THE WAGE ELASTICITY OF LABOUR SUPPLY: A SYNTHESIS OF EMPIRICAL ESTIMATES****

BY

MICHIEL EVERS*, RUUD DE MOOIJ** AND DANIEL VAN VUUREN***

Summary

This paper performs a meta-analysis of empirical estimates of uncompensated labour supply elasticities. For the Netherlands, we find that an elasticity of 0.5 for women and 0.1 for men is a good reflection of what the literature reveals. The elasticity for men hardly differs between countries, but for women some cross-country variation is found. The increasing participation rate of women may lead to a somewhat lower elasticity in the future. Both the specification of the hours function and the estimation method are found to affect elasticity estimates.

Key words: labour supply, uncompensated wage elasticity, meta-analysis

JEL Code(s): J22, H3

1 INTRODUCTION

The elasticity of labour supply with respect to the wage rate[1] plays a critical role in many economic policy analyses. For example, its value determines to a large extent the employment impact of reforms in redistributive tax-benefit systems (see e.g. Graafland et al. (2001)). Moreover, it is crucial for the magnitude of the efficiency cost of income taxation in general equilibrium models (see e.g. Ballard et al. (1985); Browning (1987)). Indeed, the larger is the elasticity of labour supply, the larger is the employment effect in response to a change in the tax rate and the higher is the excess burden of taxation.

*   Ministry of Economic Affairs, The Hague, The Netherlands
**   CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands; Erasmus Universiteit Rotterdam, Rotterdam, The Netherlands; Tinbergen Institute, Rotterdam, The Netherlands; Netspar, Tilburg, The Netherlands; CESifo, Munich, Germany
***   CPB Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM, The Hague, The Netherlands, phone: +31 70 3383 367, fax: +31 70 3383 350, e-mail: djvv@cpb.nl; Netspar, Tilburg, The Netherlands
1   The uncompensated wage elasticity of labour supply is defined as the percentage change in hours worked as a result of a one percent change in the net hourly wage rate.

In light of its importance, a large number of studies have estimated the uncompensated elasticity of labour supply. The results of this literature are reviewed in e.g. Blundell and MaCurdy (1999). It appears that there exists great variation in study results and an equally large variation in approaches to estimate the elasticity. As a result, there is little agreement among economists on the magnitude of the elasticity that should be used in economic policy analyses. To illustrate, Prescott (2004) explains the difference in hours per worker between the US and Europe entirely by the differences in redistributive tax-benefit schemes between the two continents. Alesina et al. (2005), however, maintain that this story would require an unrealistically large value of the uncompensated elasticity of labour supply.

Some studies have tried to explain the wide dispersion in empirical estimates of the uncompensated labour supply elasticity in the literature. A typical finding is that the elasticity for women exceeds that for men. Another is that the elasticity regarding the decision to participate (the extensive margin) exceeds the elasticity of the decision regarding hours worked (the intensive margin). This latter result may also explain the relatively large elasticity for women, as the participation rate among women is typically lower than for men. The rising participation rate among women in recent decades may have led to a decline in the elasticity of labour supply of women, as is for instance found by Blau and Kahn (2007).

Other study characteristics that can explain the variation in research findings have been explored by Mroz (1987). He examines the effects of economic and statistical assumptions on outcomes for married women. His findings suggest that specification and exogeneity assumptions have a substantial impact on the estimated elasticities. MaCurdy et al. (1990) explore the impact of implied model restrictions on parameter estimates in the context of maximum likelihood estimation of structural labour supply models. Their outcome is similar to Mroz'. Ericson and Flood (1997) conclude that different estimation strategies may lead to quite some dispersion between estimates, in particular when measurement error is present. Ecklöf and Sacklén (2000) show that the construction of variables from raw data may play an important role as well. They attribute the difference between the findings of Hausman (1981) and MaCurdy et al. (1990) precisely to this. Thus, it is found that the method, data, specification or estimation technique have a potentially large impact on the estimates of the model parameters, and in the end, on the (most often implicitly) estimated labour supply elasticity.

Our paper contributes to this literature by analyzing the systematic impacts of the various factors on the reported empirical estimates simultaneously. In particular, we use a sample of 209 uncompensated labour supply elasticities obtained from the literature to perform a meta analysis, i.e. we regress the elasticities on the underlying study characteristics. Apart from gender, participation, estimation method and model specification, we also explore the

impact of a number of other study characteristics and control variables used in the empirical studies. Moreover, we explore whether there are systematic differences between countries.

A second contribution of our paper is to obtain a synthesis of research results. In particular, the careful attention that is usually paid to consistent estimation of parameters in a properly defined (but not over-specified) model comes along with limited applicability of the resulting estimates for policy analysis. For instance, reduced-form estimates resulting from natural experiments are often of little use in the ex ante evaluation of new policy reforms. Our analysis aims to contribute to the synthesis of research results and, thus, on the size of the elasticities to be used in economic policy analysis.

The rest of this paper is organised as follows. Section 2 gives a brief description of issues in the empirical literature on labour supply elasticities. In section 3 we discuss the sample of elasticities used. In section 4 we discuss our strategy to estimate what causes the differences between elasticity estimates, and discuss estimation results. In this section we also produce point estimates for male and female labour supply elasticities for a selection of countries, and we perform some statistical tests. Section 5 concludes.

2 THE EMPIRICAL LITERATURE ON LABOUR SUPPLY ELASTICITIES

It appears from the literature that the estimation of the elasticity of labour supply with respect to the real after-tax wage rate is not a straightforward exercise, due to e.g. nonlinear budget constraints, unobserved wages for non-workers, and various econometric and specification issues. The many ways to deal with these problems are briefly discussed in more or less chronological order in this section. The review contains a number of seminal articles, but is certainly not meant to be exhaustive. For more complete surveys on the topic we refer to Pencavel (1986); Killingsworth and Heckman (1986) and Blundell and MaCurdy (1999).

The first empirical effort known to estimate labour supply elasticities was made by Douglas (1934) in his 'Theory of Wages'. He used aggregated data with age-sex groups for 38 US cities, collected from the 1920 Census of manufactures and examined both time series and cross-section data on hours of work and hourly earnings. Douglas found an elasticity that "is in all probability somewhere between −0.1 and −0.2". Modern labour supply studies often separate the income and substitution effects and make use of micro data instead of aggregate data. The first studies that make the distinction between income and substitution effects are Mincer (1963) and Kosters (1966).

Estimating the elasticity of labour supply under the presence of progressive taxes is not straightforward, because a linear model would not adequately represent the essential nonlinearities in the labour supply decision of individuals (for a comprehensive survey, see Moffitt (1990)). To deal with this,

Hall (1973) assumes that an individual will behave the same whether he faces the real budget constraint or a linear extension of the segment he is actually located on. A problem with this approach (as well as with a naive linear model) is that the individual's segment, and consequently his wage rate, is self-chosen and not exogenous with respect to hours worked. The consequences of this endogeneity in the regressors are investigated extensively in Mroz (1987) and, more recently, discussed in Heim and Meyer (2003). As an alternative, the instrumental variable (IV) approach offers a robust and still relatively straightforward way to obtain a consistent estimate of the uncompensated labour supply elasticity. A major problem of IV-estimation is, however, to find instruments that both satisfy the exclusion restriction and yet show 'enough' correlation with the endogenous regressor. For the net wage rate, gross wages are often used as an instrument. Studies that use both OLS and four are Mroz (1987); Blomquist (1996); Pencavel (2002); Eissa and Hoynes (2004) and Blau and Kahn (2005).

An important problem with OLS is endogenous selection. As is known, selection on endogenous factors, such as the level of income or being employed, biases the results of simple linear regression techniques. If wages are observed for all individuals, then the Tobit and two step Heckman method can take into account that only individuals with a positive amount of hours worked are observed. An important difference between the two models is that the latter approach allows for two different sets of regressors to estimate the participation and labour supply decision. A complication arises when wages are not observed for non-workers. However, the probability that an individual does not work can be estimated in a binomial model and applied in the well-known Heckman (1979) model to correct for endogenous selection. To compute wages for non-workers an often used approach is to estimate the observed wages for workers with basic regression techniques on the individual characteristics and the Mill's ratio based on the estimated participation probability. While these characteristics are also observed for non-workers, wages can be imputed for non-workers by using the estimated coefficients of the regression. Studies that explicitly take into account selection are Cogan (1981); Mroz (1987); Arellano and Meghir (1992); Blundell et al. (1998) and Devereux (2004). Mroz explicitly tests for different selection models and rejects the Tobit model for more general models that allow different specifications for participation and labour supply.

Burtless and Hausman (1978) explicitly take into account the differently sloped segments of the kinked budget curve by linking the choice of segment to the indirect utility function. This method is frequently used in labour supply studies in the 1980s and 1990s, such as Hausman (1980, 1981); Blomquist (1983); Hausman and Ruud (1984); Arrufat and Zabalza (1986); Blomquist and Hansson-Brusewitz (1990); Bourguignon and Magnac (1990); Colombino and del Boca (1990); Triest (1990); van Soest et al. (1990); Flood and MaCurdy (1992); Kuismanen (1997) and Woittiez and Kapteyn (1998).

MaCurdy et al. (1990) criticise the Hausman method for imposing too strong *a priori* restrictions on the outcomes. As an alternative, the authors propose a method with a twice differentiable convex budget constraint. The specification of the model is however more cumbersome, and the model is hardly used in empirical studies. Apart from the original article, a second application is Flood and MaCurdy (1992). As would be expected when the differentiable budget curve is a good approximation of the kinked budget curve, the results do not differ much from those obtained with the 'original Hausman method'.

The standard model of labour supply does not distinguish between the effect of wages and taxes on the decision to participate (the extensive margin) and the decision regarding hours worked (intensive margin). Yet, workers rarely choose a small number of hours. Perhaps fixed costs of entering the labour market, such as child care or transport costs, or institutional factors such as the tax system are relevant. Supply restrictions may also play a role. Mroz (1987) indeed finds evidence that the labour supply behaviour at the extensive margin differs from the behaviour at the intensive margin. The author finds that, when this effect is neglected, the estimated wage elasticity is biased upwards, because hours of work conditional on participation are relatively inelastic with respect to the net wage, while the participation decision is quite elastic with respect to the net wage. A way to model the decision to participate is to include fixed time or fixed money costs of entering the labour market. The latter can be introduced as a reduction in non-labour income if the number of hours worked is positive, so that individuals will then only supply labour above a minimum number of hours. Bourguignon and Magnac (1990) estimate a model for women that includes fixed costs and find that the uncompensated labour supply elasticity is reduced from 0.96 to 0.39. Cogan (1981) finds that the uncompensated labour supply elasticity falls from 2.45 to 0.88 after correcting for entry costs.

An alternative approach is to estimate a participation equation first and then estimate the supply function conditional on the predicted participation. van Soest (1995) includes dummy variables for certain discrete hours choices less than full time and finds a negative effect on the estimated wage coefficients. The main drawback is that the restrictions are assumed to be homogeneous across individuals. Recently, discrete choice models have become more popular for estimating labour supply elasticities. The advantage of models with discrete choice is that it is not necessary to define the entire budget constraint. Discrete choice models assume that individuals choose from a finite set of hours of work, so that only a limited number of choices need to be evaluated. Moreover, the restrictions in the Hausman model do not need to be imposed. Studies that use this approach are, amongst others, van Soest (1995); Euwals and van Soest (1999); Euwals (2001); Bonin et al. (2002) and Bargain (2005).

Recent studies often make use of policy reforms. The preferred case is to compare two randomly selected groups before and after the introduction of a policy change. One group should have experienced the change (the treated) and the other should not (the controls). This approach gives unbiased estimates if time effects are common across the two groups, and endogenous switching between groups is not allowed.[2] An important question is whether the estimators measure behavioural responses. When no structural specification is used, income and substitution effects and intertemporal and within-period effects are easily mixed up. An example of a study making use of a 'policy reform' is Saez (2003), who uses the fact that an individual pays a higher marginal tax rate if his wage increases with the inflation rate, while the income tax brackets are fixed nominally. Contrary to other studies that use tax reforms, this enables him to separately estimate the income and substitution effects.

3 CONSTRUCTING A SAMPLE OF LABOUR SUPPLY ELASTICITIES

In this section we construct a data set of empirical estimates of the uncompensated labour supply elasticity. Our focus on the uncompensated labour supply elasticity is governed by the availability of research findings. In particular, many studies do not report separate values for the compensated elasticity and the income elasticity of labour supply. Even for the uncompensated elasticity, it is not always possible to derive its implicitly estimated value from the presented research findings. Thus, we only select studies which either explicitly or implicitly report one or more estimates. It is however emphasised that the selected sample of elasticities is by no means exhaustive. Our search has primarily focussed on highly reputed academic journals and recent working papers, but was not able to include all the literature on labour supply.

Not all studies report explicit estimates of the uncompensated elasticity. Still, it is often possible to construct a consistent estimate from reported point estimates of marginal coefficients using sample statistics of hours worked and the wage rate. In particular, denote hours worked by h, w the wage rate, Y non-labour income, x a vector of control variables, and β a parameter vector of the same dimension as x. Then the hours function and the uncompensated labour supply elasticity for an individual with characteristics w, Y, and x reads as follows:

$$h = \phi(w, Y, x | \beta) \Rightarrow e := \frac{\partial \ln h}{\partial \ln w} = \frac{w}{h} \phi_w(w, Y, x | \beta), \tag{3.1}$$

2 Of course, non-generic time effects and endogenous switching can be allowed if the econometrician is capable of correcting for these.

where ϕ_w is the derivative of the hours function ϕ with respect to the wage rate w. After substituting sample means for w, h, Y and x, a consistent estimator of the elasticity e at the sample mean is obtained (i.e. for an imaginary individual whose characteristics precisely match the sample mean).

In constructing a sample of elasticities, a number of studies cannot be used. In particular, many studies based on tax reforms do not report uncompensated labour supply elasticities nor the required information to compute this figure (cf. (3.1)). Hence, these studies could not be used in our sample. Also, we have excluded estimates obtained from OLS estimation, for these are known to be inconsistent (see section 2). Ultimately, our literature search yields a sample of 209 elasticities obtained from 30 different studies. Tables 1 and 2 present some summary indicators for the sub-samples of men and women, respectively. We see that the mean elasticity for men equals 0.07, whereas for women it equals 0.43. Mean elasticities per study for men range between −0.08 and 0.18. For women, the table shows great variation across different studies: mean elasticities range from 0.03 to 2.79. If we leave out four outlier observations – which are defined as observations which differ more than two standard deviations from the sample mean – then the mean elasticity for women equals 0.34. The number of elasticities per study varies from one (various studies) to 23 (Mroz (1987)). The last column in both tables shows that although most estimates of the elasticities are significantly different from zero at a five percent significance level, a number is not significant or no information was supplied by the authors.

Figures 1 and 2 show empirical distributions for the sub-samples of men and women, respectively. The dashed lines correspond with the borders of the interval with twice the standard deviation around the mean. In Figure 2, the left border of the interval is smaller than the minimum value and therefore, not shown in the figure. Clearly, the variation in point estimates for men is much smaller than that for women. It also appears that the distribution for women is more skewed to the right. Median values of elasticities for men and women are indicated by the solid lines. For women, a substantially higher value than for men is found (0.27 and 0.08, respectively).

4 ESTIMATION

In this section we aim to identify the factors which systematically impact the magnitudes of elasticity estimates. We do this by running *meta regressions* of a set of so-called 'moderator variables' on the empirical wage elasticity of labour supply. In particular, we run regressions of the following form:

$$e = c + X\beta + \eta \tag{3.2}$$

TABLE 1 – SUMMARY STATISTICS OF STUDIES FOR THE LABOUR SUPPLY OF MEN INCLUDED IN THE META-SAMPLE

Author(s) (year of study)	Obs.	Mean	Median	Max.	Min.	St. Dev.[a]	Country	Year[b]	Sign.[c]
Blau and Kahn (2005)	6	0.07	0.07	0.13	0.01	0.04	US	1980/90/2000	6/0/0
Blomquist (1983)	2	0.08	0.08	0.08	0.08	0.00	Sweden	1973	2/0/0
Blomquist (1996)	3	0.04	0.08	0.18	−0.13	0.16	Sweden	1981	3/0/0
Blomquist and Hansson-Brusewitz (1990)	4	0.10	0.10	0.13	0.08	0.03	Sweden	1981	3/1/0
Blomquist and Newey (2002)	24	0.08	0.08	0.12	0.04	0.02	Sweden	1982	24/0/0
Bonin and Kempe (2002)	1	0.02	0.02	0.02	0.02		Germany	2000	0/0/1
Bourguignon and Magnac (1990)	1	0.08	0.08	0.08	0.08		France	1985	1/0/0
Colombino and del Boca (1990)	1	0.09	0.09	0.09	0.09		Italy	1979	0/1/0
Devereux (2004)	3	−0.04	−0.06	0.00	−0.07	0.04	US	1985/80	1/2/0
Eissa and Hoynes (2004)	2	0.07	0.07	0.09	0.05	0.03	US	1990	0/2/0
Euwals and van Soest (1999)	6	0.10	0.09	0.18	0.03	0.06	Netherlands	1988	0/0/6
Flood and MaCurdy (1992)	22	0.18	0.17	0.45	−0.24	0.17	Sweden	1983	7/3/12
Hausman (1981)	2	0.02	0.02	0.03	0.00	0.02	US	1975	0/2/0
Hausman and Ruud (1984)	1	−0.03	−0.03	−0.03	−0.03		US	1976	1/0/0
MaCurdy et al. (1990)	11	−0.08	0.00	0.00	−0.22	0.10	US	1976	4/0/7
Pencavel (2002)	8	−0.02	−0.07	0.25	−0.18	0.16	US	1983	3/5/0
Triest (1990)	5	0.03	0.05	0.06	−0.02	0.03	US	1983	2/3/0
van Soest (1995)	3	0.11	0.10	0.15	0.08	0.04	Netherlands	1987	3/0/0
van Soest et al. (1990)	2	0.17	0.17	0.19	0.15	0.03	Netherlands	1985	0/1/1
Total	107	0.07	0.08	0.45	−0.24	0.12			

Obs.: Observations; Max.: Maximum; Min.: Minimum; St. Dev.: Standard deviation.

[a]Standard deviation between point estimates within a given study.

[b]Average year of data sample or year of data.

[c]Sign.: $x/y/z$: x: Number of observations significant at 5 percent level, y: observations not significant at 5 percent level, z: no statistics available.

TABLE 2 – SUMMARY STATISTICS OF STUDIES FOR THE LABOUR SUPPLY OF WOMEN INCLUDED IN THE META-SAMPLE

Author(s) (year of study)	Obs.	Mean	Median	Max.	Min.	St. Dev.[a]	Country	Year[b]	Sign.[c]
Arellano and Meghir (1992)	5	0.49	0.49	0.68	0.29	0.16	UK	1983	5/0/0
Arrufat and Zabalza (1986)	2	1.33	1.33	2.03	0.62	1.00	UK	1974	0/2/0
Bargain (2005)	4	0.29	0.30	0.37	0.20	0.07	France	1994	4/0/0
Blau and Kahn (2007)	6	0.59	0.59	0.80	0.36	0.18	US	1980/90/2000	6/0/0
Blomquist and Hansson-Brusewitz (1990)	4	0.62	0.66	0.80	0.36	0.20	Sweden	1981	3/1/0
Blundell et al. (2000)	5	0.14	0.12	0.17	0.11	0.03	UK	1990	0/5/0
Bonin and Kempe (2002)	1	0.03	0.03	0.03	0.03		Germany	2000	0/0/1
Bourguignon and Magnac (1990)	4	0.42	0.34	0.96	0.05	0.38	France	1985	4/0/0
Cogan (1981)	2	1.67	1.67	2.45	0.88	1.11	US	1966	2/0/0
Colombino and del Boca (1990)	1	2.79	2.79	2.79	2.79		Italy	1979	1/0/0
Devereux (2004)	3	0.16	0.13	0.35	0.00	0.17	US	1985/80	0/3/0
Eissa and Hoynes (2004)	2	0.25	0.25	0.44	0.07	0.26	US	1990	0/2/0
Euwals (2001)	1	0.14	0.14	0.14	0.14		Netherlands	1988	1/0/0
Euwals and van Soest (1999)	6	0.22	0.16	0.45	0.03	0.18	Netherlands	1988	0/0/6
Hausman (1981)	4	0.85	0.94	1.00	0.53	0.22	US	1975	4/0/0
Hausman and Ruud (1984)	1	0.76	0.76	0.76	0.76		US	1976	1/0/0
Kuismanen (1997)	4	0.03	0.03	0.06	0.00	0.03	Finland	1987/93	1/3/0
Mroz (1987)	23	0.13	−0.01	2.73	−0.08	0.57	US	1975	0/23/0
Triest (1990)	11	0.43	0.27	1.12	0.03	0.37	US	1983	9/2/0
van Soest (1995)	3	0.67	0.52	1.03	0.47	0.31	Netherlands	1987	3/0/0
van Soest et al. (2002)	3	1.11	1.16	1.23	0.95	0.15	Netherlands	1995	3/0/0
van Soest et al. (1990)	3	0.45	0.42	0.59	0.35	0.12	Netherlands	1985	3/0/0
Woittiez and Kapteyn (1998)	4	0.40	0.24	1.15	0.00	0.52	Netherlands	1985	0/4/0
Total	102	0.43	0.27	2.79	−0.08	0.55			
Total (excl. outliers)[d]	98	0.34	0.26	1.23	−0.08	0.35			

Obs.: Observations; Max.: Maximum; Min.: Minimum; St. Dev.: Standard deviation.

[a] Standard deviation between point estimates within a given study.

[b] Average year of data sample or year of data.

[c] Sign.: $x/y/z$: x: Number of observations significant at 5 percent level, y: observations not significant at 5 percent level, z: no statistics available.

[d] Outliers are defined as observations which differ more than two standard deviations from the mean. Practically, this means that elasticities exceeding the value of 1.52 are omitted.
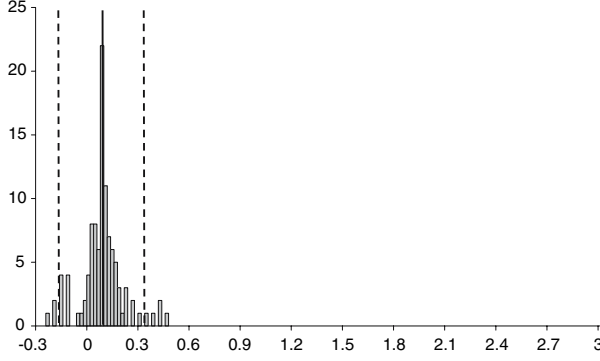
Figure 1 – Distribution of elasticities for men. *Note*: The solid line indicates the median of the empirical distribution, and the dashed lines indicate the 95% confidence interval around the mean of the distribution

Figure 2 – Distribution of elasticities for women. *Note*: The solid line indicates the median of the empirical distribution, and the dashed lines indicate the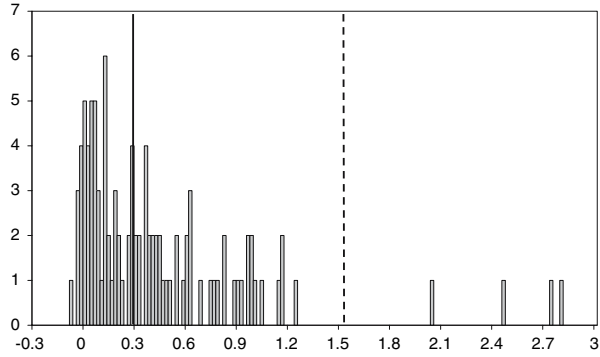 95% confidence interval around the mean of the distribution. The lower border of the confidence interval is smaller than the smallest value on the horizontal axis (−0.7), and is thus not shown in the figure

where e is the uncompensated wage elasticity, c is a constant and X is a matrix of moderator variables (see below). Study characteristics are assumed to affect the elasticity in a linear way, with slope parameters in the vector β. The error term η is assumed to be asymptotically normally distributed, independently so across different observations.[3] An OLS estimator with White heteroskedasticity-consistent standard errors will be used, as e.g. in De Mooij

---

3 It should noted that some elasticities are estimated from the same data sets, and therefore, error terms may be correlated. This argument is of limited importance for our main findings as the included empirical studies are based on many different data sources. Yet, it means that the reported standard errors may be somewhat underestimated.

and Ederveen (2003). Although weighting the observations may improve the efficiency of the estimates, it is unclear which weights should be used, and the interpretation of the effects may become more difficult (Keets and Roberts 2005). We will therefore, not put weights on the different point estimates of e.

We now turn to a brief discussion of the variables included in the moderator matrix X. The first set of variables included in the regression consists of *country dummy variables*. These dummy variables may reflect differences in institutional contexts in the different countries or cultural preferences. For France, Finland, Germany and Italy, the number of observations and studies included is too small to draw conclusions about country effects. For each of the latter three countries only one study is available in the data set (these are, respectively: Kuismanen (1997); Bonin et al. (2002); Colombino and del Boca (1990)). The second set of variables concerns the *estimation technique* used. Older studies mainly use 2SLS (Two Stage Least Squares), while more recent studies use more complex methods such as Maximum Likelihood. Two studies apply a non-parametric method (van Soest et al. (2002); Blomquist and Newey (2002)). The third set of variables indicates the *specification* that is used for the labour supply function. The different specifications imply different assumptions about the relationship between the elasticity, the wage rate and hours supplied on the labour market.

Other variables concern characteristics of the data used to estimate e, such as gender, household situation, and participation rate. *Marital status* may change the labour supply decision compared to the decision by single persons. Partners may, for instance, jointly decide on their (total) labour supply. It should be noted that only twelve observations (from two studies) for unmarried individuals were collected. The study of Euwals (2001) uses a mixed sample of both married and unmarried women, and has a separate dummy variable. Finally, the *sample participation rate* is included in order to control for the fraction of individuals who are at their extensive or intensive margins, respectively. It has been largely perceived in the literature that the decision at the extensive margin is likely to be more elastic than the decision at the intensive margin. Hence, failing to control for this variable may lead to a loss in efficiency or even biased parameters in case the sample participation rate is correlated with other moderator variables.

Estimation of equation (3.2) for men and women separately gives results displayed in Table 3. The coefficients should be interpreted as deviations from a benchmark set of study characteristics. As a benchmark, we take a study which estimates a semi-log specification for the hours function on a sample of married US residents by maximum likelihood. The estimation reported in the last column has omitted four observations from the meta sample which are more than two standard deviations away from the mean. In comparison with the other estimation for women, it can be seen that the omission of outliers implies much smaller standard errors, but that most coefficient point estimates

TABLE 3 – ESTIMATION RESULTS

	Men	Women	Women, no outliers
Finland		−0.27 (0.22)	−0.03 (0.11)
France	0.09 (0.03)	−0.29 (0.19)	−0.16 (0.16)
Germany	−0.51 (0.27)	−0.64 (0.22)	−0.46 (0.14)
Italy	0.10 (0.03)	2.12 (0.11)	
Netherlands	0.06 (0.06)	−0.26 (0.15)	−0.14 (0.12)
Sweden	0.15 (0.04)	0.01 (0.17)	0.05 (0.16)
United Kingdom		0.24 (0.14)	0.13 (0.09)
Non-parametric estimation	0.26 (0.10)	0.65 (0.16)	0.73 (0.15)
2SLS estimation	−0.11 (0.05)	−0.20 (0.22)	0.10 (0.10)
Simulation	−0.42 (0.10)	0.32 (0.30)	0.40 (0.15)
Linear specification	−0.10 (0.04)	0.21 (0.22)	0.43 (0.12)
Quadratic specification	−0.11 (0.04)	0.51 (0.28)	0.69 (0.18)
Unmarried	−0.01 (0.04)	−0.16 (0.16)	−0.08 (0.15)
Unmarried and married		0.08 (0.26)	0.21 (0.16)
Participation rate	−2.20 (0.93)	−0.26 (0.21)	−0.14 (0.20)
Constant	2.29 (0.92)	0.56 (0.27)	0.20 (0.14)
R-squared	0.38	0.36	0.42
Observations	107	102	98

Standard errors between parentheses (White heteroskedasticity-consistent).
Benchmark for dummy variables: US, Maximum Likelihood, Semi log-specification, Married.

remain qualitatively the same. Note that we cannot identify a separate country coefficient for Italy when the outliers are omitted.

Table 3 shows that elasticities for Germany appear to be systematically smaller than elasticities for other countries. This holds both for men and for women. For men, elasticities in France, Italy, and Sweden appear to be somewhat higher than in the Netherlands and the US. However, the sample only includes one single observation per country for France, Germany, and Italy, so that we should be careful drawing conclusions on these country specific effects. It is nevertheless remarkable that the Swedish elasticity for men looks systematically higher than that in other countries.

Further, we find no evidence that maximum likelihood (ML) estimation generates relatively high elasticity estimates, as was suggested earlier by MaCurdy et al. (1990).[4] Instead, ML estimates appear to be structurally

4 MaCurdy et al. (1990) claimed that the high elasticities found in articles using the Hausman model (estimated with Maximum Likelihood) are a result of the strong restrictions imposed in the model. Ecklöf and Sacklén (2000) later played down this claim, and argued that the findings of MaCurdy et al. (1990) were caused by flaws in their data.

TABLE 4 – PREDICTED ELASTICITIES

	Men	Women	Women, no outliers
Netherlands	0.12	0.49	0.49
United Kingdom		0.48	0.34
United States	−0.01	0.38	0.29
Sweden	0.12	0.62	0.62

lower than non-parametric estimates. Another interesting point from Table 3 is that the specification of the hours equation appears to impact the elasticity estimates. For men, it is found that a semi-log specification for the hours function generates higher point estimates for the elasticity than other specifications. For women, the opposite is found.

For men, the sample participation rate has a negative impact on the estimated elasticity, which is consistent with individuals being more elastic at the extensive margin than at the intensive margin. A negative coefficient is also found for women, but this is not significantly different from zero. Thus, on the basis of these estimations we cannot reject the hypothesis that women are as elastic at the extensive margin as they are at the intensive margin.

The estimated coefficients in Table 3 allow us to predict elasticities for different countries. In particular, we use the meta regression and then fix dummy variables at country-specific levels to compute particular elasticities, thereby holding other variables at their sample means (specification and estimation method). A selection of the elasticities thus obtained is presented in Table 4. Clearly, predicted elasticities are larger for women than for men in all countries. For instance, the elasticity for Dutch men equals 0.12 and for women it equals 0.49. It has been noted in the literature that this difference may be attributed to the fact that women are relatively more often located at the extensive margin (i.e. they are not participating in the labour market), and that the elasticity for women will converge towards the male level as the female participation rate goes up. In fact, Mroz (1987) concludes that married women and prime aged males are equally sensitive at the intensive margin. This is however not supported by our findings. If we impute the male participation rate into the meta regression for women, then we predict a female elasticity in the Netherlands of 0.44 or 0.47 (depending on which sample is used). Similar results are found for the other countries (not reported), be it that elasticities for the US are about 0.10 lower than for the Netherlands, while for Swedish women a higher elasticity is found.

Controlling for unobserved heterogeneity can be problematic in meta analysis. Indeed, consistency heavily depends on the ability to observe all relevant factors that determine the elasticity. We will test for the potential

TABLE 5 – ESTIMATION RESULTS FOR THE EXTENDED SPECIFICATIONS

	Men	Women	Women, no outliers
Study characteristics[a]			
Panel data	0.09 (0.06)	−0.08 (0.37)	0.35 (0.19)
In refereed journal	−0.06 (0.07)	−0.72 (0.45)	−0.15 (0.22)
Actual-desired-hours	−0.10 (0.01)	−0.12 (0.15)	−0.01 (0.12)
Measurement error	−0.01 (0.05)	0.04 (0.23)	−0.14 (0.20)
Control variables[a]			
Child	0.00 (0.05)	−1.18 (0.79)	−0.79 (0.05)
Child younger than age 6	−0.02 (0.06)	1.15 (0.83)	1.03 (0.22)
Education	−0.05 (0.03)	−0.27 (0.15)	−0.23 (0.15)
Family	−0.16 (0.04)	−0.53 (0.46)	−0.16 (0.33)
Family size	0.02 (0.04)	−2.26 (0.85)	−1.26 (0.28)
Health	0.00 (0.12)		
Age	−0.08 (0.04)	0.24 (0.34)	0.55 (0.16)
Age Squared		−0.38 (0.35)	−0.09 (0.04)
Study effects[b]	Yes	Yes	Yes
R-squared	0.50	0.61	0.69
Observations	107	102	98

Standard errors between parentheses (White heteroskedasticity-consistent).
[a] Both 'study characteristics' and 'control variables' can only take on the values 0 and 1, indicating whether the study characteristic applies and whether the control variable has been included in the specification, respectively.
[b] Study fixed effects are only included for studies with at least 8 observations. It was not possible to include fixed effects for studies with less than 8 observations as a result of multicollinearity.

omission of moderator variables. The results of the extended regressions are presented in Table 5. Note that coefficients for moderator variables reported in Table 3 are not reported in Table 5.[5]

The regression includes four dummy variables representing study characteristics. These refer to (i) the effect of using panel data instead of cross section data, (ii) whether the study has been published in a refereed journal, (iii) whether desired hours of work are used instead of actual hours of work and (iv) whether measurement error is explicitly taken into account for the observed hours of work. Furthermore, we include dummy variables which indicate whether certain control variables are used. These include age (-squared), education (-squared), health, and the presence of children. We include fixed effects per study for studies with at least 8 observations. This implies that the coefficients reported in Table 5 are predominantly identified from the 'within variation' of studies.

5 The complete results are available upon request.

Table 5 shows that most study characteristics do not have significant impact on the elasticity estimates. It is interesting to see that studies published in refereed journals tend to publish smaller elasticities than others, but this effect is not statistically significant. Controlling for the family situation and age significantly affects elasticity estimates, and for women controlling for the presence of children appears to be important as well.

While Table 5 shows the significance of each moderator variable separately, Table 6 shows results for redundant variable tests for sets of moderator variables. It appears that both the study characteristics and control variables listed in Table 5 are important in estimating labour supply elasticities. Country effects are important for women, but not statistically significant for men. The specification of the hours function and the estimation method are statistically significant at the 5% confidence level for women, but not for men. In this specification, the participation effect is not statistically significant anymore, while it was so in the earlier estimation in Table 3.

Table 7 shows the estimated coefficients for some alternative samples, using the same model specification as in Table 3. The first and second column are based on samples that contain only observations from studies published in refereed journals. Comparing estimation results with those in Table 3, we see that the main difference is that both the estimation method and the specification chosen do not have a significant impact anymore on the estimates for men. The third and fourth column of Table 7 are based on samples where observations with identical characteristics are combined. That is, whenever multiple elasticities were reported within a study for a given country, gender, estimation method, specification, etc., we included just one value, being the *average* point estimate. Following this procedure, we reduce the sample to 26 and 28 observations for men and women, respectively. Again, it is found that both the estimation method and the specification chosen do not have a significant impact anymore on the estimates for men. The smaller sample size

TABLE 6 – REDUNDANT VARIABLE TESTS

	Men		Women, no outliers	
	F-statistic	Probability	F-statistic	Probability
Country	1.7	0.14	10.3	0.00
Estimation	2.6	0.06	7.4	0.00
Specification	2.7	0.07	7.1	0.00
Marital status	25.0	0.00	4.4	0.02
Participation	0.6	0.43	0.1	0.73
Characteristics	28.3	0.00	5.3	0.00
Control variables	73.9	0.00	153.0	0.00
Study effects	0.2	0.89	6.9	0.00

TABLE 7 – ESTIMATION RESULTS FOR ALTERNATIVE SAMPLES

	Published in refereed journal		Independent observations	
	Men	Women, no outliers	Men	Women, no outliers
Finland				−0.08 (0.35)
France	0.12 (0.03)	−0.12 (0.20)	0.10 (0.03)	−0.32 (0.34)
Germany			0.00 (0.73)	−0.47 (0.45)
Italy	0.13 (0.03)		0.11 (0.03)	
Netherlands	0.27 (0.11)	−0.14 (0.12)	0.18 (0.15)	−0.08 (0.33)
Sweden	0.12 (0.03)	0.05 (0.16)	0.16 (0.04)	−0.23 (0.32)
United Kingdom		0.26 (0.09)		0.17 (0.25)
Non-parametric estimation	−0.03 (0.20)	0.73 (0.17)	0.07 (0.30)	0.67 (0.38)
2SLS estimation	−0.01 (0.06)	−0.04 (0.08)	0.00 (0.04)	−0.10 (0.35)
Simulation	−0.01 (0.21)	0.41 (0.18)	−0.19 (0.28)	0.29 (0.47)
Linear specification	−0.03 (0.07)	0.44 (0.11)	−0.02 (0.05)	0.33 (0.33)
Quadratic specification	−0.01 (0.07)	0.70 (0.18)	−0.03 (0.05)	0.72 (0.40)
Unmarried	0.03 (0.04)	−0.08 (0.17)	−0.04 (0.07)	−0.36 (0.38)
Unmarried and married		0.22 (0.15)		0.06 (0.37)
Participation rate	0.96 (1.82)	−0.14 (0.20)	−0.53 (2.63)	−0.38 (0.42)
Constant	−0.97 (1.82)	0.19 (0.14)	0.54 (2.59)	0.52 (0.41)
R-squared	0.39	0.49	0.66	0.54
Observations	78	83	26	28

Standard errors between parentheses (White heteroskedasticity-consistent).

however plays a role in this, resulting in relatively high standard errors. For women this is even worse, so we cannot draw any strong conclusions.

5 CONCLUSION

This paper reviews empirical findings on the magnitude (and sign) of the uncompensated wage elasticity of labour supply (that is, the percentage change in hours worked as a result of a one percent change in the net hourly wage rate). By using meta regressions, we also aim to identify the sources of variation in the empirical literature.

We develop a meta sample of 209 elasticities drawn from 30 empirical studies. On the basis of our meta regression, we predict the uncompensated wage elasticity of labour supply for Dutch women at around 0.5. The corresponding figure for men is predicted at 0.1. Predictions for Sweden, the UK, and the US are quite similar, although the elasticities in the UK and the US appear somewhat lower and those in Sweden somewhat higher.

Regarding the variation in study results, earlier studies principally focus on model specification and estimation techniques and in a partial setting. We

add to this by exploring a broader set of potential sources of variation and by performing a simultaneous meta analysis. Our results confirm that both the model specification and the estimation method impact elasticity estimates. Controlling for the family situation and age significantly affects estimates of the labour supply elasticity. For women, controlling for the presence of children is also found to affect such estimates. Finally, we find that women have a larger labour supply elasticity than men, even after taking into account participation rates. It is thus suggested that the elasticity at the intensive margin is more elastic for women than for men.

REFERENCES

Alesina, A., E. Glaeser and B. Sacerdote (2005), 'Work and Leisure in the U.S. and Europe; Why so different,' NBER Working Paper no. 11278.

Arellano, M. and C. Meghir (1992), 'Female Labour Supply and on the Job search: An Empirical Model Estimated Using Complementary Data Sets,' *Review of Economic Studies*, 59, pp. 537–559.

Arrufat, J. and A. Zabalza (1986), 'Female Labor Supply with Taxation, Random Preferences, and Optimization Errors,' *Econometrica*, 54(1), pp. 47–64.

Ballard, C., J. Shoven and J. Whalley (1985), 'General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States,' *American Economic Review*, 75(1), pp. 128–137.

Bargain, O. (2005), 'On Modelling Household Labor Supply with Taxation,' IZA Working Paper 1455, Bonn, IZA.

Blau, F. and L. Kahn (2005), 'Changes in the Labor Supply Behavior of Married Women: 1980–2000,' Working Paper, 11230, NBER.

Blau, F. and L. Kahn (2007), 'Changes in the Labor Supply Behavior of Married Women: 1980–2000,' *Journal of Labor Economics*, 25, pp. 393–438.

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

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

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

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

Blundell, R. and T. MaCurdy (1999), 'Labor Supply: A Review of alternative approaches,' in: O. Ashenfelter and D. Card (eds.), '*Handbook of Labor Economics*,' Vol. 3A, Ch. 27, Amsterdam, North Holland.

Blundell, R., A. Duncan and C. Meghir (1998), 'Estimating Labor Supply Responses Using Tax Reforms,' *Econometrica*, 66, pp. 827–861.

Blundell, R., A. Duncan, J. McCrae and C. Meghir (2000), 'The Labour Market Impact of the Working Families Tax Credit,' *Fiscal Studies*, 21, pp. 75–104.

Bonin, H., W. Kempe and H. Schneider (2002), 'Household Labor Supply Effects of Low-Wage Subsidies in Germany,' Discussion Paper, 637, IZA.

Bourguignon, F. and T. Magnac (1990), 'Labor Supply and Taxation in France,' *Journal of Human Resources*, 25, pp. 358–389.

Browning, E. (1987), 'On the Marginal Welfare Cost of Taxation,' *American Economic Review*, 77(1), pp. 11–23.

Burtless, G. and J. Hausman (1978), 'The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Experiment,' *American Economic Review*, 72, pp. 488–479.

Cogan, J. (1981), 'Fixed Costs and Labor Supply,' *Econometrica*, 49, pp. 945–963.

Colombino, U. and D. del Boca (1990), 'The Effect of Taxes on Labor Supply in Italy,' *Journal of Human Resources*, 25, pp. 390–414.

de Mooij, R. and S. Ederveen (2003), 'Taxation and Foreign Direct Investment: A Synthesis of Empirical Research,' *International Tax and Public Finance*, 10, pp. 673–693.

Devereux, P. (2004), 'Changes in Relative Wages and Family Labor Supply,' *Journal of Human Resources*, 39, pp. 696–722.

Douglas, P. (1934), *Theory of Wages*, Macmillan, New York.

Ecklöf, M. and H. Sacklén (2000), 'The Hausman–MaCurdy Controversy. Why Do Results Differ between Studies? *Journal of Human Resources*, 35, pp. 204–220.

Eissa, N. and H. Hoynes (2004), 'The Hours of Work Response of Married Couples: Taxes and the Earned Income Tax Credit,' *Tax Policy and Labor Market Performance*, (forthcoming).

Ericson, P. and L. Flood (1997), 'A Monte Carlo Evaluation of Labor Supply Models,' *Empirical Economics*, 22, pp. 431–460.

Euwals, R. (2001), 'Female Labour Supply, Flexibility of Working Hours and Job Mobility,' *Economic Journal*, 111, pp. 2.120–2.134.

Euwals R. and A. van Soest (1999), 'Desired and Actual Labour Supply of Unmarried Men and Women in the Netherlands,' *Labour Economics*, 6, pp. 95–118.

Flood, L. and T. MaCurdy (1992), 'Work Disincentive Effects of Taxes: An Empirical Analysis of Swedish Men,' *Carnegie–Rochester Conference Series on Public Policy*, 37, pp. 239–278.

Graafland, J., R. de Mooij, A. Nibbelink and A. Nieuwenhuis (2001), *Mimicing Tax Policies and the Labour Market*, North Holland.

Hall, R. (1973), 'Wages, Income and Hours of Work,' in: G.G. Cain and H.W. Watts (eds.), *Income Maintenance and Labor Supply*, Institute for Research on Poverty Monograph Series.

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

Hausman, J. (1981), 'The Effect of Taxes on Labour Supply,' in: H. Aaron and J. Pechman (eds.), *How Taxes Affect Economic Behavior*, Brookings, Washington D.C.

Hausman, J. and P. Ruud (1984), 'Family Labor Supply with Taxes,' *American Economic Review*, 74, pp. 242–248.

Heckman, J. (1979), 'Sample Selection Bias as a Specification Error,' *Econometrica*, 46, pp. 931–959.

Heim, B. and B. Meyer (2003), 'Structural Labor Supply Models when Budget Constraints are Nonlinear,' Working Paper, Duke University, Northwestern University.

Keef, S. and L. Roberts (2004), 'The Meta-Analysis of Partial Effect Sizes,' *British Journal of Mathematical and Statistical Psychology*, 57, pp. 97–129.

Killingsworth, M. and J. Heckman (1986), 'Female Labor Supply: A Survey,' in: O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, Vol. I, North-Holland, Amsterdam, pp. 103–204.

Kosters, M. (1966), 'Effects of an Income Tax on Labor Supply,' in: A. Harberger and J. Martin (eds.), *The Taxation of Income from Capita*, Studies of Government Finance, Washington, D.C.

Kuismanen, M. (1997), 'Labour Supply, Unemployment and Income Taxation: An Empirical Application,' Working Paper, Government Institute for Economic Research, Helsinki, University College London.

MaCurdy, T., P. Green and H. Paarsch (1990), 'Assessing Empirical Approaches for Analyzing Taxes and Labor Supply,' *Journal of Human Resources*, 25, pp. 415–490.

Mincer, J. (1963), 'Labour Force Participation of Married Women: A Study of Labor Supply, Aspects of Labor Economics,' *Universities-National Bureau Conference Series No.* 14, pp. 63–105.

Moffitt, R. (1990), 'The Econometrics of Kinked Budget Constraints,' *Journal of Economic Perspectives*, 4(2), pp. 119–139.

Mroz, T. (1987), 'The Sensitivity of an Empirical Model of Married Woman's Hours of Work to Economic and Statistical Assumptions,' *Econometrica*, 55, pp. 765–799.

Pencavel, J. (1986), 'Labor Supply of Men,' in: O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, North-Holland.

Pencavel, J. (2002), 'A Cohort Analysis of the Association between Work Hours and Wages among Men,' *Journal of Human Resources*, 37, pp. 251–274.

Prescott, E. (2004), 'Why Do Americans Work So Much More than Europeans,' NBER Working Paper no. 10316.

Saez, E. (2003), 'The Effect of Marginal Tax Rates on Income: A Panel Study of 'Bracket Creep,' *Journal of Public Economics*, 87, pp. 1231–1258.

van Soest, A., I. Woittiez and A. Kapteyn (1990), 'Labor Supply, Income Taxes, and Hours Restrictions in the Netherlands,' *Journal of Human Resources*, 25, pp. 517–558.

Triest, R. (1990), 'The Effect of Income Taxation on Labor Supply in the United States,' *Journal of Human Resources*, 25, pp. 491–516.

van Soest, A. (1995), 'Structural Models of Family Labor Supply: A Discrete Choice Approach,' *Journal of Human Resources*, 30, pp. 63–88.

van Soest, A., M. Das and X. Gong (2002), 'A Structural Labour Supply Model with Flexible Preferences,' *Journal of Econometrics*, 107, pp. 345–374.

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