Luxembourg Employment Study Working Paper Series

Working Paper No. 4

LIS/LES Wage Imputations

Hermann Buslei

January 1997

Luxembourg Income Study (LIS), asbl

LIS/LES Wage Imputations

Hermann Buslei

ZEW, Mannheim

January, 1997

Abstract
The Luxembourg Employment Study (LES) data do not contain any information about wages and income either at the household or at the individual level except for a few countries. But with respect to other work-related variables, it is a rich data set and the number of observations is generally much higher than in other household data sets. For many questions that could be investigated using LES data, it might be useful to have additional information about the personal wage rate or household income. This additional information might be imputed from another data set, possibly a Luxembourg Income Study (LIS) data set. A general estimation method for complementary data sets was developed by ARELLANO and MEGHIR (1992). For some analyses however, it is not necessary to resort to the general estimation method which is rather complex. A more simple approach was already used by ARELLANO and MEGHIR (1992) to estimate a labour supply function in the presence of job-search for married women in the United Kingdom based on the U.K. Labour Force Survey and the U.K. Family Expenditure Survey. A similar estimation technique, but without considering job-search, will be used to estimate a labour supply function for married women in Germany based on the German Socioeconomic Panel (GSOEP) and the German Labour Force Survey (GLFS). These data bases are comparable to the LIS, respectively the LES, data sets. As the estimations for the cases of the United Kingdom and Germany both lead to plausible results, similar models might be used also for other countries in the LES data base. Thus, for those countries for which data sets in LIS and LES are available for nearly the same year the availability of variables that are necessary for an estimation of a model similar to the one used for Germany is checked.

Acknowledgement
For helpful discussions I would like to thank Bertrand Koebel, Florian Kraus, François Laisney, Viktor Steiner and Elke Wolf. All remaining errors are my own.

1 Introduction

For most countries, the Luxembourg Employment Study (LES) data do not contain any information about wages and income either at the household or at the individual level. Exceptions are Austria 1991 (total personal income), the Czech Republic 1994 (total household income), Poland 1994 (total personal earnings), and the United States 1990 (wages/earnings per hour, total personal earnings, total personal income, total household income) (see FÖRSTER et al., 1996). But with respect to other work-related variables it is a rich data set. The data include information about the work status, the employment characteristics of the first and second job, the previous work experience of persons not in employment, the search for employment, education and training and the labor force status.

For many questions that could be investigated using LES data it would be useful to have additional information about the personal wage rate or household income. The additional information necessary to investigate these questions, especially the wage, might be imputed from another data set, possibly a Luxembourg Income Study (LIS) data set. The LES and LIS data sets are complementary in that they have some elements in common.

A general technique for the estimation using complementary data sets was developed by ARELLANO and MEGHIR (1992).[1] In this technique the estimation using complementary data sets is seen as a moment estimation problem. At the expense of some efficiency loss there exist simpler methods that are appropriate for the estimation using complementary data sets. One of these will be described and used to estimate labour supply of married women in West-Germany using the German Socioeconomic Panel (GSOEP) and the German Labor Force Survey (GLFS). Most of the existing studies of labour supply of married women are solely based on the GSOEP. Accordingly, it is interesting to see whether the results of these also hold on the basis of the GLFS. Thus, one goal of the paper is a contribution to the question of how elastic the labour supply of married women in Germany is.

The second goal is to confirm by this example that the method already used by ARELLANO and MEGHIR (1992)[2] for the case of the U.K. leads to plausible results for the case of Germany too and could thus generally be seen as a valuable instrument for the estimation of labour supply in countries for which large labour force surveys and smaller household budget surveys are available. This is the case for the countries in the LES data base. As the GSOEP and the GLFS are comparable to the LIS respectively the LES data sets, for those countries wage imputations should generally be possible. Although it should generally be feasible to impute wages in the LES data sets, it has to be proved that the data sets are available for nearly the same year and that the variables necessary for the imputation are available in both the LIS and LES data sets. For countries for which the neccessary variables are available the imputation could be done straightforwardly by using SPSS or SAS because only an OLS and a Probit procedure are needed.[3]

---

[1] For a similar independently developed estimation technique (instrumental variables with moments from two data sets) see ANGRIST and KRUEGER (1990).

[2] It should be mentioned that the model used here is simpler than the one of ARELLANO and MEGHIR (1992) because job-search is not considered.

[3] The LIS and LES data sets can be used by sending SPSS or SAS command files via e-mail to the CEPS in Luxembourg. No other statistical packages can be used.

The paper proceeds as follows. In section 2 some of the questions that could be investigated using LES data and for which it would be desirable to have additional information on wage and household income are discussed. In section 3 the method used for the estimation of the labour supply of married women in West-Germany as well as the results of this estimation are decribed. In section 4 it will be checked whether LES and LIS data sets are available for nearly the same year and whether the variables used in the estimation of the labour supply of married women in West-Germany are available in these data sets. Section 5 concludes. In the Appendix, some descriptive statistics and estimation results are shown.

2 Why is it desirable to impute wages in the LES?

The first (perhaps trivial) question to answer is why one should use a LES data set to estimate a model in which the wage is a relevant variable, although it is not available in the LES data set. There are generally two reasons. The first is that there might be other variables in the model that are available in the LES data set and not in the LIS data set. An example for such a variable is the job-search variable. The second reason is that the number of observations in the LES data sets is generally considerably higher than in the LIS data sets so that one might hope that the estimated parameters are more efficient. Furthermore, if one is interested in an international comparison of parameters of the same model, an additional advantage of the LES data sets is that they are conducted in nearly the same way in each country.

The next question to answer is in which cases one should impute the wage[4] from a LIS data set and not use those variables in the LES data set as regressors for which it is well known from other estimations that they might explain the wage well. Suppose that there is a structural model that includes the wage as a regressor. All other variables included in the structural equation might be available in the LES data set. Suppose additionally that in the LES sample there are variables that are able to explain the wage of a person, like variables related to education or vocational training. If you are solely interested in the parameters of variables that are independent of the wage, there is surely no need to impute the wage from another data set. But if you are interested in the parameter of the wage itself, or the parameters of variables that are correlated with the wage, you have to include the wage as a regressor. For example, if you want to know how labour supply reacts to a change of the wage rate, perhaps to predict labour supply given information about the future development of wages, it will surely not suffice to include those variables in the LES data set in your equation that might explain the wage. You have to know exactly how these variables affect the wage. This information can only be provided using another data set.

A well known paper using complementary data sets is by ARELLANO and MEGHIR (1992). They use the U.K. Labour Force Survey (LFS) 1983 and the U.K. Family Expenditure Survey (FES) 1983 to estimate a labour supply model in the presence of job search. The special advantage of the LFS in this case is that it includes information about job-search and the data set is much larger than the FES. On the other hand, the advantage of the FES is that it contains information on wages. Both data sets have a large number of variables in common.

Even though labour supply perhaps seems to be the most important application of techniques for complementary data sets, these might also be useful for other questions that can be investigated

[4] To impute the wage in a LES data set means that an estimated value for the wage is assigned to every element in the LES data set. This value is based on the information about the wage in the LIS data set and some common variables of both the LES and the LIS data set. A method for the imputation will be described below.

using LES data sets. One that was included in the model of ARELLANO and MEGHIR (1992), but that could also be of interest of its own is the search decision of people either having or not having a job. Another application, for which unfortunately information is available only for Luxembourg 1992, Poland 1994, Spain 1993 and the United Kingdom 1989, is participation in education and training.

3 An application to the labour supply of married women in West-Germany

Although an international comparison of the labour supply of married women on the basis of the LIS/LES data sets would certainly be interesting, the estimation will only be done here for (West-) Germany. The estimation method may straightforwardly be applied to other countries for which the necessary variables exist. As household income is available in the GLFS but not for most of the countries in the LES data base, for those countries it would be necessary to additionally estimate a household income equation on the basis of a LIS data set.[5] Although it is generally possible to use the same method for other countries, it should be mentioned that the estimation for every country requires specific knowledge about the variables entering the wage equation, the participation equation and the hours equation. Surely the best specification might differ a lot between the countries.

Like ARELLANO and MEGHIR (1992), I use a two-step estimator for the hours equation. First a wage equation is estimated using the GSOEP-West 1993. As mentioned above, this data set is comparable to the LIS data sets.[6] Only those variables may be included in the equation that are also available in the GLFS of the same year. It will be seen below that this is a strong restriction for the estimation of a wage equation on the basis of the GSOEP. Given the estimated parameters it is possible to impute the wage in the GLFS. In the second step the imputed wage is used as a regressor in the hours equation. The hours equation was estimated by OLS on the sample of working women in the GLFS 1993. As also mentioned above this data set is comparable to the LES data sets.

3.1 Estimation of the gross wage equation on the basis of the GSOEP-West 1993

The wage equation was estimated on the basis of a GSOEP-West 1993 sample for employed married women living with their husband in the same household. This allows to generate information about the husband in order to explain the labour force participation of the spouse. Self-empoyed women, those working in the agricultural sector and those in vocational training were excluded.[7]

To correct for a possible sample selection bias in the log gross wage equation, a participation equation for married females was estimated on the basis of the GSOEP 1993. The resulting inverse Mills-Ratio was included as a regressor in the log gross wage equation. The coefficient of the inverse Mills-Ratio was not significant. Thus, it was not included in the gross wage equation used

---

5 This was also done by ARELLANO and MEGHIR (1992) for the case of the U.K.

6 The GSOEP samples for 1984, 1989 and 1991 are included in the LIS data base. As the GLFS 1995 is publicly available now in Germany, it might also soon be available in the LES data base.

7 For details of the definition of this group see the Appendix.

for the imputation in the GLFS 1993. The results of the participation equation are shown in Table 7 and those of the gross wage equation including the inverse Mills-Ratio are shown in Table 8, in the Appendix.

The arguments entering the wage equation are potential work experience, the level of general education, the level of vocational training, economic sectors, regional variables and nationality. The estimation results are given in Table 1. The reference person has a middle degree of general education and vocational training, she works in the steel or electric industry, she lives in the southern states (Baden-Württemberg or Bayern), and her nationality is German.

Table 1: Wage equation (ln hourly gross wage), married women, GSOEP-West 1993

Variable		Coeff.	t-Value
	constant	2.875	39.20
potential experience	potential experience	0.012	2.14
	potential experience squared /100	-0.021	-1.96
general education	no degree	-0.021	-0.50
	high degree	0.118	2.78
	no answer	0.132	1.63
vocational training	no degree	-0.086	-3.30
	high degree	0.290	6.42
	no answer	0.002	0.04
nationality	foreigner	-0.058	-2.02
economic sector	primary sector	0.019	0.39
	other industries	-0.177	-4.12
	trade	-0.173	-4.47
	cultural activities	0.058	1.13
	financial sector	0.150	3.02
	public sector	-0.001	-0.03
	other services	-0.059	-1.55
	other sectors / no answer	-0.091	-1.82
regional variables	northeastern states	0.111	2.81
	northwestern states	-0.006	-0.15
	Nordrhein-Westfalen	-0.013	-0.50
	Rheinland-Pfalz, Saarland	-0.093	-1.94
	Hessen	0.027	0.80
	R^2	0.26	
	number of observations	888	

Note: GSOEP-West, 1993, married women living in one household with their husband, age 20-59.

The coefficients, especially those for the human capital variables general education and vocational training as well as the potential labor market experience have the expected sign.[8]

[8] Unfortunately the firm size is not included in the GLFS 1993. The firm size is able to explain a lot of the wage variation. The R^2 rises substantially if firm size dummies are included additionally in the log gross wage equation. In a specification with the addition of three firm size dummies, it rose from 0.257 to 0.336 and all dummies were highly significant.

3.2 Estimation of the hours equation on the basis of the GLFS

In order to correct for a possible sample selection bias in the hours equation using only those observations with positive hours supply, a participation equation was estimated using the subsample of married women in the GLFS 1993.

Participation Equation

The reference person has a middle degree of general education and vocational training; in case she has children, the youngest child is between 13 and 17 years old, she lives in a community of medium size (20000 to 500000 inhabitants), her nationality is German. Her husband has neither a high degree of general education nor a high degree of vocational training. He is neither a blue collar worker nor a public sevant. He is working full time.

One of the explaining variables in the participation equation is the „other household income". This variable, defined as household net income minus net earnings of the spouse is not known in the GLFS. It was imputed using the GSOEP-West 1993 in the following steps. First, it was estimated a net wage equation based on the GSOEP. The estimation results of this equation are shown in Table 8 in the Appendix. In the net wage equation the inverse Mills-ratio was significant. Thus, for the net wage imputation in the GLFS an estimate for the inverse Mills-ratio is needed. To get this estimate, in the sescond step, a participation equation on the basis of the GLFS with the net income of the other persons in the household, which is reported in the GLFS, as a regressor, was estimated. In the third step, the coefficients of the net wage equation including the one for the inverse Mills-ratio and the corresponding variables in the GLFS (including the estimated inverse Mills-ratio) were used to impute the net wage per hour. The imputed net wage per hour was used in the fourth and last step to calculate the other household income[9].

The results of the participation equation are shown in Table 2.[10]

[9] Other household income = net household income as reported in the GLFS minus imputed net wage rate multiplied by hours worked per month; hours worked per month = hours worked per week as reported multiplied by 4.

[10] The coefficients of the participation equation do only slightly change if instead of the other household income as defined above, the income of the other persons in the household is used as a regressor. Morevover, the coefficient for the income of the other persons in the household (0.086) only slightly differs from the one for other household income (0.085).

Table 2: Participation equation (Probit), married women, GLFS (West) 1993

	Variables	Coefficient	t-Value
	constant	-3.112	-27.30
age	age	0.234	39.97
	age squared /100	-0.300	-44.08
	under age 25	0.243	6.74
general education	no degree	-0.353	-7.14
	high degree	0.119	5.49
	no answer	0.257	6.93
vocational training	no degree	-0.293	-21.97
	high degree	0.294	12.76
	no answer	-0.325	-9.52
children	number of children	-0.374	-47.81
	youngest child under age of 3	-0.446	-19.64
	age of youngest child between 3 and 5	-0.027	-1.24
	age of youngest child between 6 and 12	0.188	10.62
community size	small community (<20000 inhabitants)	-0.033	-2.79
	large city (>500000 inhabitants)	-0.008	-0.50
regional variables	northeastern states	-0.054	-2.68
	northwestern states	-0.185	-11.14
	Nordrhein-Westfalen	-0.319	-22.31
	Rheinland-Pfalz, Saarland	-0.255	-13.11
	Hessen	-0.152	-8.02
nationality	foreigner from European country	0.022	0.70
	foreigner from non European countries	-0.221	-9.45
other income	household income - wage income of wife /1000	-0.085	-27.32
husbands charac-teristics	age	-0.004	-3.68
	high degree general education	-0.087	-4.87
	high degree vocational training	-0.015	-0.98
	blue collar worker	-0.120	-9.60
	public servant	-0.059	-3.41
	not full time employed	-0.247	-15.15
	% correctly predicted	66,10%	
	Mc Kelvey/Zavoina R^2	0.22	
	number of observations	67148	

Note: married women living in one household with their husband, age 20-59.

Hours Equation

The hours equation is estimated on the basis of the subsample of married women with a positive amount of hours worked.

Table 3: Hours equation, married women, GLFS 1993

Variables		Coefficient	t-Value	adjusted t-Value
	constant	8.601	2.48	2.52
age	age	0.095	0.91	0.92
	age squared /100	-0.384	-2.94	-2.94
children	number of children	-2.265	-9.95	-9.60
	youngest child under age of 3	-7.656	-21.32	-20.61
	age of youngest child between 3 and 5	-6.278	-25.06	-24.27
	age of youngest child between 6 and 12	-3.582	-16.61	-15.97
wage	ln gross hourly wage (imputed)	10.711	15.74	15.95
economic sector	primary sector	-0.663	-2.44	-2.62
	other industries	1.401	5.08	5.32
	trade	-2.914	-11.91	-12.60
	cultural activities	-4.190	-15.37	-15.91
	financial sector	-5.223	-19.40	-21.06
	public sector	-4.023	-18.59	-20.35
	other services	-2.994	-14.37	-15.35
other income	other household income	-0.787	-15.09	-14.87
	other household income * number of children	-0.146	-4.66	-4.28
community size	small community (<20000 inhabitants)	-0.053	-0.42	-0.41
	large city (>500000 inhabitants)	1.204	7.00	7.21
regional variables	northeastern states	-0.678	-2.97	-3.05
	northwestern states	-0.182	-0.96	-0.95
	Nordrhein-Westfalen	-0.132	-0.66	-0.66
	Rheinland-Pfalz, Saarland	1.060	4.68	4.58
	Hessen	0.360	1.68	1.71
nationality	foreigner	3.661	15.66	14.94
husbands charac-teristics	age	-0.038	-2.90	-2.82
	high degree general education	0.249	1.34	1.32
	high degree vocational training	-0.471	-2.95	-2.83
	blue collar worker	-0.961	-6.68	-6.79
	public servant	-0.901	-4.95	-5.08
	not full time occupied	0.570	2.59	2.53
correction	inverse Mills-Ratio	1.636	2.25	2.20
	R^2	0.194		
	number of observations	33047		

Note: GLFS (West), 1993, married women living in one household with their husband, age 20-59.

The arguments include age, children related variables, log gross hourly wage, other household income and a dummy variable for nationality. To control for a possible sample selection bias the inverse Mills-ratio, calculated on the basis of the participation equation described above, is included. The gross hourly wage was imputed using the the coefficients of the gross wage equation estimated on the basis of the GSOEP 1993 (see Table 1). The variable other household income is defined as household net income minus the imputed net wage of the wife. As the wage variable and

the other income variable are based on estimated parameters, the standard errors have to be adjusted. The adjusted t-values (based on heteroskedasticity consistent standard errors) are also given in Table 3. The difference between the t-values and the adjusted t-values is rather small.[11]

The variables for the potential experience, the degree of education and vocational training that are included in the wage equation and the participation equation are excluded in the hours equation. Thus, identification of the parameters in the hours equation is ensured.

The estimated parameter for „ln gross hourly wage" has a plausible value. Calculated at the sample mean the labour supply elasticity is about 0.38.[12] Compared to other estimates mostly based on the GSOEP-West this value is considerably lower than those in the early works of FRANZ and KAWASAKI (1981) and FRANZ (1985) and those in the most recent studies by STRØM and WAGENHALS (1991), LAISNEY et al. (1993), HUJER and SCHNABEL (1994) and HUJER, GRAMMIG and SCHNABEL (1994). It is in line with the result of KAISER (1992) and considerably higher than the result from MERZ (1990).[13] The result is also in line with the findings of ARELLANO and MEGHIR (1992) for the case of Great Britain using the data from 1983.

3.3 An estimation using a proxy for wage in the GLFS

In the GLFS 1993 the personal net income and the household net income are available. Both are given in 15 income groups.[14] Additionally it was asked what the main source of income was. For those for whom the main source of income is employment income, the personal income might be used as a proxy for the monthly net earnings. Divided by the hours worked this might give a proxy for hourly net wage. This was used as a regressor in the hours equation.

The obvious shortcomings of this approach are that, first, personal income is obviously on average higher than net wage. But more important is the possibility that the discrepancy between personal income and net wage systematically differs between population groups. For example, the child benefits might often be reported as income of the mother. Accordingly the net hourly wage, calculated on the basis of personal income, is systematically influenced by the presence of children.

As personal income per hour might be correlated with the same omitted variables than hours supply, a personal income equation was estimated first. The results are shown in Table 9 in the Appendix. Surprisingly, the effect of children is negative. It was expected to be positive because child allowances augment the personal income. On the other hand, women with children work

---

[11] ARELLANO and MEGHIR (1992) derive the adjusted standard errors by explicitly using the information on the errors of the imputed variables. The adjusted t-values in Table 3 are based on the WHITE procedure (see GREENE, 1993). Thus, the information on the errors of the imputed is not taken into account.

[12] If instead of the other household income the income of the other persons in the household is used a regressor in the hours equation, the wage effect on hours is slightly higher. The elasticity is 0.39. The results are not shown here, but may be obtained on request.

[13] For an overview of the results see BUSLEI et al. (1996) and ZIMMERMANN (1993).

[14] Personal income was set equal to the middle value of the interval of every income group. For the highest income group (equal to or more than 7500 DM) it was set equal to the mean value of this income group in the GSOEP.

fewer hours and might be less flexible so that more of them might have to accept poorly paid jobs than women without children.[15]

The estimated personal income per hour was used as a regressor in the hours equation. The estimation results of this equation are shown in Table 10 in the Appendix. The net personal income elasticity (0.27), measured at the mean of the sample of all women employed, is lower than the gross wage elasticity that resulted using the imputed gross wage as a regressor. It is also lower than the net wage elasticity that results using the imputed net wage (0.32).[16] The lower value for the elasticity of personal income could be explained by the fact that an increase of personal income does only to a certain extent mean an increase of net wage (wage effect) and to another extent an increase in other income that should reduce labour supply (income effect).

When comparing the results it should be kept in mind that in the estimation based on personal income in the GLFS only those individuals for which employment income is the most important source of income were included. There is no question in the GSOEP about which source of income was the most important. As there is information about the different sources of income in the GSOEP it might be possible to construct approximately the necessary information. This will not be done here. The estimation on the basis of personal net income was carried out only to validate to a certain extent the result for the gross wage elasticity based on the imputed gross wage. As the difference between the two values for the net wage elasticity is rather small and the direction of the deviation is as expected, the evidence from the estimation based on personal net income supports to a certain extent the results that were obtained by using the imputed wage.

4 Countries in the LIS/LES data base

For the estimation on the basis of complementary data sets, it would be desirable that the data were drawn in the same year. Unfortunately, this is only exceptionally the case for the LIS and LES data sets. But one might argue that an estimation is still valid if the time span between the drawings is not too large. As Table 4 shows, this might be fulfilled for all data sets in LES except of course for Slovenia where no data set is available in LIS.

In the following, I check whether the variables used in the estimation for West-Germany are available in the LES and LIS data files. This does not mean that for other countries other specifications could not be superior, but it seems that the included variables are more or less the basic variables that are suggested by economic theory.

In the case of Germany other household income had a significant influence on the hours supply. In the LES Data base household income is available only for the Czech Republic and for the U.S. For Austria and Poland it might be constructed from total personal income. Therefore for the other countries it seems necessary to estimate household income using a LIS Data base, as was also done by ARELLANO and MEGHIR (1992) for the case of Britain.

---

[15] If, additionally to the variables in Table 1, the number of children is included in the wage equation based on the GSOEP, their influence is significantly negative. But the coefficient is small and much smaller than the coefficient in the equation based on the GLFS.

[16] The results are not shown here, but may be obtained on request.

Table 4: LES[17] and LIS data sets

	LES	LIS
Country	Year	Year
Austria	1991	1991
Czech Republic	1994	1992
Finland	1990	1991
Hungary	1993	1991
Luxemburg	1992	1991
Norway	1990	1991
Poland	1994	1992
Spain	1993	1990
Sweden	1990	1992
United Kingdom	1989	1991
United States*	1990	1991
Slovenia	1994	

Note: *The United States are a special case because wage information is available in the LES data set.

Reference: LIS-Database General description, FÖRSTER et al (1996, 32).

4.1 Availability of variables for the wage and other income equation in LIS and LES data sets

To impute the wage in the LES data set in the same way as it was done in the German case, it is necessary that the variables used in the wage equation are availabe in the LIS as well as the LES data sets. Accordingly I check the availability for both data sets. I do not check whether the variables for a participation equation on the basis of the LIS data sets that are needed for the correction of a possible sample selection bias in the wage equation, are available. It seems plausible that there are enough variables for a suitable specification in every LIS data set.

It should be noted here that in the German case some of the variables in the two data sets have slightly different values. Although not knowing if this is also the case for other countries, I suspect it is. If sufficient knowledge exists on which values mean rather the same or can be aggregated so that the same aggregate variable results in both data sets, this seems not generally to be a serious problem. For an international comparison it seems to be rather costly, but indispensable, to collect this knowledge for every country to be compared.

Availability in LIS

Gross wage or net wage and hours information are available in all LIS files listed in Table 4. Thus the dependent variable gross wage per hour or alternatively net wage per hour that were used above can be constructed for every country. [18]

---

[17] In between two additional data sets for France and Slovakia have been added to the LES data base.

[18] Economic theory suggests that the relevant variable for hours supply is the net wage. But it seems not to be clear whether individuals are fully aware of their tax burden so that the relevant variable is gross wage (see KÖNIG et al., 1995). A more elaborate model of hours supply would explicitly take taxation into account (see for example LAISNEY et al., 1993).

As Table 5 shows, the variable general education is available for every country. As age is available in every country too, it is possible to construct the variable potential experience for every country. Information on vocational training is not available in the Czech Republic, Finland, Poland, Spain and the U. K. For those countries one should check whether information about vocational training can be substituted by information about occupation. I suspect that this is possible but rather time expensive. The economic sector is known except for the U. K. Some regional information is available for every country. Information about nationality is lacking for some countries.

Table 5: Availability of variables for a wage equation in the LIS data sets

	Age	General education	Vocational training	Economic sector	Region	Nationality
Austria	x	x	x	x	x	x
Czech Republic	x	x		x	x	
Finland	x	x		x	x	x
Hungary	x	x	x	x	x	x
Luxembourg	x	x	x	x	x	x
Norway	x	x	x	x	x	
Poland	x	x		x	x	
Spain	x	x		x	x	
Sweden	x	x	x	x	x	x
United Kingdom	x	x			x	

Note: x: is available.

Availability in LES

Mostly all the variables necessary for the wage equation are available in the LES data sets. Age, general education, economic activity of the establishment/industry and region are available for all countries listed in Table 4. Information on vocational training is available for all countries except for Finland 1990. Nationality is only available for Finland, Luxembourg, Spain and the U.K.

To conclude: for Austria, Hungary, Norway and Sweden the estimation along the lines of the estimation described above for the case of Germany is surely possible. For the other countries it seems essential that a substitute for vocational training can be constructed (skill variable) perhaps using the information about occupation. Because information on occupation is missing for the U. K. there might be serious problems in this case.

The other income equation can be specified nearly in the same way as the wage equation. Additionally one should include information about the husband (age, education, economic sector, ...). Thus, the imputation should be possible if the wage imputation is possible. The exceptions are the Northern countries Finland, Norway and Sweden (see FÖRSTER et al., 1996), where the household serial number is missing and thus, no information about the husband can be generated.

4.2 Availability of variables for the participation and the hours equation in LES

For the participation equation the necessary information on age, general education, vocational traning (except for Finland) and children is available. Age and number of children as well as husbands characteristics can be calculated for all countries except for the nordic countries Finland, Norway and Sweden (see FÖRSTER et al., 1996). The urban/rural indicator that was significant in

the German case is not available for the Czech Republic, Hungary, Luxembourg, Norway and the United Kingdom. Nationality, that was also significant in the German case, is only available for Finland, Luxembourg, Spain and the UK.[19]

Given that it is possible to estimate a wage equation and an equation for other income on the basis of a LIS data set with independent variables that are also availbale in the corresponding LES data set, all necessary variables for the hours equation (except nationality, and vocational training for the case of Finland) are available in all LES data sets. Problems only arise for the nordic countries because household and personal information cannot be matched to generate information about the children and the husband.

5 Conclusion

One goal of this paper was to make a contribution to the question how elastic the labor supply of married women in West-Germany is. Most of the existing studies on labour supply of married women in West-Germany are based on the GSOEP. In this paper the hours supply of married women was estimated on the basis of the GLFS 1993 and the GSOEP 1993. One advantage of the GLFS is that it is much larger than the GSOEP. As the GLFS does not contain information about the wage rate, this was imputed from the GSOEP. To impute the wage, a similar, but somewhat simpler framework than the one used by ARELLANO and MEGHIR (1992) for the case of Britain was used. It was found that the (imputed) wage had a very significant influence on hours supply. The gross wage elasticity of hours supply was 0.38. Some of the values reported in the literature are considerably higher, some considerably lower and some are nearly the same. Thus the result may provide an incentive to investigate what determines the difference between the estimates. One possible explanation might be that some studies who show a high elasticity use another estimation approach, which lets the same process govern both the participation and the hours decision whereas these are decoupled in the selectivity approach followed here (for the influence of the estimation method on the coefficients in labour supply models, see MROZ, 1987).

The results above confirm that the method already used by ARELLANO and MEGHIR (1992) for the case of Great Britain leads to plausible results for the case of Germany too and that it thus can generally be seen as a valuable instrument for the estimation of labour supply for the countries in the LES data base.

Moreover I have checked that in most of the countries in the LES the variables that are needed for the wage imputation are included in the LES and the LIS files.

The LES data sets as well as the GLFS are cross-section data. The estimation of labour supply from cross-section data can not adequately tackle the life-cycle aspects of labour supply. Hence these estimations are normally second best given the availability of panel data. Yet, one advantage of the labour force surveys is the higher number of observations. The second advantage is that they have been conducted in several countries with a nearly identical concept. One might suppose that estimations on such a basis should be adequate for an international comparison. Such a comparison based on the LIS and LES data seems to be a fruitful and interesting task for further research.

[19] The variable ethnicity might be used instead of nationality. It is available for the Czech Republic, Finland, Hungary and the UK.

Appendix A: Variable definitions and descriptive statistics

Brief definition of variables that are not self explaining

GSOEP-West (1993)

hours worked: average normal working time per week as reported; if there was a missing value for average normal working time, the working time according to the contract was used.

hourly gross (net) wage: gross (net) wage per month as reported divided by *hours worked* multiplied by 4.

potential experience (potexp): age - years of schooling -6.

degree of general education: no degree, middle degree: Hauptschule, Realschule, high degree: Abitur or höhere Fachschule.

degree of vocational training: no degree, middle degree: Lehre, high degree: Fachschule, Fachhochschule, Ingenieur, Beamtenausbildung, Universität.

primco: energy and water supply; mining; manufacture of chemicals, chemical products and man-made fibres; manufacture of rubber and plastic products; quarrying of sand, clay and stones; manufacture of other non-metallic mineral products; construction.

steel: manufacture of basic metals and fabricated metal products; manufacture of machinery and equipment; manufacture of electrical and optical equipment; manufacture of transport quipment.

oind: manufacture of wood and wood products; manufacture of furniture; manufacture of pulp, paper and paper products, publishing and printing; manufacture of textiles and textile products; manufacture of food products, beverages and tobacco.

trade: wholesale and retail trade.

cult: education, research, sports.

finance: financial intermediation; other business activities (legal, accounting and similar services).

puboth: transport, storage and communication; membership organizations; private households; public administration; compulsory social security.

oserv: hotels and restaurants; real estate, renting and similar services; cleaners, sewage and refuse disposal; health sector.

osecna: industries, services and trades, that cannot be related to one of the sectors above, no answer and wrong answers

GLFS (1993)

hours worked: normal working time per week as reported; if there was a missing value for normal working time, actual working time in the week of the interview was used.

net personal income: middle value of the interval of the income groups in the GLFS (15 different income groups), for the highest income group (equal to or more than 7500 DM) it was set equal to the mean value of this income group in the GSOEP.

net personal income per hour: *net personal income* divided by *hours worked* multiplied by 4.

income of other persons in the household: mean of net household income groups in the GLFS (15 different groups) minus *net personal income*

potential experience (potexp): age - years of schooling -6.

degree of general education: no degree, middle degree: Hauptschule, Realschule, high degree: Abitur or Fachhochschulreife

degree of vocational training: no degree, middle degree: Lehre, high degree: Fachschule, Meister, Techniker, Fachhochschule, Universität.

economic sectors (see definitions of GSOEP).

foreigner from a European country: foreigners from western european countries including those who are not members of the European Union.

Descriptive Statistics

The descriptive statistics refer to the subsample of married women who are

- gainfully employed,
- living in a household with their husband (being head of household or spouse of head of household),
- 20 to 59 years old,
- not working in the agricultural sector,
- not in school (university) or on vocational training,
- not self-employed.

Moreover, it is required that

- the hourly wage can be computed from the sample (hourly netwage >0, hourly gross wage > 0, hours worked >0).
- the hourly gross wage is higher than 7.5 DM and the hourly net wage is higher than 5 DM.

Table 6: Descriptive Statistics GSOEP-West 1993, GLFS-West 1993

		GSOEP		GLFS	
	Variable	Mean	Std	Mean	Std
work time	mean working time per week	30.21	11.19	28.36	10.89
wage	hourly gross wage	21.33	8.83	20.86	4.43
	hourly net wage	14.03	5.93	13.60	2.98
age	age	41.17	9.43	41.04	9.58
pot. exp.	potential experience	24.89	9.61	25.45	10.14
general education	no degree	0.03	0.16	0.01	0.12
	high degree	0.11	0.31	0.14	0.35
	no answer	0.01	0.09	0.09	0.29
vocational training	no degree	0.24	0.43	0.19	0.39
	high degree	0.10	0.30	0.12	0.32
	no answer	0.09	0.29	0.09	0.29
children	number of children	0.64	0.90	0.69	0.91
	youngest child under age of 3			0.05	0.22
	age of youngest child between 3 and 5			0.08	0.26
	age of youngest child between 6 and 12			0.18	0.38
	# children at the age under 4	0.07	0.25		
	# children at the age of 4 or 5	0.08	0.28		
other persons	number of other persons in household	0.42	0.72		
	income of other persons in household			3633	1782
handicap	degree of handicap over 60 %	0.01	0.10		
care	persons in household with need for Care	0.01	0.12		
other inc.	other household income / 1000	3.38	1.69	3.73	1.92
	other household income /1000 * # children			2.59	4.04
economic sector	primary sector	0.05	0.22	0.06	0.24
	steel	0.10	0.30	0.10	0.31
	other industries	0.07	0.26	0.09	0.28
	trade	0.20	0.40	0.18	0.39
	cultural activities	0.08	0.27	0.09	0.29
	financial services	0.07	0.26	0.09	0.28
	public sector	0.17	0.37	0.17	0.37
	other services	0.20	0.40	0.21	0.41
	other sectors / no answer	0.06	0.23	-	-
regional variables	northeastern states	0.10	0.30	0.10	0.30
	northwestern states	0.10	0.30	0.13	0.33
	Nordrhein-Westfalen	0.29	0.46	0.22	0.42
	Rheinland-Pfalz, Saarland	0.06	0.24	0.09	0.29
	Hessen	0.10	0.30	0.08	0.27
community size	small community (<20000 inhabitants)			0.43	0.50
	large city (>500000 inhabitants)			0.16	0.37
nationality	foreigner	0.09	0.29		
	foreigner from european country			0.03	0.16
	foreigner from non european country			0.05	0.21
husbands charac-teristics	age	44.53	10.28	44.12	10.19
	high degree general education	0.15	0.35	0.19	0.40
	high degree vocational training	0.25	0.43	0.27	0.44
	blue collar worker	0.48	0.50	0.41	0.49
	public servant	0.11	0.31	0.11	0.32
	not full time occupied	0.15	0.36	0.12	0.33

Note: married women, living in one houeshold with their husband, age 20-59, employed.

Appendix B: Estimation results

Participation Equation GSOEP-West

The reference person has a medium degree of general education (Hauptschule, Realschule) and vocational training (Lehre), if she has any children living with her in the household they are between 13 and 18 years old, she lives in the southern states (Baden-Württemberg or Bayern), her handicap indicator is less than 60% and her nationality is German. Her husband has not a high degree of general education and not a high degree of vocational training. He is neither a blue collar worker nor a public servant and he is working fulltime.

Table 7: Participation equation; GSOEP-West 1993

	Variable	Coeff.	t-Value
	constant	-0.069	-0.12
age	age	0.054	1.75
	age squared /100	-0.113	-3.13
general education	no degree	0.170	1.44
	high degree	0.325	2.53
	no answer	-0.121	-0.52
vocational training	no degree	-0.142	-1.91
	high degree	0.071	0.54
	no answer	0.050	0.38
children	number of children	-0.223	-5.77
	child at the age under 4	-0.932	-9.13
	child at the age of 5 or 6	-0.214	-2.16
oth. persons	other persons in household	0.060	1.35
care	person in household with need for care	-0.170	-0.79
other income	other household income (/1000)	-0.140	-6.30
handicap	degree of handicap over 60%	-0.564	-2.40
nationality	foreigner	-0.167	-1.98
regional variables	northeastern states	-0.020	-0.17
	northwestern states	-0.211	-2.01
	Nordrhein-Westfalen	-0.190	-2.53
	Rheinland-Pfalz, Saarland	-0.313	-2.35
	Hessen	0.134	1.27
husbands charac-teristics	age	0.015	2.05
	high degree general education	-0.208	-1.87
	high degree vocational training	0.240	2.40
	blue collar worker	0.146	1.91
	public servant	0.030	0.26
	not full time occupied	-0.163	-1.64
	% correctly predicted	67.02	
.	**Mc Kelvey/Zavoina R^2**	0.26	
	number of observation	2001	

GSOEP-West, 1993, Married women living in one household with their husband, age 20-59.

Table 8: Wage equation with selectivity correction; GSOEP-West, 1993

		ln hourly gross wage		ln hourly net wage	
	Variable	Coeff.	t-Value	Coeff.	t-Value
	constant	2.863	37.34	2.324	31.58
potential experience	potential experience	0.012	2.17	0.017	3.35
	potential exp. squared /100	-0.021	-2.01	-0.035	-3.39
general education	no degree	-0.019	-0.46	-0.052	-1.31
	high degree	0.119	2.80	0.125	3.07
	no answer	0.128	1.58	0.065	0.84
vocational training	no degree	-0.088	-3.34	-0.047	-1.85
	high degree	0.289	6.40	0.321	7.40
	no answer	0.003	0.08	0.019	0.48
economic sector	primary sector	0.019	0.37	0.015	0.32
	other industries	-0.178	-4.13	-0.181	-4.37
	trade	-0.173	-4.46	-0.128	-3.44
	cultural activities	0.059	1.14	0.083	1.69
	financial sector	0.149	3.00	0.166	3.47
	public sector	-0.001	-0.02	0.041	1.09
	other services	-0.059	-1.56	-0.028	-0.78
	other sectors / no answer	-0.089	-1.79	-0.079	-1.66
regional variables	northeastern states	0.110	2.76	0.103	2.70
	northwestern states	-0.009	-0.23	-0.021	-0.59
	Nordrhein-Westfalen	-0.016	-0.60	-0.016	-0.60
	Rheinland-Pfalz, Saarland	-0.097	-2.00	-0.106	-2.27
	Hessen	0.028	0.82	0.024	0.74
nationality	foreigner	-0.058	-2.04	-0.064	-2.32
correction term	inverse Mills Ratio	0.018	0.56	0.075	2.20
	R^2	0.26		0.29	
	number of observations	888		888	

Note: married women living in one household with their husband, age 20-59.

Table 9: Personal income equation; Dependent variable: logarithm of hourly personal income; GLFS-West

	Variable	Coeff.	t-Value
	constant	2.081	100.00
potential experience	potential experience	0.047	33.59
	potential experience squared /100	-0.092	-34.28
general education	no degree	-0.284	-11.22
	high degree	0.206	22.40
	no answer	0.213	11.42
vocational training	no degree	-0.206	-30.23
	high degree	0.263	26.99
	no answer	-0.140	-8.39
economic sector	primary sector	-0.029	-2.80
	other industries	-0.186	-19.28
	trade	-0.171	-20.53
	cultural activities	0.053	5.15
	financial sector	0.026	2.67
	public sector	-0.021	-2.44
	other services	-0.105	-13.05
regional variables	northeastern states	0.043	5.72
	northwestern states	-0.073	-10.27
	Nordrhein-Westfalen	-0.074	-11.65
	Rheinland-Pfalz, Saarland	-0.092	-10.53
	Hessen	-0.027	-3.45
nationality	foreigner	-0.079	-8.83
children	number of children	-0.107	-26.29
husbands characteristics	age	-0.003	-6.04
	high degree general education	0.026	3.35
	high degree vocational training	-0.013	-2.02
	blue collar worker	-0.127	-24.67
	public servant	0.007	0.92
	not full time occupied	0.055	7.86
correction term	inverse Mills-Ratio	0.476	32.12
	R^2	0.28	
	number of observations	28349	

GLFS-West, 1993, married women living in one household with their husband, age 20-59, main source of income is income from employment.

Table 10: Hours equation, married women, net personal income, GLFS 1993

	Variables	Coeff.	t-Value
	constant	20.920	7.34
age	age	0.037	0.38
	age squared /100	-0.291	-2.37
children	number of children	-1.446	-6.52
	youngest child under age of 3	-8.098	-20.33
	age of youngest child between 3 and 5	-5.943	-23.11
	age of youngest child between 6 and 12	-2.852	-12.58
personal income	ln hourly net personal income	7.712	13.01
economic sector	primary sector	0.054	0.21
	other Industries	1.608	6.08
	trade	-2.209	-9.55
	cultural activities	-3.453	-13.34
	financial Sector	-2.861	-11.74
	public sector	-3.013	-14.34
	other services	-1.530	-7.33
other income	other household income	-0.853	-14.86
	other household income * number of children	-0.185	-5.52
community size	small community (<20000 Inhabitants)	0.228	1.81
	large city (>500000 Inhabitants)	0.696	4.17
regional variables	northeastern states	-0.166	-0.80
	northwestern states	-0.325	-1.78
	Nordrhein-Westfalen	-0.002	-0.01
	Rheinland-Pfalz, Saarland	-0.166	-0.72
	Hessen	-0.081	-0.40
nationality	foreigner	3.41	15.14
husbands charac-teristics	age	-0.014	-1.11
	high degree general education	0.035	0.18
	high degree vocational training	-0.153	-0.96
	blue collar worker	0.152	1.09
	public servant	-1.091	-5.97
	not full time occupied	-1.113	-4.51
correction term	inverse Mills-Ratio	1.519	2.31
	R^2	0.165	
	Number of observations	28349	

GLFS (West), 1993, married women living in one household with their husband, age 20-59, main source of income is income from employment.

References

ANGRIST, JOSHUA D. and ALAN B. KRUEGER (1990): The Efect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples, Working Paper # 274, Industrial Relations Section, Princeton University.

ARELLANO, MANUEL and COSTAS MEGHIR (1992): Using Complementary Data Sources: An Application to Labour Supply and Job Search, Review of Economic Studies 59, 537-559.

BUSLEI ET AL. (1996): Möglichkeiten und Grenzen der Entkopplung von Erwerbstätigkeit und sozialer Sicherung, unpublished report for the „Kommission für Zukunftsfragen der Freistaaten Bayern und Sachsen“, Bonn.

FÖRSTER, MICHAEL, ANNE HELLIESEN and JOHN EVIND KOLBERG (1996): LES: The New Challenge, An Introduction to Concepts, Documentation, and Background Statitics, Luxembourg Employment Studies Working Paper Series, No. 1, June 1996, Luxembourg.

GREENE, WILLIAM H. (1993): Econometric Analysis, Second edition, Englewood Cliffs, New Jersey.

HUJER, REINHARD, und REINHOLD SCHNABEL (1994): „The Impact of Regional and Sectoral Labor Market Conditions on Wages and Labor Supply - An Empirical Analysis for Married Women Using West-German Panel Data“, Empirical Economics, Vol. 19, No. 1, 19-35.

HUJER, REINHARD, JOACHIM GRAMMIG, and REINHOLD SCHNABEL (1994): A Comparative Empirical Analysis of Labour Supply and Wages of Married Women in the FRG and the USA, Jahrbücher für Nationalökonomie und Statistik, Bd. 213/2, Stuttgart, 129-147.

KÖNIG, HEINZ, FRANÇOIS LAISNEY, MICHAEL LECHNER AND WINFRIED POHLMEIER (1995): Tax Illusion and Labour Supply of Married Women: Evidence from German Data, Kyklos, 48, 347-368.

LAISNEY, FRANÇOIS, MICHAEL LECHNER, ARTHUR VAN SOEST and GERHARD WAGENHALS (1993): A Life Cycle Labour Supply Model with Taxes Estimated on German Panel Data: The Case of Parallel Preferences, The Economic and Social Review, 24, 335-368.

MROZ, THOMAS. A. (1987): The Sensitivity of an Empirical Model of Married Women´s Hours to Work to Economic and Statistical Assumptions, Econometrica, 55, 765-799.

ZIMMERMANN, KLAUS F. (1993): Labour Responses to Taxes and Benefits in Germany, Welfare and Work Incentives, A North European Perspective, A. B. ATKINSON und G. V. MOGENSEN (Hrsg.), Oxford, 192-240.