



TI 2006-017/3

Tinbergen Institute Discussion Paper

# What Explains the Variation in Estimates of Labour Supply Elasticities?

*Michiel Evers<sup>1</sup>*

*Ruud A. de Mooij<sup>1,2,3,4</sup>*

*Daniël J. van Vuuren<sup>3,5</sup>*

<sup>1</sup> *Erasmus Universiteit Rotterdam,*

<sup>2</sup> *Tinbergen Institute,*

<sup>3</sup> *CPB, The Hague,*

<sup>4</sup> *CESifo,*

<sup>5</sup> *Free University Amsterdam.*

**Tinbergen Institute**

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

**Tinbergen Institute Amsterdam**

Roetersstraat 31

1018 WB Amsterdam

The Netherlands

Tel.: +31(0)20 551 3500

Fax: +31(0)20 551 3555

**Tinbergen Institute Rotterdam**

Burg. Oudlaan 50

3062 PA Rotterdam

The Netherlands

Tel.: +31(0)10 408 8900

Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to [driessen@tinbergen.nl](mailto:driessen@tinbergen.nl).

Most TI discussion papers can be downloaded at <http://www.tinbergen.nl>.

# **What Explains the Variation in Estimates of Labour Supply Elasticities?<sup>1</sup>**

**Michiel Evers**  
**Erasmus University Rotterdam**

**Ruud A. de Mooij**  
**Erasmus University Rotterdam, Tinbergen Institute, CPB, CESifo**

**Daniël J. van Vuuren**  
**Free University Amsterdam and CPB**

## **Abstract**

This paper performs a meta-analysis of empirical estimates of uncompensated labour supply elasticities. We find that much of the variation in elasticities can be explained by the variation in gender, participation rates, and country fixed effects. Country differences appear to be small though. There is no systematic impact of the model specification or marital status on reported elasticities. The decision to participate is more responsive than is the decision regarding hours worked. Even at the intensive margin, we find that the elasticity for women exceeds that for men. For men and women in the Netherlands, we predict an uncompensated labour supply elasticity of 0.1 (or 0.2 if an alternative specification is preferred) and 0.5, respectively. These values are robust for alternative samples and specifications of the meta regression.

*Key words: labour supply, meta-analysis, uncompensated elasticity.*

*JEL code: J22; H2.*

<sup>1</sup> The authors thank Sjef Ederveen, Rob Euwals, Henri de Groot, Egbert Jongen, Pierre Koning, Cees Withagen, and in particular Arthur van Soest and Isolde Woittiez for useful comments and discussions.

# 1 Introduction

The elasticity of labour supply with respect to the wage rate 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 (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 bigger is the employment effect in response to a change in the tax rate and the higher is the excess burden of taxation.

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 value 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. Alessina 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 robust 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 (2005). Mroz (1987) examines the effects of economic and statistical assumptions on outcomes for married women. He finds 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. His 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. Hence, 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 impact of the various factors on the reported empirical estimates simultaneously. In particular, we use a sample of 239 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 primary 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, therefore, 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. Section 3 explores the sources of variation in more detail by performing a meta analysis. Section 4 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.<sup>2</sup> 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 IV are Mroz (1987), Blomquist (1996), Pencavel (2002), Eissa and Hoynes (2004) and Blau and Kahn (2005).

Another 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

<sup>2</sup> See Moffitt (1990) for a comprehensive survey.

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 90s, 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 and Woittiez (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. 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 and Kempe (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.<sup>3</sup> 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.

<sup>3</sup> Of course, non-generic time effects and endogenous switching can be allowed if the econometrician is capable of correcting for these.

## 3 Meta-analysis

### 3.1 The meta sample

This section aims to identify the sources of variation in empirical estimates of the uncompensated labour supply elasticity in the literature. To that end, we first construct a ‘meta sample’ using empirical estimates of the elasticity found in the literature. Subsequently, the variation in reported elasticities is explained by the variation in study characteristics of the underlying primary studies, i.e. we run meta regressions. Our focus on the uncompensated labour supply elasticity is governed by the availability of research findings. In particular, we cannot obtain separate values from many primary studies 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 selected studies which either explicitly or implicitly report one or more estimates. It is however emphasised that the selected meta-sample 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.

Explicit elasticities can be obtained from studies that use the double log specification. If the elasticity is not reported explicitly, it is still possible to construct a consistent estimate from reported point estimates of marginal coefficients if we use 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  $\beta$  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  $x$  read as follows:

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

where  $\phi_w$  is the derivative of the hours function  $\phi$  with respect to the wage rate  $w$ . After substituting sample means for  $w$ ,  $Y$  and  $x$ , a consistent estimator of  $e$  at the sample mean is obtained (i.e. for an imaginary individual whose characteristics precisely match the sample mean). Apart from elasticity values, we also collect information on standard errors of the estimated elasticities. Yet, it is impossible to retrieve consistent estimates of standard errors of  $e$  as long as the estimated covariance matrix for  $\beta$  is unknown. Unfortunately this is often the case since primary studies usually do not report full covariance matrices. A straightforward simplification is the assumption that off-diagonal elements cancel out, so that the Delta method can be applied:

$$(3.2) \sigma_e^2 = \frac{w^2}{h^2} \cdot \left[ \frac{\partial \phi_w}{\partial \beta} \right] \Sigma_\beta \left[ \frac{\partial \phi_w}{\partial \beta} \right]'$$

where  $\Sigma_\beta$  denotes the estimated covariance matrix for  $\beta$  with off-diagonal elements set to zero, and  $\partial \phi_w / \partial \beta$  is the row vector with derivatives of  $\phi_w$  with respect to the parameters in  $\beta$ . The matrix  $\Sigma_\beta$  is reported in nearly all studies; in case the entire (original) matrix is reported it can be substituted into (3.2) to apply the ‘real’ Delta method (e.g., p. 297 in Greene, 1993).

In constructing a meta sample, a number of studies cannot be used because of missing sample statistics or because the study does not allow us to compute elasticity values. For instance, many studies based on tax reforms do not report uncompensated labour supply elasticities or information to compute this figure. Hence, these studies could not be used in our sample. Ultimately, our literature search yields a sample of 239 elasticities obtained from 32 studies. Table 3.1 presents some summary indicators from the sample. We see that the mean value of the elasticity in the sample is 0.24, with mean values for men and women of respectively 0.07 and 0.41. The table shows great variation across different studies: the mean of the elasticities ranges from -0.24 to 2.79. The difference between males and females is apparent. The range for men is from -0.24 to 0.13, while for women it is from -0.19 to 2.79.<sup>4</sup> The number of elasticities per study varies from one (Burtless and Hausman, 1978) to 25 (Mroz, 1987). The last column in Table 3.1 shows that although most estimates of the elasticities are significantly different from zero at a five percent significance level, a reasonable number is not or no information was supplied by the authors.

**Table 3.1 Summary statistics of studies in the meta-sample**

Author(s) (year of study)	Gender	Obs.	Mean	Median	Max.	Min.	St. Dev. <sup>1</sup>	Country	Year <sup>2</sup>	Sign. <sup>3</sup>
Arellano, Meghir (1992)	female	5	0.49	0.49	0.68	0.29	0.16	UK	1983	5/0/0
Arrufat, Zabalza (1986)	female	2	1.33	1.33	2.03	0.62	1.00	UK	1974	0/2/0
Bargain (2005)	female	4	0.29	0.30	0.37	0.20	0.07	France	1994	4/0/0
Blau, Kahn (2005)	female	12	0.54	0.56	0.80	0.31	0.16	US	1980/90/2000	12/0/0
Blau, Kahn (2005)	male	6	0.07	0.07	0.13	0.01	0.04	US	1980/90/2000	6/0/0
Blomquist (1983)	male	2	0.08	0.08	0.08	0.08	0.00	Sweden	1973	2/0/0
Blomquist (1996)	male	4	0.00	-0.02	0.18	-0.13	0.15	Sweden	1981	4/0/0
Blomquist, Hansson-Brusewitz (1990)	female	4	0.62	0.66	0.80	0.36	0.20	Sweden	1981	3/1/0
Blomquist, Hansson-Brusewitz (1990)	male	4	0.10	0.10	0.13	0.08	0.03	Sweden	1981	3/1/0
Blomquist, Newey (2002)	male	24	0.08	0.08	0.12	0.04	0.02	Sweden	1982	24/0/0
Blundell et al. (2000)	female	5	0.14	0.12	0.17	0.11	0.03	UK	1990	0/5/0
Bonin, Kempe (2002)	female	1	0.03	0.03	0.03	0.03		Germany	2000	0/0/1
<b>Table 3.2 (continued)</b>	<b>Summary statistics of studies in the meta-sample</b>									
Bonin, Kempe (2002)	male	1	0.02	0.02	0.02	0.02		Germany	2000	0/0/1

<sup>4</sup> Saez (2003) reports elasticities based on a sample containing both men and women and it is therefore classified as ‘both’

Bourguignon, Magnac (1990)	female	5	0.30	0.30	0.96	-0.19	0.43	France	1985	5/0/0
Bourguignon, Magnac (1990)	male	2	-0.02	-0.02	0.08	-0.13	0.14	France	1985	2/0/0
Burtless, Hausman (1978)	male	1	0.00	0.00	0.00	0.00		US	1972	0/1/0
Cogan (1981)	female	2	1.67	1.67	2.45	0.88	1.11	US	1966	2/0/0
Colombino, Del Boca (1990)	female	1	2.79	2.79	2.79	2.79		Italy	1979	1/0/0
Colombino, Del Boca (1990)	male	1	0.09	0.09	0.09	0.09		Italy	1979	0/1/0
Devereux (2003)	male	8	0.18	0.18	0.21	0.16	0.02	US	1980/90	8/0/0
Devereux (2004)	female	3	0.16	0.13	0.35	0.00	0.17	US	1985/80	0/3/0
Devereux (2004)	male	3	-0.04	-0.06	0.00	-0.07	0.04	US	1985/80	1/2/0
Eissa, Hoynes (2004)	female	3	0.17	0.07	0.44	0.02	0.23	US	1990	1/2/0
Eissa, Hoynes (2004)	male	3	0.02	0.05	0.09	-0.07	0.08	US	1990	0/3/0
Euwals (2001)	female	1	0.14	0.14	0.14	0.14		Netherlands	1988	1/0/0
Euwals, Van Soest (1999)	female	6	0.22	0.16	0.45	0.03	0.18	Netherlands	1988	0/0/6
Euwals, Van Soest (1999)	male	6	0.10	0.09	0.18	0.03	0.06	Netherlands	1988	0/0/6
Flood, MaCurdy (1992)	male	22	0.18	0.17	0.45	-0.24	0.17	Sweden	1983	7/3/12
Hausman (1981)	female	4	0.85	0.94	1.00	0.53	0.22	US	1975	4/0/0
Hausman (1981)	male	2	0.02	0.02	0.03	0.00	0.02	US	1975	0/2/0
Hausman, Ruud (1984)	female	1	0.76	0.76	0.76	0.76		US	1976	1/0/0
Hausman, Ruud (1984)	male	1	-0.03	-0.03	-0.03	-0.03		US	1976	1/0/0
Kuismanen (1997)	female	4	0.03	0.03	0.06	0.00	0.03	Finland	1987/93	1/3/0
MaCurdy et al. (1990)	male	11	-0.08	0.00	0.00	-0.22	0.10	US	1976	4/0/7
Mroz (1987)	female	25	0.12	-0.01	2.73	-0.08	0.55	US	1975	0/25/0
Pencavel (2002)	male	8	-0.02	-0.07	0.25	-0.18	0.16	US	1983	3/5/0
Saez (2003)	both	8	0.15	0.01	1.30	-0.22	0.49	US	1980	1/7/0
Triest (1990)	female	11	0.43	0.27	1.12	0.03	0.37	US	1983	9/2/0
Triest (1990)	male	5	0.03	0.05	0.06	-0.02	0.03	US	1983	2/3/0
Van Soest (1995)	female	3	0.67	0.52	1.03	0.47	0.31	Netherlands	1987	3/0/0
Van Soest (1995)	male	3	0.11	0.10	0.15	0.08	0.04	Netherlands	1987	3/0/0
Van Soest et al. (2002)	female	3	1.11	1.16	1.23	0.95	0.15	Netherlands	1995	3/0/0
Van Soest et al. (1990)	female	3	0.45	0.42	0.59	0.35	0.12	Netherlands	1985	3/0/0
Van Soest et al. (1990)	male	2	0.17	0.17	0.19	0.15	0.03	Netherlands	1985	0/1/1
Woittiez, Kapteyn (1998)	female	4	0.40	0.24	1.15	0.00	0.52	Netherlands	1985	0/4/0
All studies	female	112	0.41	0.28	2.79	-0.19	0.53			
"	male	119	0.07	0.08	0.45	-0.24	0.12			
"	all	239	0.24	0.10	2.79	-0.24	0.42			

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

<sup>1</sup> Standard deviation between point estimates within a given study

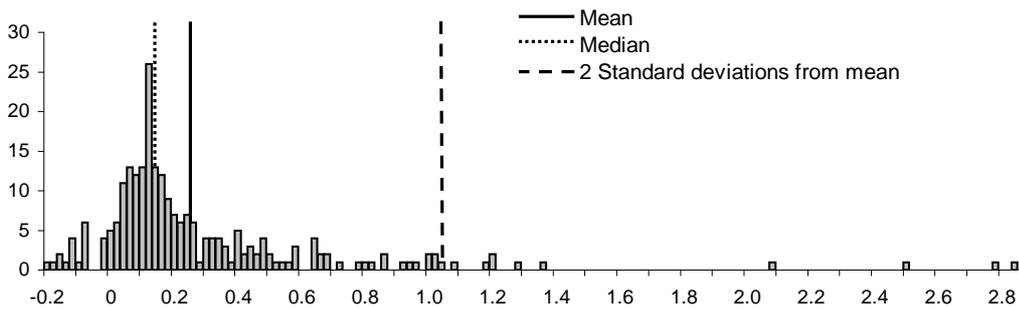
<sup>2</sup> Average year of data sample or year of data

<sup>3</sup> 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

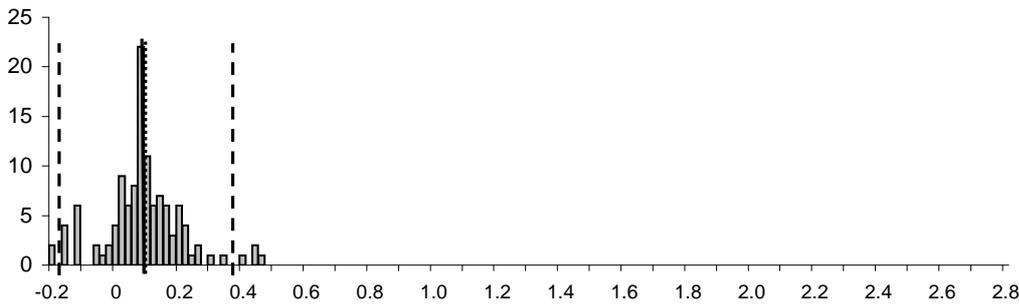
Figure 3.1 shows the empirical distribution of the 239 point estimates observed. Figure 3.2 and 3.3 show empirical distributions for the sub samples of men and women, respectively. The dashed lines correspond with the borders of the interval with twice the standard deviation

around the mean. In Figure 3.1 and 3.3, the left border of the interval is smaller than the minimum value and therefore not shown in the figures. Figure 3.3 shows that the median (0.28) for women is to the left of the mean (0.41), as a result of some extreme values in the right tail. For men, the mean (0.072) and the median (0.076) are more similar. Clearly, the variation in point estimates for men is much smaller than that for women.

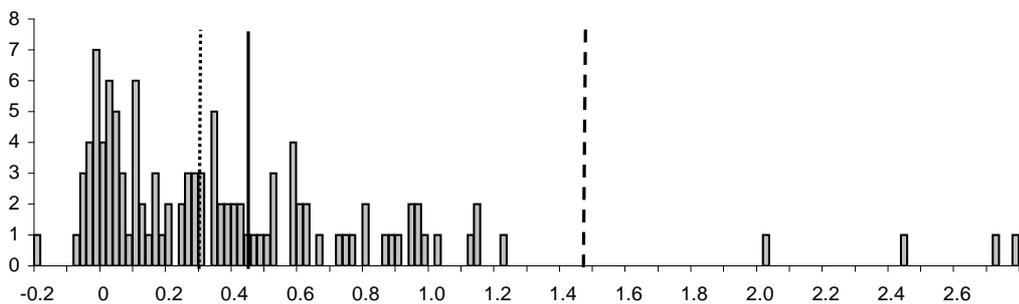
**Figure 3.1** Distribution of elasticities



**Figure 3.2** Distribution of elasticities for men



**Figure 3.3** Distribution of elasticities for women



### 3.2 The base regression

The meta-analysis takes the form of the following regression:

$$(3.3) e = c + X\beta + \eta$$

where  $e$  is the uncompensated elasticity,  $c$  is a constant and  $X$  is a matrix of moderator variables (see below). Study characteristics affect the elasticity in a linear way, with slope parameters in the vector  $\beta$ . The error term  $\eta$  is assumed to be asymptotically normally distributed, independently so across different observations.<sup>5</sup> An OLS estimator with White heteroskedasticity-consistent standard errors will be used, as e.g. in De Mooij 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 (Keef and Roberts, 2005). We will therefore not put weights on the different point estimates of  $e$ .

We now briefly discuss 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 and Kempe, 2002; Colombino and Del Boca, 1990). The second set of variables concerns the *estimation technique* used. Older studies mainly use OLS and 2SLS, while more recent studies use more complex methods such as Maximum Likelihood. Two included studies use 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 relation between the elasticity, the wage rate and labour hours supplied.

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 variable *mixed study* concerns studies that estimate the labour supply decision for a sample containing both married and unmarried individuals. There are two studies with a mixed sample, Euwals (2001) and Devereux (2003). A third category, labelled *both sexes* only concerns the study by Saez (2003) who uses a mixed sample of both men and women. 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

<sup>5</sup> It should be 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 primary studies are based on many different data sources. Yet, it means that the reported standard errors may be somewhat underestimated.

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. The participation rate is interacted with gender in order to allow for different effects for males and females, implicitly recognising that gender may be a determinant of the labour supply decision at the extensive margin.

Estimation of equation (3.3) gives results displayed in Table 3.2. The coefficients should be interpreted as deviations from a benchmark set of study characteristics. As a benchmark, we take a study for the US, using a maximum likelihood estimator, a double log specification, for male workers. The difference between the specifications in column (1) and (2) is the inclusion of sample participation rates in the latter. The estimation reported in column (3) has omitted 4 observations from the meta sample which are more than three standard deviations away from the mean. In comparison with the other estimations, it can be seen that the omission of outliers implies much smaller standard errors, but that most coefficient point estimates remain qualitatively the same.

Regarding countries, we read from Table 3.2 that the point estimates for France, Germany, the Netherlands, and the United Kingdom are close to zero, and never significantly different from this value. Otherwise stated, we are unable to reject the null hypothesis that labour supply elasticities in these mentioned countries differ from the elasticity for the United States. More in general, hardly any evidence can be found supporting the hypothesis that elasticities differ between countries.

Further, it is interesting to see that Maximum Likelihood estimation appears to generate relatively high elasticity estimates, as was suggested earlier by MaCurdy et al. (1990).<sup>6</sup> Indeed, other estimators typically produce smaller elasticities. This effect is however insignificant, i.e. we cannot reject the hypothesis that different estimation methods generate structurally different estimates. Another interesting point from Table 3.2 is that the specification of the hours equation does not appear to have much impact on the elasticity estimates. The only exception is the quadratic specification which might have a positive impact compared to e.g. the log-linear specification. A similar finding was reported by Ericson and Flood (1997).

As expected, 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. It is also apparent from Table 3.2 that there is a significant difference between males and females. Yet, the extent of this difference cannot be directly read from the parameter estimates reported, because of the interaction with the gender-specific participation rates. We

<sup>6</sup> 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.

will therefore explore this difference in more detail by ‘predicting’ elasticities for males and females from our estimation results.

**Table 3.3 Estimation results**

	(1)	(2)	(3)
Finland	-0.42 (0.14)	-0.26 (0.13)	-0.20 (0.09)
France	-0.07 (0.11)	-0.06 (0.11)	-0.04 (0.12)
Germany	-0.12 (0.18)	-0.23 (0.12)	-0.18 (0.12)
Italy	1.19 (0.86)	1.14 (0.79)	0.07 (0.04)
Netherlands	0.10 (0.11)	0.08 (0.11)	0.13 (0.11)
Sweden	0.07 (0.05)	0.11 (0.05)	0.11 (0.05)
United Kingdom	0.00 (0.08)	0.04 (0.10)	0.05 (0.09)
Discrete estimation	-0.99 (0.54)	-0.96 (0.55)	-0.31 (0.15)
Non-parametric estimation	-0.86 (0.53)	-0.76 (0.54)	-0.15 (0.12)
OLS estimation	-0.12 (0.13)	-0.09 (0.12)	0.01 (0.10)
TOLS estimation	-0.14 (0.12)	-0.09 (0.12)	0.01 (0.07)
Linear specification	-0.01 (0.11)	0.02 (0.11)	0.12 (0.07)
Log linear specification	-0.06 (0.13)	0.01 (0.12)	0.10 (0.07)
Quadratic specification	0.12 (0.16)	0.19 (0.15)	0.28 (0.13)
Simulation	0.88 (0.54)	0.80 (0.55)	0.29 (0.10)
Unmarried	-0.01 (0.12)	-0.02 (0.11)	-0.02 (0.12)
Mixed study	0.17 (0.13)	0.18 (0.11)	0.19 (0.10)
Female	0.39 (0.05)	-0.44 (0.36)	-0.30 (0.37)
Both sexes	0.23 (0.14)	-0.92 (0.45)	-0.49 (0.42)
Participation-rate * Female		-0.43 (0.25)	-0.12 (0.20)
Participation-rate * Male		-1.18 (0.43)	-0.74 (0.39)
Constant	0.06 (0.13)	1.17 (0.45)	0.64 (0.37)
R-squared	0.33	0.35	0.32
Observations	239	239	235

*Standard errors between parentheses (White heteroskedasticity-consistent)*

*Benchmark for dummy variables respectively: US, Maximum Likelihood, Double log-specification, Married and Male Panels (2) and (3) only differ in that the latter has omitted 4 outlier observations.*

The estimated coefficients in Table 3.2 allow us to predict elasticities for some specific cases. In particular, we use the meta regression and then insert one or more dummy variables to compute particular elasticities, thereby holding other variables at their sample means. A selection of the elasticities thus obtained are presented in Table 3.3. Clearly, the predicted elasticities are larger for women than for men in all countries. For instance, the elasticity for Dutch males lies between 0.07 and 0.16, while for females it lies between 0.48 and 0.52. The inclusion of sample participation rates – i.e. going from (1) to (2) – affects the predicted elasticities for men. Controlling for this variable leads to higher predicted elasticities for men. Perhaps omitted variable bias plays a role, rendering incorrect estimates if the sample participation rate is left out as a moderator variable in the meta-regression. More in particular,

this finding suggests that much of the difference between men and women found in panel (1) can be attributed to the higher participation rate of men. Our results however indicate that, even after controlling for this ‘participation effect’ women still have a higher labour supply elasticity than men. Finally, the omission of outliers – panel (3) – leaves the predicted elasticities roughly unchanged.

**Table 3.4** Predicted elasticities

	(1)	(2)	(3)
Women (Netherlands)	0.52	0.50	0.48
Men (Netherlands)	0.07	0.16	0.16
Women (United Kingdom)	0.40	0.47	0.36
Men (United Kingdom)	0.02	0.24	0.13
Women (United States)	0.38	0.40	0.31
Men (United States)	0.01	0.24	0.16
Women (Sweden)	0.54	0.59	0.53
Men (Sweden)	0.12	0.42	0.31

*Column numbers correspond to specifications/sample selections in Table 3.2.*

By setting participation rates in the meta regression (3) equal to one, we can also simulate elasticities at the intensive margin. Taking the Netherlands as an example (i.e. impose a one for the Netherlands dummy), we obtain a point estimate for the labour supply elasticity at the intensive margin of 0.44 for Dutch women and 0.03 for Dutch men. Hence, we confirm that the difference in elasticities between men and women, which is generally found in the literature, cannot be fully explained by the difference in participation rates. Indeed, differences also exist at the intensive margin. This contrasts with the suggestion by Mroz (1987) that married women and prime aged males are equally sensitive at the intensive margin.

### 3.3 Robustness for the specification

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. This section performs several robustness tests for the potential omission of moderator variables. Moreover, we include study fixed effects. Thus, we explore whether the results are robust for the inclusion of other moderator variables. Our starting point is model specification (3) in Table 3.2. The results of the extended regressions are presented in Table 3.4. Note that coefficients for moderator variables reported in Table 3.2 are not reported in Table 3.4.<sup>7</sup>

<sup>7</sup> The complete results are available upon request.

The regression results in column (4) are from a specification that includes five dummy variables representing study characteristics. These refer to (i) whether fixed costs of participation on the labour market are included in the primary study, (ii) the effect of using panel data instead of cross section data, (iii) whether the study has been published in a refereed journal, (iv) whether desired hours of work are used instead of actual hours of work, and (v) whether measurement error is explicitly taken into account for the observed hours of work. The second regression in column (5) introduces dummy variables which indicate whether certain control variables are used in primary studies. These include age (-squared), education (-squared), health, and the presence of children. The final regression in column (6) of Table 3.4 combines the two former specifications, and adds study fixed effects for ten studies that report at least 8 elasticities. Hence, this last regression removes much of the between-study variation in our sample, and coefficients are for a large part identified from the ‘within variation’ of studies.

Table 3.4 shows that some of the study characteristics matter for the elasticities. For instance, studies that include fixed costs in the regression report systematically higher elasticities. Published studies tend to produce smaller elasticities than unpublished studies, suggesting that it is easier to publish ‘moderate’ elasticity estimates than outliers. We will come back to this in section 3.5. Studies that correct for measurement error report higher elasticities. Statistical theory predicts that if an explanatory variable suffers from measurement error, then the estimated coefficient will be biased towards zero (attenuation bias), so that our finding is consistent with theory. In panel (5) four out of the nine control variables matter for the estimated elasticities. This holds in particular for education and age variables. However, when study fixed effects are introduced in the final column of Table 3.4, hardly any study characteristic or control variable is found to differ from zero significantly. It should however be noted that the within variation of studies is often small (i.e. studies employ similar specifications in different estimation runs), leading to high standard errors and limited scope for obtaining strong results from statistical tests.

With the estimates from Table 3.4, we have again predicted elasticities for specific cases as we did in Table 3.3. These estimates for the uncompensated elasticity of labour supply change somewhat due to the inclusion of other moderator variables, but never significantly so. Therefore, we do not report these values here. Overall, it is fair to state that the main results from section 3.2 remain qualitatively the same and that the results from Table 3.3 carry over to this section.

**Table 3.5 Estimation results for the extended specifications**

	(4)	(5)	(6)
<i>Study characteristics<sup>a</sup></i>			
Fixed costs	0.21 (0.10)		-0.07 (0.16)

Panel data	0.13 (0.05)		-0.32 (0.24)
In refereed journal	-0.25 (0.04)		-0.35 (0.41)
Actual-desired-hours	-0.10 (0.09)		-0.07 (0.07)
Measurement error	0.13 (0.06)		0.00 (0.06)
<i>Control variables<sup>a</sup></i>			
Child		-0.10 (0.13)	0.24 (0.16)
Child younger than age 6		0.08 (0.11)	0.18 (0.15)
Education		-0.23 (0.08)	0.06 (0.07)
Family		-0.21 (0.18)	-0.12 (0.33)
Family size		0.01 (0.18)	-0.10 (0.49)
Health		-0.05 (0.19)	-0.02 (0.48)
Age		0.37 (0.08)	0.22 (0.20)
Age Squared		-0.16 (0.09)	-0.12 (0.04)
Age dummy variables		0.28 (0.07)	0.09 (0.31)
Study effects <sup>b</sup>	No	No	Yes
R-squared	0.47	0.46	0.66
Observations	235	235	235

Standard errors between parentheses (White heteroskedasticity-consistent)

<sup>a</sup> 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.

<sup>b</sup> 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.

While Table 3.4 shows the significance of each moderator variable separately, Table 3.5 shows results for redundant variable tests for sets of moderator variables in the last specification of Table 3.4. We see from Table 3.5 that the combined gender and participation variables are by far the most important moderator variables. Second, both country effects and the estimation method should not be ignored altogether, although we have seen that in particular the magnitude of (the point estimates of) country effects was small. Still, it is remarkable that the estimation method appears to have some impact on elasticity point estimates. Furthermore, the results suggest that we should not care so much about the exact specification, marital status and the five study characteristics (see Table 3.4 for the latter). Also, control variables included in the primary studies seem to matter for the outcomes of studies on labour supply elasticities.

**Table 3.6** Redundant variable test for specification (3) in Table 3.4

	F-statistic	Log Likelihood ratio	Prob. (F-Stat)	Prob. (Log. Lik.)
Country	3.26	26.80	0.00	0.00
Estimation	4.80	22.72	0.00	0.00
Specification	1.63	7.97	0.17	0.09
Marital status	1.59	3.93	0.21	0.14
Gender / Participation	25.24	79.19	0.00	0.00
Characteristics	0.40	2.50	0.85	0.78
Control variables	2.22	23.58	0.02	0.01
Study effects (8)	2.62	9.56	0.05	0.02

### 3.4 Robustness for the sample

Table 3.6 shows the estimated coefficients for alternative samples. The specification used is the same as specification (3) in Table 3.2. The first columns of Table 3.6 show results for samples of women and men, respectively. The third estimation is based on a sample that contains only observations from studies published in refereed journals, while the last estimation is based on a sample 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 70 observations. We find that among the elasticities published in refereed journals, the model specification appears to have more impact than for other elasticity estimates. In particular, the earlier mentioned effect of the quadratic specification is now more pronounced. There is no clear explanation for this, and perhaps unobserved study effects may simply play a role. If we only allow ‘independent observations’, these specification effects vanish. Note that the latter sample generates relatively high standard errors due to the reduction in sample size.

**Table 3.7** Estimation results for alternative samples

	Female	Male	Published in refereed journal	Independent observations
Finland	-0.09 (0.12)			-0.28 (0.12)

France	-0.27 (0.19)	0.10 (0.03)	-0.12 (0.14)	-0.13 (0.13)
Germany	-0.54 (0.17)	-0.22 (0.14)		-0.29 (0.15)
Italy		0.10 (0.03)	0.03 (0.04)	-0.03 (0.00)
Netherlands	-0.17 (0.13)	0.12 (0.04)	0.09 (0.11)	-0.03 (0.14)
Sweden	0.03 (0.16)	0.15 (0.04)	-0.05 (0.07)	0.01 (0.08)
United Kingdom	0.12 (0.10)		0.16 (0.09)	0.07 (0.21)
Discrete estimation	0.09 (0.20)		-0.19 (0.18)	
Non-parametric estimation	0.79 (0.19)	0.15 (0.06)	0.13 (0.14)	0.42 (0.32)
OLS estimation	0.12 (0.14)	-0.21 (0.04)	-0.05 (0.11)	-0.22 (0.11)
TSLS estimation	0.05 (0.11)	-0.11 (0.05)	0.07 (0.06)	-0.13 (0.10)
Linear specification	0.39 (0.12)	-0.11 (0.04)	0.31 (0.08)	0.05 (0.10)
Log linear specification		-0.10 (0.04)	0.26 (0.08)	-0.06 (0.10)
Quadratic specification	0.66 (0.19)	-0.12 (0.04)	0.51 (0.13)	0.19 (0.14)
Simulation	0.34 (0.13)	-0.32 (0.07)	0.32 (0.11)	-0.07 (0.16)
Unmarried	-0.11 (0.16)	0.00 (0.04)	0.01 (0.12)	0.01 (0.08)
Mixed study	0.18 (0.17)	0.26 (0.05)	0.41 (0.12)	0.16 (0.17)
Female			-0.31 (0.56)	-0.26 (0.38)
Mixed study, both sexes			-0.51 (0.62)	-0.64 (0.42)
Participation-rate * Female	-0.08 (0.20)		-0.22 (0.19)	-0.34 (0.23)
Participation-rate * Male		-1.18 (0.48)	-0.85 (0.61)	-0.86 (0.41)
Constant	0.21 (0.14)	1.28 (0.47)	0.59 (0.58)	0.92 (0.40)
R-squared	0.38	0.44	0.41	0
Observations	108	119	185	70

*Standard errors between parentheses (White heteroskedasticity-consistent)*

### 3.5 Publication bias

Publication bias occurs if not all estimated effect sizes are published, because of endogenous selection. This may be due to journal editors and referees selecting significant results or authors leaving insignificant results in the file drawer. Since unpublished results are either not registered or difficult to find, there exist no formal tests to directly detect publication bias. However, one can explore indirect evidence to see whether publication patterns are consistent with the presence of publication bias. A general approach is to correlate the observed effect sizes with design features of studies that are risk factors for publication (Begg, 2002). The mostly used factors are sample size or standard errors. We follow Card and Krueger (1995), who propose a regression of the estimated effect size on the standard error and a constant. In theory, the slope parameter for the standard error should equal zero, because there exists no systematic relationship between a coefficient's point estimate and its standard error. However, if publication bias is present, then the coefficient might be unequal to zero. In particular, if studies finding significant elasticities are more likely to be published, then we can expect the slope coefficient to be positive. We estimated such a regression for a sample of 195 observations for which standard errors could be computed, and obtained a coefficient of 1.96

for the standard error (with a standard error of the coefficient of 0.24). Hence, on the basis of a simple t-test, we cannot reject the presence of publication bias in the literature on labour supply elasticities.

## 4 Conclusion

This paper aims to identify the sources of variation in empirical estimates of the uncompensated labour supply elasticity. Earlier studies principally focussed on a limited number of sources of variation, such as model specification and the estimation technique. Moreover, these studies explored this on a partial basis. We add to this by exploring a broader set of potential sources of variation and by means of a *simultaneous* meta analysis. To that end, we develop a sample of 239 elasticities drawn from 32 empirical studies in the literature. Thereby, we explore the systematic impact of a great number of study characteristics. One interesting finding is that the model assumption on the relation between hours worked and the wage rate – be it linear, quadratic, log-linear or double-log – mostly does not have a significant impact on the elasticity estimates. Only a quadratic specification appears to produce higher elasticity estimates than other specifications. A second finding is that the difference between elasticities among countries is small. Looking at four particular countries, the US, UK, Netherlands, and Sweden, there is no evidence for different elasticity values. Finally, we find that females have a larger labour supply elasticity than males, even after controlling for participation rates. This suggests that the elasticity of hours worked with respect to the net wage rate (the intensive margin) is more elastic for females than for males. Thus, in the near future female elasticities will indeed become lower as participation rates increase, but on the basis of our findings it is questionable whether the relatively low male level will ever be achieved.

A test statistic proposed by Card and Krueger (1995) is used to explore the presence of ‘publication bias’ in the empirical literature. We find that the presence of publication bias cannot be rejected. That is, the elasticity estimates presented in the literature do not appear to be a randomly selected sample of estimated labour supply elasticities.

Another aim of this paper is to achieve a ‘synthesis’ of research results for certain special cases. For instance, what would be a reasonable estimate for the uncompensated elasticity for women in the Netherlands? Using our meta regression, we predict the uncompensated elasticity of labour supply for Dutch women at around 0.5. The corresponding figure for men is predicted at 0.1 or 0.2, depending on the preferred specification. Predictions for Sweden, the UK, and the US are qualitatively the same as for the Netherlands.

## 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, 537-559.

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

Ballard, C.L., J.B. Shoven and J. Whalley, 1985, General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States, *American Economic Review*, vol. 75, no. 1, 128-137.

Bargain, O., 2005, *On Modelling Household Labor Supply with Taxation*, IZA Working paper 1455, Bonn, IZA.

Begg, C., 1994, Publication Bias. In: H. Cooper and L.V. Hedges, eds., *Handbook of Research Synthesis*, Chapter 25, Russell Sage Foundation, New York.

Blau, F., and L. Kahn, 2005, *Changes in the Labor Supply Behavior of Married Women: 1980-2000*, Working paper, 11230, NBER.

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

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

Blomquist, S., 1996, Estimation methods for male labor supply functions: How to take account of nonlinear taxes, *Journal of Econometrics*, 70, 383-405.

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

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

Blundell, R., and T. MaCurdy, 1999, Labor supply: A review of alternative approaches. In: O. Ashenfelter D. Card, eds., *Handbook of Labor Economics*, vol. 3A, Ch. 27, Amsterdam, North Holland.

- Blundell, R., A. Duncan, J. McCrae, and C. Meghir, 2000, The labour market impact of the working families tax credit, *Fiscal Studies*, 21, 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, 358-389.
- Browning, E.K., 1987, On the Marginal Welfare Cost of Taxation, *American Economic Review*, vol. 77, no.1, 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, 488-479.
- Card, D., and A. Krueger, 1995, Time-Series Minimum-Wage Studies: A Meta-analysis, *American Economic Review: Papers and Proceedings*, 85, 238-243.
- Cogan, J., 1981, Fixed Costs and Labor Supply, *Econometrica*, 49, 945-963.
- Colombino, U., and D. del Boca, 1990, The Effect of Taxes on Labor Supply in Italy, *Journal of Human Resources*, 25, 390-414.
- Devereux, P., 2003, Changes in Male Labor Supply and Wages, *Industrial and Labor Relations Review*, 56, 409-428.
- Devereux, P., 2004, Changes in Relative Wages and Family Labor Supply, *Journal of Human Resources*, 39, 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, 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, 431-460.

Euwals R., and A. van Soest, 1999, Desired and actual labour supply of unmarried men and women in the Netherlands, *Labour Economics*, 6, 95-118.

Euwals, R., 2001, Female Labour Supply, Flexibility of Working Hours and Job Mobility, *Economic Journal*, 111, 2.120-2.134.

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, 239-278.

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

Greene, W., 1993, *Econometric Analysis*, 2<sup>nd</sup> edition, New York: Macmillan.

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, 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, 242-248.

Heckman, J., 1979, Sample selection bias as a specification error, *Econometrica*, 46, 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, 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.

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

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.

MaCurdy, T., P. Green, and H. Paarsch, 1990, Assessing Empirical Approaches for Analyzing Taxes and Labor Supply, *Journal of Human Resources*, 25, 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, 63-105.

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

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

Mroz, T., 1987, The Sensitivity of an Empirical Model of Married Woman's Hours of Work to Economic and Statistical Assumptions, *Econometrica*, 55, 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, 251-274.

Prescot, E.C., 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, 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, 517-558.

van Soest, A., 1995, Structural Models of Family Labor Supply: A Discrete Choice Approach, *Journal of Human Resources*, 30, 63-88.

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

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

Woittiez, I., and A. Kapteyn, 1998, Social interactions and habit formation in a model of female labour supply, *Journal of Public Economics*, 70, 185-205.