542
	(0.0277)	(0.0241)	(0.0286)	(0.0291)
Socio	0.21142	0.23198	0.20780	0.34013
	(0.0343)	(0.0330)	(0.0322)	(0.0491)
Nchild	0.01465	0.03680	0.01853	0.01538
	(0.0090)	(0.0090)	(0.0094)	(0.0138)
Funp/reg				0.02032
				(0.0085)
Munp/reg				-0.01549
				(0.0068)
South	0.11558	0.13836	0.08662	0.09764
	(0.0204)	(0.0343)	(0.0203)	(0.0343)
Ln L	-1296.47	-1216.99	-1305.24	-1211.82

NOTE:The selection index is a function of the individual, geographical and demand side variables. The selectivity effect was statistically significant in all other years than in 1987.

standard probit technique) separately from the participation and labour supply equations, because we asumed mutual independence between the stochastic error terms. We do not report any statistical tests conserning the assumption made above.[27] This should be kept in mind when evaluating the results. As we mentioned earlier, in 1989 we have only 34 employment seekers, and that can be seen also from the the value of the log–likelihood. So, some doubts should be made when looking these results.

In all years increase in age decreases the probability of finding a job. Education, as expected, increses this probability and in year 1987 Educ15 predicts this probability perfectly. In years 1987 and 1989 increase in the number of children increase the probability of finding a job but the effect is reverse in recesion years. Again, being a white–collar worker increase the probability as it is also the case when husband is working.

Let us now turn into the participation equations results. These equations are estimated simultaneously with the labour supply functions, which also takes into account nonlinear taxes. Chosen procedure means that the likelihood function is quite difficult to estimate and it is very sensitive to the labour supply spesification. We estimated the model with different starting values and we also used both numerical and analytical derivatives.

The properties of the parameter estimates are broadly plausible. Age affects negatively on labour supply participation. Again, education has strong positive effect of increasing the participation likelihood. Our results also show clearly that decision to participate in the labour market should be modelled as separate from the choice of hours of work.

Our experience shows that incorporating separate participation equation into the sample likelihood lessen the importance of certain parameters in the labour supply equation, like e.g. child dummies and especially education. This suggests that the importance of the children and education is more on the probability of participation rather than the decision of how many hours to work.

Number of dependent children in the family seem to increase the wife's probability to participate in years 1987 and 1989. In the case of years 1991 and 1993 the effect is reverse.

It may be the case that in normal economic conditions market wage exceeds the reservation wage and vice versa (coefficient for the years 1991 and 1993 are not statistically significants). It is very difficult to speculate what is behind this result and it needs further

[27] Basically, there are two ways to proceed in this area. Either to follow technique (Generalised residuals approach) developed Bera et al.[2], Gourieroux et al.[14], Lee[25] or technigue (χ^2 goodness of fit statistic) proposed in econometrics by Heckman[19] and Andrews[1]. Both approaches are not straightforward to implement and progressing in this area is on the author's research agenda.

Table 3: Employment Probability Indeces

Employment Probability Indeces.				
Variables	*1987*	*1989*	*1991*	*1993*
Constant	1.65030	1.19469	1.29975	1.31268
	(0.6300)	(0.6820)	(0.5647)	(0.5711)
Age		-0.00419	-0.00007	-0.00360
		(0.0107)	(0.0078)	(0.0070)
Age2534	0.07978			
	(0.2046			
Age3549	0.13830			
	(0.1816)			
Educ10	-0.16858	0.13456	0.20685	0.09197
	(0.1501)	(0.1244)	(0.1246)	(0.1173)
Educ12	0.15420	0.35687	0.39379	0.39187
	(0.2208)	(0.3078)	(0.1714)	(0.1643)
Educ14	0.29893	0.15231	0.85440	0.24191
	(0.4292)	(0.4290)	(0.4049)	(0.1218)
Educ15		-0.27210	1.42638	0.47341
		(0.3057)	(0.5547)	(0.2272)
Husb	0.10374	0.80000	0.43950	0.28706
	(0.1834)	(0.1758)	(0.1202)	(0.1131)
Stat	0.35524	0.19414	0.22002	0.29657
	(0.1354)	(0.1676)	(0.1140)	(0.1061)
Nchild	0.12898	0.07966	-0.04892	-0.09321
	(0.0796)	(0.0893)	(0.0564)	(0.0533)
Funp/reg	-0.06212	-0.05873	-0.06403	-0.02866
	0.05722	(0.0765)	(0.0576)	(0.0245)
South	0.61506	0.31001	0.0770	-0.04748
	(0.3452)	(0.4132)	(0.2650)	(0.1838)
Exog	-0.0001	-2.36e-06	-2.17e-06	4.26e-06
	(2.7e-06)	(3.04e-06)	(8.29e-06)	(4.56e-06)
Unskill	-1.13691	-0.44377	-2.66028	-0.58443
	(0.2211)	0.3512	0.4206	(0.1735)
Ln L	-213.40	-142.69	-326.70	-419.48

Table 4: Participation Probabily Indeces

Participation Probability Indeces.

Variables	*1987*	*1989*	*1991*	*1993*
Constant	1.8337	1.8394	1.9199	1.7737
	(0.2041)	(0.2441)	(0.2723)	(0.2100)
Age		-0.0177	-0.0216	-0.2121
		(0.0047)	(0.0052)	(0.1645)
Age5059	-0.5578			
	(0.0881)			
Educ10	0.0540	0.1124	0.1655	-0.0812
	(0.0774)	(0.0813)	(0.0796)	(0.0912)
Educ12	0.0838	0.0906	0.2347	0.0184
	(0.0921)	(0.9730)	(0.0938)	(0.1096)
Educ14	0.0809	-0.1214	0.1873	0.2823
	(0.0137)	(0.1444)	(0.1497)	(0.1608)
Educ15	0.4127	0.1748	0.4932	0.2852
	(0.1552)	(0.1351)	(0.1285)	(0.1489)
Stat	1.1061	0.19414	1.24532	1.1314
	(0.0714)	(0.1676)	(0.0723)	(0.0848)
Nchild	0.0127	0.07966	-0.0156	-0.0517
	(0.0422)	(0.0893)	(0.0454)	(0.0489)
Cdum1	-0.7133	-0.4731	-1.2050	-0.7184
	(0.1283)	(0.1415)	(0.1672)	(0.1390)
Cdum2	0.2214	0.5786	0.2098	0.5389
	(0.1518)	(0.1635)	(0.1888)	(0.1802)
Cdum3	0.3065	0.7282	0.1526	0.3435
	(0.1501)	(0.1592)	(0.1734)	(0.1759)
Cdum4	0.3324	0.7046	0.2693	0.3685
	(0.1098)	(0.1170)	(0.1174)	(0.1186)
Funp/reg	0.0332	0.0495	-0.1635	-0.0427
	(0.0396)	(0.0757)	(0.0511)	(0.0281)
Munp/reg	-0.0491	-0.0731	0.0523	0.0220
	(0.0215)	(0.0766)	(0.0292)	(0.0226)
South	-0.0235	-0.1329	-0.4633	-0.1384
	(0.1403)	(0.1322)	(0.1702)	(0.1390)

investigation. For all years, it is clear that the presence of the young children reduces the participation probability. Surprisingly, coefficient for the south is negative in all years, but the parameters are inaccurately estimated for all other years than 1991. Business–cycle variables change the signs but they are also quite inaccurately estimated. It still seem to be that during the recession females participation likelihood decreases if female unemployment rate per region goes up but if male unemployment rate goes up it will increase the participation likelihood.

Let us next turn to the labour supply equations. Now, it is crucial to realise that these results are **conditional** to the employment and participation equations. In labour supply analysis the two main variables we (usually) are interested in are the net wage and the unearned income. Obviously, this is because we are keen on knowing the income and wage elasticities. Table 5 shows the results. Results are based on the functional form presented in equation 9.[28]

Our results indicate that in years 1987 and 1989 individuals are more or less on their labour supply curve. Net wage coefficient has a positive sign as theory predicts, and coefficients for those years are presicely estimated. Results for the unearned income variable also satisfy theoretical expectations. Increase in exogenous (unearned) incomes decreseas the optimal/desired hours of of work. It should be kept on mind that we actually estimate the parameters in the individual's utility function. For the recession years the results are mixed, as expected. It is very likely that individuals cannot be on their labour supply curve when the demand side restrictions are strong. Still, for all years compensated wage elasticities are positive. In 1987 and 1989 compensated wage elasticities are 0.07 and 0.045 correspondingly.

Our compensated wage elasticity estimates are somewhat lower than usually got for females in other labour supply studies. The only relevant comparison for our results is the study conducted by Ilmakunnas[23]. In her study, the estimated compensated wage elasticity was found to be as high as 0.29. She used the same 1987 data set and practically the same sample selection procedure. The diffrence is that she used linear labour supply function (which is a debatable choice) and did not take into account different labour market groups in her sample likelihood. Our opinion is that the elasticity got in this study is more plausible in it's magnitude. Considering the structure of Finnish labour markets it is hard to believe that hours are very sensitive to the changes in net wages.

In years 1991 and 1993 the elasticities are 0.048 and 0.03 correspondingly, but these figures are not presicely estimated. The results show that the same estimation approach (and probably the same labour supply spesification either) is not valid through the years and it can be very misleading to use elasticities obtained from one cross–section study. There is a lot of work to be done and the next natural step is to move to the dynamic labour supply models. This is not a straightforward task to do, because if one wants to take e.g. taxation into account, the sample likelihood function will become a very tedious one. Another problem is that in Finland there are no suitable data sets available for proper dynamic empirical modelling.

Estimation results for the other covariates in our labour supply functions are more or less plausible. Presence of young kids decrease the desired working hours in all years (but not as strongly when the model is estimated without latent index function), but in 1989, 1991 and 1993 also the elderly children decreases the desired labour supply compared to the childless female.

[28] We also estimated the sample likelihood functions for all years using the linear logarithmic functional form. Results are on the line shown here and are available from the author.

Table 5: Labour Supply Functions

Labour Supply Functions

Variables	*1987*	*1989*	*1991*	*1993*
Constant	1.8409	1.8400	1.8409	1.8512
	(0.2222)	(0.2425)	(0.3187)	(0.4101)
Ln W	0.1112	0.1098	0.0199	-0.0009
	(0.0586)	(0.0112)	(0.0188)	(0.0028)
Exog. inc	-0.0031	-0.0047	-0.0079	-0.0059
	(0.0012)	(0.0058)	(0.0022)	(0.0039)
Cdum1	-0.2147	-0.3658	-0.3825	-0.2892
	(0.0626)	(0.0619)	(0.0816)	(0.0763)
Cdum2	-0.0327	-0.1542	-0.0807	-0.2495
	(0.0638)	(0.0635)	(0.0750)	(0.0813)
Cdum3	0.0691	-0.0890	-0.1801	-0.1835
	(0.0648)	(0.0633)	(0.0683)	(0.0844)
Cdum4	0.1269	-0.0469	-0.0497	-0.0257
	(0.0494)	(0.0494)	(0.0501)	(0.0591)
Age		0.1278	0.0374	0.0784
		(0.0169)	(0.0191)	(0.0225)
Age*Age		-0.1746	-0.06128	-0.0868
		(0.0127)	(0.2234)	(0.0263)
Age2534	0.5608			
	(0.0562)			
Age3549	0.2307			
	(0.0501)			
σ	0.7632	0.5636	0.7756	0.7930
	(0.1009)	(0.0101)	(0.1266)	(0.2659)

**NOTE: Dependent variable is Hours/1000.
Unearned income variable is Exog.inc/10000.
Square of age variable is Age*Age/100.**

7 Conclusions

In this paper we have estimated a model of labour supply which allows for job seeking, discouraged workers and and nonlinear taxes. We use flexible functional form to estimate labour supply behaviour unlike in other earlier made Finnish studies. We find that female labour supply is not so elastic as earlier results might suggest. We apply our model to the four different cross–section data sets. In years 1987 and 1989 our model gives results that fullfills theory predictions and mean elasticities are 0.07 and 0.045 respectively. These elasticities are clearly lower than previus results and we argue that our estimation procedure is more complete way to estimate labour supply functions from the cross–section data sets.

We also find that in recession years, 1991 and 1993, our model is inadequate to take account the fact that individuals are not on their labour supply curve. The first lesson to learn from our simple exercise is that it is misleading to use mean elasticities estimated from the cross–section studies in simulation models when the macroeconomic environment has changed. We need to develop empirical dynamic labour supply models with (nonlinear) taxation before we can use our elasticities in dynamic simulations. This paper has no attemp to give robust estimates for elasticities over time. In addition, we should not use mean elasticities, cause it is clear that different income groups react differently to the tax reforms.

¿From the practical point of view our results do not show that decreasing marginal tax rates will have a significant effect on amount of labour supply for those already in the labour force. This is not to say, that decrease in marginal tax rates do not have any effect at all. On the contrary, reforms may have a great effect on labour supply participation rates. It seems to be that in Finland the discrete choise between labour market status is the most important one and different kind of welfare reforms should be formed to courage individuals to change their labour market regime.

The model we have estimated are so called single–error–term model. This is a someway implausible way to approach the problem, because we assume that all variance in hours conditional on covariates is measurement error. In other words, our model assumes that the estimated parameters are identical for all observations, so there is only one utility maximizing choice in the population. On the contrary, it is quite natural to assume that some of the observed distribution of the observations over the constraint is because there are differences in the parameters of the utility function. Introducing the second stochastic term into our sample likelihood function is the natural step to proceed. Another field which has got too little attention in this paper is statistical testing. As we mentioned earlier, there are no standard tests available, and we have to program the modified tests ourselves. This is clearly a very important field to take under the research agenda.

References

[1] D.W.K. Andrews. The Chi–Square Diagnostic Tests for Econometric Models: Introduction and Applications. *Journal of Econometrics*, 37:135–156, 1988.

[2] A.K. Bera, C. M. Jarque, and L.F. Lee. Testing for the Normality Assumptions in Limited Dependent Variable. *International Economic Review*, 25:563–578, 1984.

[3] S. Blomquist. Restrictions in Labour Supply Estimation: Is the Macurdy Critique Correct? *Economics Letters*, 47:229–235, 1995.

[4] S. Blomquist. Estimation Methods for Male Labor Supply Functions: How to Take Account of Nonlinear Taxes. *Journal of Econometrics*, 70:383–405, 1996.

[5] R.W Blundell. Labour Supply and Taxation: A Survey. *Fiscal Studies*, 13:15–40, 1992.

[6] R.W. Blundell, A. Duncan, and C. Meghir. Estimating Labour Supply Responses Using Tax Reforms. *The Institut for Fiscal Studies*, W95/7, 1995.

[7] R.W. Blundell, J. Ham, and C. Meghir. Unemployment and Female Labour Supply. *Economic Journal*, 97:44–64, 1987.

[8] R.W. Blundell, J. Ham, and C. Meghir. Unemployment, Discouraged Workers and Female Labour Supply. *Mimeo, University College London*, 1995.

[9] R.W Blundell and C Meghir. Selection Criteria for a Microeconometric Model of a Labour Supply. *Journal of Applied Econometrics*, 1:55–80, 1986.

[10] K. Burgett and D.T. Mortensen. *Labour Supply Under Uncertainty*. JAI Press, Connecticut, 1978. In Research in Labour Economics. Edited by R.G. Ehrenberg.

[11] G. Burtless and J. A. Hausman. The Effect of Taxation on Labour Supply. *Journal of Political Economy*, 86:1103–1130, 1978.

[12] J.C. Cragg. Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durables Goods. *Econometrica*, 39:829–844, 1971.

[13] A. Duncan. Labour Supply Decision and Non–Convex Budget Sets: The Case of National Insurance Contributions in the UK. *The Institut for Fiscal Studies*, W90/7, 1990.

[14] C. Gourieroux, A. Monfort, and A. Renault, E. Trognon. Generalized Residuals. *Journal of Econometrics*, 34:5–32, 1987.

[15] S Grossman and O. Hart. Implicit Contracts, Moral Hazard and Unemployment. *American Economic Review Paers and Proceedings*, 71:301–307, 1981.

[16] J.C. Ham. Estimation of a Labour with Censoring Due to Unemployment and Underemployment. *Review of Economic Studies*, 49:335–354, 1982.

[17] J.A. Hausman. The Effect of Wages, Taxes and Fixed Costs on Women's Labor Force Participation. *Journal of Public Economics*, 14:161–194, 1980.

[18] J.A. Hausman. The Econometrics of Nonlinear Budget Sets. *Econometrica*, 53:1255–1282, 1985.

[19] J.J. Heckman. The Chi–Squared Goodness of Fit Statistic for Models with Parameter Estimated from Microdata. *Econometrica*, 52:1543–1547, 1978.

[20] J.J. Heckman. Sample Selection Bias as a Spesification Error. *Econometrica*, 42:153–162, 1979.

[21] J.J. Heckmann. *Comment*. University of Chicago Press, Chicago, 1983. In Behavioral Simulation Methods in Tax Policy Analysis. Edited by M. Feldstein.

[22] D. Hum and W. Simpson. Labour Supply Estimation and Public Policy. *Journal of Economic Surveys*, 8:57–78, 1994.

[23] S. Ilmakunnas. Income Taxation, Hours Restrictions and Labour Supply. *Labour Institute for Economic Research*, 113:1–41, 1992.

[24] M.G. Kendall and A. Stuart. *The Advanced Theory of Statistics, Vol 2.* Griffin, 1974.

[25] L-F. Lee. Tests for the Bivariate Normal Distribution in Econometric Models wit Selectivity. *Econometrica*, 52:843–864, 1984.

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

[27] E. Malinvaud. *The Theory of Unemployment Reconsidered.* New York, Halstead, 1977.

[28] R. Moffitt. The Econometrics of Kinked Budget Constraints. *Journal of Economic Perspectives*, 4:119–139, 1990.

[29] T.A. Mroz. The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions. *Econometrica*, 55:765–800, 1987.

[30] R.A Pollack and T.J Wales. Demographic Variables in Demand Analysis. *Econometrica*, 49:1533–1554, 1981.

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

[32] N. Stern. *On the Spesification of Labour Supply Functions*. Cambridge University Press, Cambridge, 1986. In Unemployment, Search and Labour Supply. Edited by R.W. Blundell and I. Walker.

Appendix 1

union=1, if the respondent is a member of an union
age=Age of the respondent
age2= Age squared
age2534= 1, if the respondent's age is between 25 and 34. Otherwise zero.
age3549= 1, if the respondent's age is between 35 and 49. Otherwise zero.
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 university degree from 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= Exp. 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.
exo= Unearned income.
exo+hnet= Unearned income+husband's net incomes.
Helsinki= 1, if respondent lives in the Helsinki metropolitan area (in 1987 and 1989 in administrative district Uusimaa).
south=South Finland.
west=West Finland.
east=East Finland.
middle=Middle Finland.
north=North Finland.
lapl=Lapland.
noccu=Occupation not specified.
manufac=Manufactory.
retcat=Retail and Catering.
transp=Transportation.
bankins=Banking and Insurance.
pubsec=Public sector administrative work.
educres=Education and Research.
healtsoc=Health and social services.
priserv=Other private sector services.

All money term variables are expressed in 1993 money.

Appendix 2

State income tax schedule — 1987

taxable income	*tax at lower bound*	*margin. tax rate*
15 600 – 21 800	10	6
21 800 – 27 000	382	13
27 000 – 32 200	1058	19
32 200 – 41 600	2046	23
41 600 – 53 000	4208	28
53 000 – 76 000	7400	29
76 000 – 102 000	14 070	33
102 000 – 159 000	22 650	38
159 000 – 265 000	44 310	45
265 000 – 475 000	92 010	50
475 000 –	197 010	51

State income tax schedule — 1989		
taxable income	*tax at lower bound*	*margin. tax rate*
36 000 – 51 000	50	11
51 000 – 63 000	1700	21
63 000 – 89 000	4220	26
89 000 – 140 000	10 980	32
140 000 – 250 000	27 300	37
250 000 –	68 000	44

State income tax schedule — 1991 and 1993		
taxable income	*tax at lower bound*	*margin. tax rate*
40 000 – 56 000	50	7
56 000 – 70 000	1170	17
70 000 – 98 000	3550	21
98 000 – 154 000	9430	27
154 000 – 275 000	24 550	33
275 000 –	64 480	39

Appendix 3

Among participants: Descriptive Statistics – 1987,1989,1991,1993

Variables	*1987*	*1989*	*1991*	*1993*
Hours	1962.42(503.63)	1855.58(560.10)	1876.37(570)	1855(569.06)
union	0.75(0.43)	0.71(0.45)	74.14(0.44)	0.80(0.40)
age2534	0.29(0.45)			
age3549	0.54(0.49)			
age5059	0.17(0.38)			
age		41.29(8.47)	41.88(8.34)	42.00(8.34)
educ10	0.28(0.45)	0.30(0.46)	0.32(0.46)	0.30(0.46)
educ12	0.17(0.38)	0.18(0.39)	0.21(0.40)	0.21(0.41)
educ14	0.06(0.24)	0.05(0.21)	0.06(0.23)	0.08(0.27)
educ15	0.06(0.25)	0.07(0.26)	0.11(0.31)	0.13(0.34)
cdum1	0.14(0.35)	0.13(0.34)	0.10(0.30)	0.12(0.32)
cdum2	0.13(0.33)	0.12(0.33)	0.12(0.32)	0.17(0.37)
cdum3	0.12(0.32)	0.12(0.32)	0.12(0.33)	0.10(0.29)
cdum4	0.26(0.44)	0.25(0.44)	0.29(0.46)	0.25(0.43)
workexp	18.26(8.89)	19.50(9.32)	19.83(8.97)	19.80(9.04)
jobdur	8.51(8.14)	8.60(8.34)	8.78(8.42)	9.40(8.43)
permjob	0.79(0.40)	0.77(0.41)	0.79(0.40)	0.78(0.41)
phusb	0.88(0.32)	0.86(0.33)	0.81(0.39)	0.77(0.41)
hwage	50.56(24.80)	57.03(55.89)	60.47(32.28)	62.68(50.27)
exo	5022(7326)	6387(16217)	4095(8385)	7463(13477)
exo+hnet	80707(46769)	97786(70744)	96952(54864)	92403(54373)
south	0.25(0.44)	0.25(0.44)	0.25(0.43)	0.26(0.44)
west	0.16(0.36)	0.16(0.36)	0.15(0.36)	0.14(0.34)
east	0.18(0.38)	0.19(0.39)	0.22(0.42)	0.19(0.39)
middle	0.16(0.36)	0.14(0.34)	0.13(0.34)	0.14(0.34)
north	0.18(0.39)	0.18(0.39)	0.18(0.38)	0.19(0.39)
lapl	0.07(0.26)	0.08(0.27)	0.07(0.25)	0.08(0.27)

Among Unemployed Seekers: Descriptive Statistics – 1987,1989,1991,1993				
Variables	*1987*	*1989*	*1991*	*1993*
age		45.0(11.0)	43.0(9.22)	43.18(9.12)
age2534	0.35(0.48)			
age3549	0.38(0.49)			
age5059	0.27(0.45)			
cdum1	0.17(0.38)	0.15(0.36)	0.16(0.37)	0.13(0.33)
cdum2	0.10(0.30)	0.09(0.29)	0.10(0.30)	0.16(0.36)
cdum3	0.12(0.32)	0.06(0.24)	0.11(0.31)	0.07(0.25)
cdum4	0.22(0.42)	0.18(0.38)	0.21(0.41)	0.24(0.43)
workexp	13.7(9.2)	18.26(10.98)	17.50(10.53)	17.62
phusb	0.78(0.42)	0.5(0.5)	0.58(0.49)	0.65(0.47)
educ10	0.35(0.48)	0.26(0.45)	0.35(0.48)	0.41(0.49)
educ12	0.12(0.32)	0.06(0.24)	0.12(0.32)	0.10(0.30)
educ14	0.02(0.13)	0.03(0.17)	0.009(0.09)	0.04(0.20)
educ15	0(0)	0.09(0.29)	0.009(0.09)	0.03(0.18)
exo	5877(8315)	6352(13314)	4588(8501)	6594(11013)
exo+hnet	64475(43884)	75087(43905)	83957(61412)	89602(58660
south	0.03(0.18)	0.09(0.29)	0.13(0.34)	0.19(0.39)
west	0.07(0.25)	0.18(0.39)	0.20(0.40)	0.15(0.35)
east	0.32(0.47)	0.30(0.46)	0.25(0.43)	0.16(0.37)
middle	0.20(0.40)	0.15(0.36)	0.13(0.33)	0.19(0.39)
north	0.30(0.0.46)	0.18(0.38)	0.20(0.40)	0.18(0.39)
lapl	0.08(0.28)	0.12(0.32)	0.09(0.289	0.13(0.33)

Among Non–Participants: Descriptive Statistics – 1987,1989,1991,1993				
Variables	*1987*	*1989*	*1991*	*1993*
age		43.00(11.68)	42.11(11.26)	42.17(11.16)
age2534	0.30(0.46)			
age3549	0.34(0.47)			
age5059	0.37(0.48)			
cdum1	0.29(0.45)	0.32(0.47)	0.38(0.48)	0.36(0.48)
cdum2	0.07(0.25)	0.06(0.23)	0.05(0.21)	0.04(0.20)
cdum3	0.05(0.22)	0.05(0.22)	0.05(0.22)	0.05(0.21)
cdum4	0.12(0.33)	0.10(0.30)	0.12(0.32)	0.13(0.33)
phusb	0.72(0.45)	0.69(0.46)	0.69(0.46)	0.64(0.48)
exo	6302(11030)	8112(12916)	8802(13316)	12093(16492)
exo+hnet	73711(53360)	89965(54160)	97097(52218)	95408(106072)
educ10	0.25(0.44)	0.26(0.44)	0.31(0.46)	0.37(0.48)
educ12	0.16(0.36)	0.16(0.37)	0.18(0.38)	0.20(0.40)
educ14	0.05(0.22)	0.06(0.24)	0.05(0.22)	0.05(0.22)
educ15	0.03(0.18)	0.06(0.24)	0.07(0.25)	0.06(0.25)
south	0.20(0.40)	0.22(0.42)	0.21(0.41)	0.22(0.41)
west	0.14(0.34)	0.42(0.33)	0.14(0.34)	0.13(0.34)
east	0.26(0.43)	0.24(0.42)	0.23(0.42)	0.22(0.41)
middle	0.13(0.34)	0.15(0.36)	0.16(0.36)	0.17(0.38)
north	0.21(0.44)	0.18(0.38)	0.17(0.37)	0.17(0.38)
lapl	0.07(0.25)	0.08(0.27)	0.08(0.28)	0.09(0.29)

Appendix 3

1987

Cross–Tabulation of Female Hours vs. Number of Children								
n. of kids	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	304	29	31	151	216	51	23	805
1	144	20	25	127	132	56	12	516
2	114	23	22	141	123	55	11	489
3	56	14	11	25	29	20	1	156
4+	25	4	4	6	8	8	1	57
Total	644	90	93	450	508	190	48	2023

Cross-Tabulation of Female Hours vs. Status								
Status	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	493	32	42	85	245	73	24	994
1	151	58	51	365	263	117	24	1029
Total	644	90	93	450	508	190	48	2023
NOTE: 1=White–collar worker, 0=Blue–collar worker								

Cross-Tabulation of Female Hours vs. Union Membership

Union	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	467	45	30	104	68	59	26	799
1	117	45	63	346	440	131	22	1224
Total	644	90	93	450	508	190	48	2023

NOTE: 1=Union member, 0=Not a member

Cross-Tabulation of Female Hours vs. Level of Education

Educ.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
1	323	34	43	164	241	75	27	907
2	268	40	42	214	216	95	18	893
3	53	16	8	72	51	20	3	223
Total	644	90	93	450	508	190	48	2023

NOTE: 1= Less than 9 years of education, 2=9–13 years of education
3= More than 13 years of education.

1989:

Cross–Tabulation of Female Hours vs. Number of Children

n. of kids	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	264	38	40	223	146	67	16	794
1	97	25	42	171	83	47	10	475
2	132	36	47	187	96	42	9	549
3	59	8	12	42	17	23	2	103
4+	25	2	4	10	8	7	0	56
Total	577	109	145	633	350	186	37	2037

Cross-Tabulation of Female Hours vs. Status

Status	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	437	55	53	171	176	90	25	1007
1	140	54	92	412	174	96	12	1030
Total	577	109	145	633	350	186	37	2037

NOTE: 1=White-collar worker, 0=Blue-Collar worker

Cross-Tabulation of Female Hours vs. Union Membership

Union	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	448	50	51	135	77	77	26	864
1	129	59	94	498	273	109	11	1173
Total	577	109	145	633	350	186	37	2037

NOTE: 1=Union member, 0=Not a member

Educ.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
1	263	45	54	219	162	81	19	843
2	243	54	70	327	155	85	17	951
3	71	10	21	87	33	20	1	243
Total	577	109	145	633	350	186	37	2037

NOTE: 1= Less than 9 years of education, 2=9–13 years of education
3= More than 13 years of education.

Cross-Tabulation of Female Hours vs. Occupation

Occup.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	341	2	4	22	13	3	1	386
1	53	33	33	124	101	60	8	412
2	43	20	20	116	47	33	13	295
3	12	6	11	24	10	7	2	72
4	23	10	5	68	21	6	2	135
5	6	5	4	50	13	2	0	80
6	15	8	26	75	37	10	2	173
7	65	11	27	120	84	50	7	361
8	22	14	15	34	24	15	2	126
Total	577	109	145	633	350	186	37	2037

NOTE: 0=unspecified, 1=Manufacturing, 2= Retail and Catering, 3=Transport industry, 4= Banking and Insurance, 5=Public Sector, 6= Education and Research, 7= Health and Social Services, 8=Other Private Sector Services

1991:

Cross–Tabulation of Female Hours vs. Number of Children

n. of kids	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	288	25	49	258	91	68	19	798
1	139	15	39	209	68	35	11	516
2	158	29	38	200	52	32	13	522
3	71	12	11	51	21	15	3	184
4+	46	2	4	11	4	6	1	76
Total	702	83	141	729	439	156	47	2094

Cross-Tabulation of Female Hours vs. Status

Status	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	543	34	55	166	105	54	36	993
1	159	49	86	563	131	102	11	1101
Total	702	83	141	729	439	156	47	2094

NOTE: 1=White-collar worker, 0=Blue-Collar worker

Cross-Tabulation of Female Hours vs. Union Membership

Union	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	584	35	49	122	53	64	37	944
1	118	48	92	607	183	92	10	1150
Total	702	83	141	729	439	156	47	2094

NOTE: 1=Union member, 0=Not a member

Cross-Tabulation of Female Hours vs. Level of Education

Educ.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
1	289	29	48	197	90	52	20	725
2	341	45	66	408	109	75	24	1068
3	72	9	27	124	37	29	3	301
Total	702	83	141	729	439	156	47	2094

NOTE: 1= Less than 9 years of education, 2=9–13 years of education 3= More than 13 years of education.

Cross-Tabulation of Female Hours vs. Occupation

Occup.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	423	0	0	2	0	0	0	425
1	85	15	23	117	76	32	22	370
2	53	12	28	114	35	26	15	283
3	12	5	3	22	6	4	0	52
4	26	7	13	80	11	8	2	147
5	13	1	9	54	6	4	0	87
6	16	7	28	91	25	15	2	184
7	58	20	24	195	53	53	4	407
8	16	16	13	54	24	14	2	139
Total	702	83	141	729	439	156	47	2094

NOTE: 0=unspecified, 1=Manufacturing, 2= Retail and Catering, 3=Transport industry, 4= Banking and Insurance, 5=Public Sector, 6= Education and Research, 7= Health and Social Services, 8=Other Private Sector Services

1993:

Cross–Tabulation of Female Hours vs. Number of Children

n. of kids	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	226	24	37	171	57	43	10	568
1	108	13	18	107	35	29	10	326
2	141	20	21	140	43	29	9	403
3	69	5	10	32	13	16	1	140
4+	22	1	2	7	0	1	2	35
Total	566	63	88	457	148	112	32	1466

Cross-Tabulation of Female Hours vs. Status

Status	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	411	30	24	102	60	44	21	692
1	155	33	64	355	88	68	11	774
Total	566	63	88	457	148	112	32	1466

NOTE: 1=White-collar worker, 0=Blue-Collar worker

Cross-Tabulation of Female Hours vs. Union Membership

Union	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	480	26	17	49	27	34	22	655
1	86	37	71	108	121	78	10	811
Total	566	63	88	457	148	112	32	1466

NOTE: 1=Union member, 0=Not a member

Cross-Tabulation of Female Hours vs. Level of Education

Educ.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
1	196	23	26	134	47	28	14	468
2	312	28	45	234	75	61	13	768
3	58	12	17	89	26	23	5	230
Total	566	63	88	457	148	112	32	1466

NOTE: 1= Less than 9 years of education, 2=9–13 years of education 3= More than 13 years of education.

Cross-Tabulation of Female Hours vs. Occupation

Occup.	0	1–1000	1001–1500	1501–2000	2001–2200	2201–3000	3000+	Total
0	272	5	2	24	7	3	2	315
1	61	14	10	63	49	34	15	246
2	58	8	15	63	16	8	3	171
3	12	6	4	17	5	7	1	52
4	25	6	5	51	11	8	2	108
5	14	1	5	28	6	2	0	56
6	26	6	18	49	12	13	1	125
7	75	16	26	131	28	31	5	300
8	23	7	9	31	14	6	3	93
Total	566	63	88	457	148	112	32	1466

NOTE: 0=unspecified, 1=Manufacturing, 2= Retail and Catering, 3=Transport industry, 4= Banking and Insurance, 5=Public Sector, 6= Education and Research, 7= Health and Social Services, 8=Other Private Sector Services