ERASMUS

9203/A

ECONOMETRIC INSTITUTE

A SYMMETRIC APPROACH TO THE LABOR MARKET

B.M.S VAN PRAAG, J.P. HOP AND E. EGGINK

REPORT 9203/A

Erasmus

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM-THE NETHERLANDS

A SYMMETRIC APPROACH TO THE LABOR MARKET

by

Bernard M.S. van Praag
J. Peter Hop
Evelien Eggink

In this paper we present a modification of the usual labor market model. Three wage concepts are introduced. We assume that the worker has a minimum wage in mind for which she is willing to participate and that the employer has a maximum wage in mind, which he is willing to pay for a specific type of labor. If the female does not work we assume that the institutional wage does not lie between those wage levels. This model is estimated on a cross-section data set of 6352 working and non-working married females. We employ a flexible simulated EM-algorithm where we simulate from individually varying sets with positive or zero measure in $\mathbb{R}^3$ and $\mathbb{R}^4$. The estimated model indicates that the presence of young children does not only reduce the female's willingness to participate but also the employer's willingness to hire. On the basis of the model we differentiate unemployment as being due to unwillingness of the worker, the employer or both to supply and demand of labor at the present institutional wage level. It is estimated that about 40% of female employment may be attributed to the "insider advantage". A political evaluation of some effects of wage measures is provided. It is found that, in addition to the manipulation of wages and wage costs, the supply of childcare facilities will have a considerable impact on female participation.

Keywords: Female labor supply, simulated EM-algorithm, set-valued estimation, employer's maximum wage, childcare.

A SYMMETRIC APPROACH TO THE LABOR MARKET

by

Bernard M.S. van Praag *

J. Peter Hop **

Evelien Eggink ***

November, 1991

---

* Econometric Institute, Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.
Scientific Council for Government Policy (WRR), The Hague

** Econometric Institute, Erasmus University Rotterdam

*** Econometric Institute, Tinbergen Institute, Erasmus University Rotterdam

1. Introduction.

Nowadays the employment problem is one of the major issues in economic policy. Economic research has reacted on that problem by a vast number of contributions both in the field of macro- and micro-economics. Traditionally within micro-economics a distinction is made between the male and the female labor market. This paper is a contribution to the micro-economic analysis of female employment. It is clearly impossible to review the vast literature in any depth and we refer to the excellent surveys by Pencavel (1986) and Killingsworth and Heckman (1986) as points of departure. The empirical problems gave rise to a specialization in econometrics, which is surveyed by Heckman and MaCurdy (1986).

In the Netherlands like in many other countries there is a severe unemployment problem, both for the male and the female population. Like in a commodity market in the labor market demand may equal supply, provided that the price of labor is flexible. Then wages are equilibrium prices. The phenomenon of a large persistent unemployment together with a labor shortage on some specific segments of the labor market suggests that wages do not perform their role as equilibrium prices. This is not very surprising as wages in the Dutch labor market are more or less centrally determined by collective wage negotiations between employers and trade unions. The collective contracts are rather strictly maintained by both parties; in most sectors they are also binding by law for workers and employers who are not a member of the organizations participating in the negotiations. The result of those negotiations will be called the institutional wage rate, say w. The unemployment problem seems to be particularly pertinent with respect to married women.

Since the seminal work by Heckman (1974) it is recognized that labor supply behavior cannot be adequately studied if the sample consists of workers only, due to selectivity bias.

We thank Joop Hartog, Bernard Salanié and Jules Theeuwes for their helpful comments on earlier drafts. Two anonymous referees made us aware of major shortcomings, we are grateful for their inspiring comments which lead to a thorough revision. Financial support for this study has been given by the Netherlands Organization for the Advancement of Scientific Research (NWO). The data base has been created by Aldi Hagenaars and Bernard Van Praag in cooperation with the Joint Press Service, The Hague. The authors are solely responsible for remaining errors and the opinions expressed.

The selectivity bias can be corrected by assuming that individuals have a reservation wage w_{min}, such that they refuse to participate if the institutional wage rate w is below their w_{min}. This is assumed to be the case for the non-working individuals in the sample. It is not generally recognized that there is a second reason for being unemployed, viz., the behavior of the employer. He may not want to hire particular individuals for the institutional wage rate.[1] In that sense the approach is one-sided as it ignores employer behavior as a cause of being unemployed.

In this paper we will focus on the two-sidedness of the employment relation, where both the worker and the firm are considered.[2] In the labor market individuals are looking for jobs offered by firms and firms are looking for workers. As an approximation we assume the following model. The firm looks for labor of a specific quality and the employer has a maximum price in mind which he is willing to pay for that type of labor. That maximum price is an amount of w_{max} per hour, depending on the quality of the labor. It is implicitly determined by the expected marginal productivity of the additional worker. Similarly the worker has a minimum price in mind for which she is willing to offer her labor. This price, w_{min} per hour, is based on utility maximization considerations. Finally, there is an institutional wage w for that type, which is established in negotiations between the union of employers and the trade union. It is well conceivable that employer and worker find each other in the informal market at a wage level different from the institutional w, but it is legally forbidden to conclude such a deal. So in this paper it is assumed for a participant that she wants to work for the institutional wage ($w_{min} \leq w$) *and* that the employer is willing to pay her that wage rate ($w \leq w_{max}$). Thus, by comparing the three wage rates we may find the following classification into six cases:[3]

---

1 Recently Blundell, Ham and Meghir (1987) attempted to incorporate this feature.

2 Independently Van Soest and Kapteyn (1989) consider a similar but not identical model; their assumptions are more restrictive than the assumptions we make. See footnote 7.

3 In most segments of the Dutch labor market the informal economy is negligible and we shall ignore the possibility in this paper. In fact we only refer to the possibility of an informal market in the applications of the model considered in section 5. Furthermore we ignore the equality signs as they lead to events of probability zero.

A. $w_{min} < w < w_{max}$

The worker is employed in the formal market.

B. $w_{min} < w_{max} < w$

The worker would be employable in the informal market, but the institutional wage rate is too high for the employer.

C. $w < w_{min} < w_{max}$

The worker would be employable in the informal market, but the institutional wage rate is too low for the worker.

D. $w_{max} < w < w_{min}$

The institutional wage is too high for the employer and too low for the worker.

E. $w < w_{max} < w_{min}$

The employer wants to hire but the worker is not interested.

F. $w_{max} < w_{min} < w$

The worker is willing but the employer is not interested.

Most, if not all, labor market studies implicitly assume that workers are hired day by day on the basis of market decisions. This implies that even after twenty years in a job a worker is assumed to satisfy the market conditions, as if he would be hired anew that day. Obviously this is not realistic in a modern welfare state. For the employee it is fairly easy to quit if $w < w_{min}$, but for the Dutch employer it is very difficult to fire a worker, even if $w_{max} < w$, due to legal protection of the worker in her job. Moreover, in most jobs workers who have known the firm for some time have a higher value to their employer than on the market in general, where it is supposed they start afresh as an anonymous applicant. Properly speaking, we have to differentiate between w_{max} (present job) and w_{max} (general), denoted by w^{p}_{max} and w^{g}_{max} respectively. In most cases for an employed worker there will hold that $w^{p}_{max} > w^{g}_{max}$, as the worker has knowledge and experience specific to her present job. Moreover, in most cases, the transaction cost of firing the person and hiring and training of the new worker are considerable. In order to stay in her present job the relevant concept is w^{p}_{max}. Thus many people who would have no chance at the labor market at large, given the relevant variable w^{g}_{max} may still be observed to be working, as the employer decides on continuation on the basis of w^{p}_{max}. Actually these remarks are the core of the "insider-outsider" theory of Lindbeck and Snower (1988). It is the market in which we are interested, i.e., the relevant concept in this paper is w^{g}_{max}.

The objective of this paper is to estimate the three wage rates w, w_{min} and w_{max} for females of various qualifications. Relevant aspects are the age, the education level and the number of children to be cared for where we differentiate between children below 4, between 4 and 12, and older children. Also the number of hours the woman works is included. Finally, at least on the supply side, other income, i.e., the income of the husband, is expected to be relevant.

From the results it will be possible to calculate for the existing volume of female unemployment what fraction is due to lacking demand by the employers and what fraction may be ascribed to lacking supply of the workers. Estimates of the effect of wage increases and the supply of childcare free of charge can then be assessed.

We consider a sample of 6352 female non-breadwinners drawn from the Dutch population in 1983. In the data set we have a variable for the reservation wage w_{min} which has only been filled in by those females who either participate or are looking for a job. A proxy for w_{max} can be constructed for the same groups. This proxy is based on the self-perceived employability of the respondents, if they would apply for a new job. This empirical operationalization of the maximum wage rate and hence the resulting employment concept abstains from the insider-outsider effect. Consequently the employment chance for workers is consistently underestimated. Comparison with observed employment frequencies will reveal that about one third of observed employment has to be ascribed to the insider-outsider effect.

As for most of the non-workers most variables are unobserved we are confronted with an incomplete data set with a sample selection problem. For many observations likelihoods become rather weird integrals which are not readily evaluated in a routine and computationally feasible way. A classical iteration method to solve this type of problem is application of the EM-algorithm (Dempster, Laird and Rubin (1977)). In this method the data are "completed" by replacing incomplete data by their conditional expectations. However, in many cases these conditional expectations are just as difficult to evaluate as the original likelihoods. Therefore, instead of using one of the well-known methods for estimating this kind of models (see e.g. Maddala (1983)), we replace these conditional expectations by averages of a *finite* number of simulations from the conditional distributions. We call this method, which is related to the method of simulated scores, the Simulated EM-algorithm (SEM).

This paper builds on a theoretical paper by Van Praag and Hop (1987) and describes a method which has been independently developed by Ruud (1991). Our technical approach is different from and more efficient than the methods suggested by Ruud. Research in the same spirit on time series is performed by Hajivassiliou and McFadden (1987). McFadden (1989) developed the method for discrete choice situations. See also Pakes and Pollard (1989) and Gouriéroux and Monfort (1989) for theoretical considerations. Recently Laroque and Salanié (1989) estimated a macro-economic fix-price model by a similar method based on a time series of macro-data. A recent application on panel-data is by Keane (1989).[4] See Gouriéroux et al. (1987) for related work in the context of hypothesis testing.

In Section 2 we outline the theoretical model, while Section 3 deals with the SEM-algorithm. In Section 4 we describe the simulation technique in detail. The data and results are presented in Section 5. In Section 6 we discuss the political-economic implications of our results for the Dutch female labor market. Section 7 concludes.

[4] The American, French and Dutch authors worked independently of each other.

2. The economic model.

The traditional model of labor supply (Becker (1965)) assumes that the choice between income y and leisure le (le=T−h; T=24; h=hours worked per day) is made in such a way that a quasi-concave utility function $U(c,T-h)$ is maximized subject to a budget constraint $c=y_{NL}+w*h$, where y_{NL} stands for non-labor income and w for hourly wages net of taxes.

This model is too simple to describe reality. At first the budget curve is not necessarily a straight line, but it may be rather kinked (e.g. Hausman (1985)). This is due to the fact that frequently non-labor income (e.g. unemployment benefits) is reduced if one is starting to participate, and to the fact that participation causes additional (mainly fixed) costs, for instance travel cost, homecare, clothing and specifically childcare (cf. Cogan (1981)). A second point of tension between reality and theory is the assumption that hours can be fixed at will by the worker.

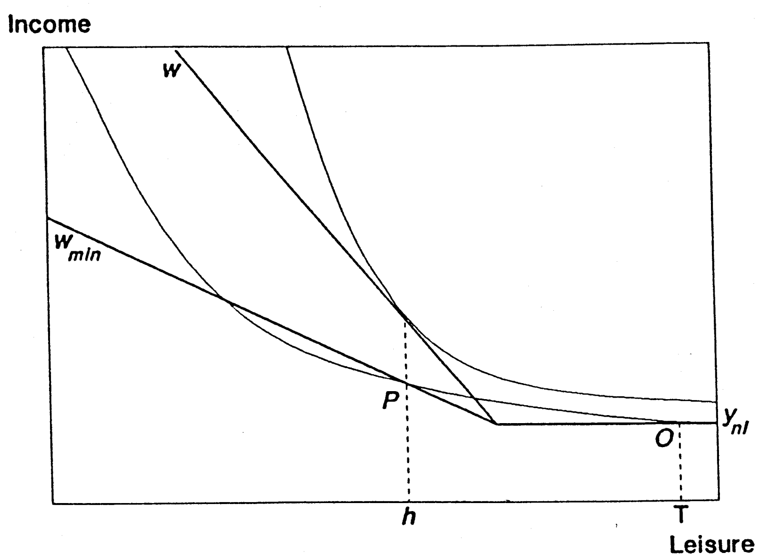

Figure 2.1. The choice for labor and leisure for a worker.

Looking at figure 2.1, depicting our model in the case of a worker, it follows that situation O standing for non-participation is just as attractive as situation P where one works h hours for w_{min} per hour. For less than w_{min} one does not supply h hours of labor. Notice that w_{min} is the slope of the budget curve at P, where the worker is just indifferent between working for h hours or supplying no hours at all; w_{min} is not the traditional *reservation wage*, i.e., the slope of the indifference curve at O.

Just like the producer is rather flexible in his investment decisions *ex ante* but investment is fixed *ex post*, the average female employee has some room in choosing the hours supply before she accepts a job, but it is in most cases hardly possible to change the hours supply after acceptance of a job. Then there is only the choice of keeping the job or looking for another job with different working conditions. It follows that, even if we assume that the worker would have chosen an optimal position when accepting the job, we have to recognize that workers who have worked for some time mostly are not at the optimum which the static neoclassical model prescribes but rather at a suboptimal position which is considered "satisficing" enough to keep the job (see Simon (1979)). In short we assume the supplied hours to be predetermined for those who work.

Over the course of life it is quite probable that preferences for consumption and leisure will change. If one grows older and/or the household becomes larger, the need for leisure (including time for home production) may intensify and one would prefer a continuous change in labor supply. As this is impossible, one may have to choose for leaving the job altogether. Thus, given h, there is a critical w_{min}-level; if the institutional wage w falls below that level one stops working. Notice that the institutional wage may stay constant over time and w_{min} may increase as a result of shifting preferences. It depends among others on age and the presence of children, but also on the hours h one is actually working. Another reason for quitting can be a shift of the budget curve, when the husband gets more income (y_{NL}).

Non-participating women have to be induced into working by a wage exceeding their w_{min}, which is determined as in Figure 2.1. Again they assign different values of utility to different supplies h.

A similar case holds for the employer. There is an upper wage level w_{max} which the employer is willing to pay. If, by the institutional wage, the worker costs more than w_{max}, she will not be hired. Notice that this can again be based on a profit maximizing argument. The maximum wage will also vary in time, e.g., with the age of the female, the capital intensity of the firm, the technology, etc. Finally it will depend on the supply of hours, which is not wholly flexible.

We describe the variables and the equations informally, where we ignore for the moment the fact that not all the variables are observed for every individual. At the end of this section the model equations are given in full.

The supply of hours h.

For the reasons given in the discussion above, and others (see e.g. Mroz (1987)) we assume a labor supply function in reduced form, which does not depend directly on the wage w but only on more basic variables. It describes the number of hours h supplied by those who participate, while for those who are non-participating h has to be interpreted as the number of hours they would supply if they would get an acceptable wage offer. The number of hours supplied is only observed for workers. Table 2.1 gives the average number of hours worked per age bracket.

Table 2.1. Hours supplied by working females.
(number of respondents in parentheses)

	Average number of hours h	
Under 25	34.6	(396)
25 to 29	30.0	(699)
30 to 34	21.0	(442)
35 to 44	19.5	(791)
45 to 54	18.5	(278)
55 and older	17.6	(35)
Whole sample	25.8	(2641)

Instead of the absolute number of hours h, worked per week, we take the logarithm, so the used variable can take on negative values. This is necessary because we make the usual assumption that the error is normally distributed with expectation zero. As we do not know the working experience in our sample we use ($age - 16$) as a proxy, as schooling is compulsory until 16. This is also the appropriate variable for women who do not have any working experience but whose aspiration level is that of their age-mates.
We assume that supply is at first increasing with age and later on decreasing. Furthermore we introduce the variable fs, which stands for the family size without any equivalence scaling, and two dummies for the age of the children, D_{ch4} and D_{ch12} being equal to one when the youngest child is below the age of four or between four and twelve respectively. The education level is modeled by four dummies $E_2, \dots, E_5$, where the reference level is the lowest level in the Dutch education system. In this research only individuals who are not the main breadwinner of the household are considered;

they all have a spouse who is employed.[5] Thus we can include the wage rate and the hours worked by the main breadwinner. Except for the dummies all variables are measured in logarithms.

The minimum wage w_{min}.

The worker's preferences are reflected in the (net) minimum wage w_{min} she requires to be employed, which we measure in logarithms. We model the relationship as being parabolic in $\ln(h)$. The assumption is, that it is rather difficult for a female to work for only a few hours a week, but that when she has started to organize her life as a working woman, there will be economies of scale if her working day is prolonged, resulting in a lower hourly wage. However, after a certain optimal hours supply she has to be seduced into work by a higher wage to work full time. This change in attitude may be expected to be in the range of twenty hours a week. We expect that females with higher education will set their wage demands higher. Therefore we include four dummies for the education levels. The minimum wage also depends on age (or experience): the relationship is again assumed to be parabolic. In the beginning of life individuals are learning. This implies that young women will ask a relatively low price and that that price will increase when the woman grows older. Similar to the labor supply h we expect that the household situation will have an impact, where the female with a larger or younger family at home will expect a higher wage to be seduced into work. Finally there may be an income motive. Conform to standard micro-economic theory we assume that the wage demand is set higher when there is other household income; w_{min} will increase as the husband earns more.

The variable w_{min} is observed for workers, who had to answer the questions:

> "Suppose your job were at risk. If you could keep your job by accepting a wage reduction, would you do so?" (yes/no)
>
> "If yes, what would be the maximum reduction rate (in percentage of your present wage)?"

The frequencies to these questions are presented in table 2.2.

---

[5] We excluded from the sample 18 observations where both partners were male and 16 observations where both partners were female. Finally 204 non-breadwinners were excluded, because they were males instead of females.

Table 2.2. Frequencies of the maximum reduction rate.

	Frequencies of reduction rate	
N.A. (Unemployed)	3731	
No (0 %)	554	(23 %)
Yes 1 to 5 %	508	(21 %)
6 to 10 %	827	(36 %)
11 % and more	481	(20 %)
Unknown (missing)	251	
	6352	(100 %)

If the response to the first part of the question is "yes", and the maximum reduction rate is α, the variable w_{min} is given by $(1-\alpha)w$. If the response to the first question is "no", we assume that the worker's wage is equal to the critical wage level w_{min}, thus $w_{min} = w$.
The active job-seekers were explicitly asked to fill in their reservation wage and the number of hours per week they would like to work.
Table 2.3 gives some insight of the average w_{min} for working females and job-seekers per age bracket.

Table 2.3. Average minimum wage per hour for working and job-seeking females.
(in dutch guilders; number of respondents in parentheses)

	Average minimum wage w_{min}			
	workers		job-seekers	
Under 25	9.05	(369)	8.32	(24)
25 to 29	9.66	(649)	10.10	(41)
30 to 34	10.51	(400)	9.08	(49)
35 to 44	10.28	(679)	9.00	(57)
45 to 54	10.58	(240)	9.47	(16)
55 and older	10.82	(31)	18.86	(2)
Whole sample	9.88	(2368)	9.19	(189)

The institutional wage w.

The third equation to be specified is the *wage*-equation, the institutional wage measured in logarithms. This is specified quite traditionally as a (quadratic) function of $\ln(age - 16)$ to cover the age-dependency in institutional wage scales. The educational dummies (E_2, ... ,E_5) are used to set different levels for different levels of education. The institutional wage is reported by the respondent in terms of net amounts. As the Dutch tax and family allowance system depends on the presence and the age of the children, the two dummies for the age of the children (D_{ch4} and D_{ch12}) are included in the $\ln(w)$-equation to capture this effect. In table 2.4 the average wage is given by age and by education.

Table 2.4. Average wage per hour for working females.

(in dutch guilders; number of respondents in parentheses)

age group:	Average wage w		education level:	Average wage w	
Under 25	10.07	(396)	level 1 (lowest)	9.73	(600)
25 to 29	10.70	(699)	level 2	10.66	(747)
30 to 34	11.23	(442)	level 3	10.59	(627)
35 to 44	11.31	(791)	level 4	12.51	(162)
45 to 54	11.87	(278)	level 5 (highest)	13.46	(505)
55 and older	11.75	(35)			
Whole sample	10.92	(2641)		10.92	(2641)

The maximum wage w_{max}.

The maximum wage w_{max}, also measured in logarithms, reflects the employer's preferences. It would be possible to derive a formula from a formal marginal productivity model. However, in this paper we immediately postulate a specification, based on informal intuition. We assume that the maximum wage will depend on the education level of the woman to be employed. The five levels are again represented by the four dummies indicating the education level relative to the lowest level. Further, we assume that age as a proxy for experience and physical fitness will play a role. We assume that young girls have no experience, so their productivity will be low. During their life they learn at the work place, and this will result in a maximum wage which rises with experience. After a certain moment experience can not increase any further, and as the female grows older, she will become less productive, at least in the perception of most employers. So we assume a parabola in years of experience, which is again measured by $\ln(age - 16)$. Although it has formally no place in the theory nor in conventional political wisdom, direct observation leads us to believe that the household situation of the female is important from the viewpoint of the employer. More precisely we believe that a female with children at home is no attractive bargain for most employers. The problem, at least in the Netherlands, is that childcare is expensive and difficult to obtain and that in 1983, the year the data were collected, there was practically no institutional childcare. This implies for the Dutch situation in 1983 that any illness of one of the children, the babysit or the schoolteacher leads to difficulties; there is practically no choice for the working mother than to take a leave of absence. The employer feels morally obliged to comply, but it tends to make females with children less reliable and more expensive as part of the work force. It is therefore that we hypothesize a negative influence of the presence of children. Finally, we assume that the number of hours supplied to the labor market may have a effect on w_{max} as there are economies of scale. However two workers with 20 hours a week each may be more productive than one worker who works 40 hours. So we assume a parabolic relation with h, but we make no assumptions about the direction of the effect. Like the other wage variables, we estimate w_{max} as a net amount. Evidently, the corresponding gross wage cost for the employer is much higher, but this is no problem as there is a one-to-one relationship between net and gross wage. In the Netherlands gross wage cost is about 1.8 times the net wage, accruing to the worker.

As noted before, we operationalize w_{max} on the basis of the worker's perception of her own labor market position. In the survey the following question was included:

> "If you would lose your job, do you expect to find another job within one year?"

Using a reasoning analogous to Van Praag (1991), we equate the answer "very improbable" to a subjective probability estimate of 1/8 to get another job; similarly the other categories are equated to subjective probability estimates of 3/8, 5/8 and 7/8 respectively. The frequencies and the assigned subjective probabilities of this question are given in table 2.5.

Table 2.5. Self-perceived probability of finding a job within one year for working females.

	Frequencies of finding a job		$P(u_i)$
N.A. (Unemployed)	3731		
Very probable	410	(19 %)	7/8
Probable	557	(25 %)	5/8
Improbable	646	(29 %)	3/8
Very improbable	596	(27 %)	1/8
Don't know	194		
Unknown (missing)	218		
	6352	(100 %)	

For workers we may assume that $w_{min} \leq w$, for otherwise they would quit. Hence, the doubt whether they would get a new job is caused by uncertainty whether the employer will pay at least their present wage. We assume that individuals assign a subjective probability distribution to their w_{max}; that subjective distribution of beliefs is assumed to be a lognormal distribution with log-variance σ_j^2 equal to that of the wage distribution in their reference group j, where the reference group is defined by common education and age bracket.[6] The σ_j^2 are estimated from the subsample of workers. The log-expectation of w_{max} is denoted by μ_{max}. It follows that $\ln(w_{max,t}) - \ln(w_t) \sim N(\mu_{max,t} - \ln(w_t), \sigma_j^2)$. We are interested in μ_{max}. The random $w_{max,t}$ can be expressed as $\ln(w_{max,t}) = \mu_{max,t} + \varepsilon$, so we have

$$\ln(w_{max,t}) - \ln(w_t) = \mu_{max,t} - \ln(w_t) + \varepsilon$$

Then there holds

$$P(\ln(w_{max,t}) - \ln(w_t) > 0) = P\left(\frac{\ln(w_t) - \mu_{max,\,t}}{\sigma_j} < \frac{\varepsilon}{\sigma_j} \right)$$

Someone, who answers that she would "very improbably" find a new job implies

$$P\left(\frac{\ln(w_t) - \mu_{max,\,t}}{\sigma_j} < \frac{\varepsilon}{\sigma_j} \right) = 1/8$$

and as ε/σ_j is assumed to be standard normal we get

$$[\mu_{max,t} - \ln(w_t)]/\sigma_j = u_{1/8}$$

or equivalently

$$\mu_{max,t} - \ln(w_t) = \sigma_j u_{1/8}$$

where $u_{1/8}$ stands for the normal quantile corresponding to 1/8. In a similar way we deal with the responses "improbable", "probable" and "very probable". Another question asked to workers was:

> "If you compare your wage with that of others, doing the *same* work, do you get less, about equal or more than the others?"

---

6 We distinguish six age brackets and five education levels, yielding 30 reference groups.

Table 2.6 shows the frequencies and the subjective probabilities to this question.

Table 2.6. Self-perceived probabilities of earning more or less than others doing the same work, for working females.

	Frequencies of comparing wages		$P(u_i)$
N.A. (Unemployed)	3731		
Less than others	421	(16 %)	5/6
About equal	2042	(80 %)	3/6
More than others	106	(4 %)	1/6
Unknown (missing)	52		
	6352	(100 %)	

Similarly we interpret this response as a "guesstimate" about the maximum wage distribution. Someone, who answers that she earns more than others in a similar job, believes that it will be difficult to get a comparable job for the same wage. The three response categories are assumed to correspond to subjective probabilities of 1/6, 3/6 and 5/6 of getting other work. If both answers are given, we have two estimates of $\mu_{max} - \ln(w_t)$ of which we take the average as our observation for $\mu_{max,t} - \ln(w_t)$. In case that one of the questions is answered we use the remaining response index. The categories "do not know" and "not relevant" are interpreted as non-response.
Active job-seekers responded to the question:

"Do you expect to find another job within one year?"

In table 2.7 the frequencies and the subjective probabilities to these questions are given.

Table 2.7. Self-perceived probability of finding a job within one year for job-seeking females.

	Frequencies of finding a job		$P(u_i)$
N.A. (Employed)	2621		
Very probable	8	(4 %)	7/8
Probable	22	(10 %)	5/8
Improbable	86	(41 %)	3/8
Very improbable	95	(45 %)	1/8
Don't know	23		
Unknown (missing, not seeking)	3497		
	6352	(100 %)	

If we compare table 2.5 and table 2.7, we see that job-seeking females are much more pessimistic about finding a job than working females are.
In a similar way as before this question is translated into a qualitative statement. For individuals, who are neither working nor job-seeking, $\ln(w_{max}) - \ln(w)$ can not be constructed.

In the following we write $\ln(w_{max})$ instead of μ_{max}, keeping in mind the interpretation as the log mean of a subjective believe distribution. The dependent variable will be $\ln(w_{max}) - \ln(w)$ as explained in section 4; in the summary below the w_{max}-equation is given explicitly.

Table 2.8 gives some insight of the average of $(\ln(w_{max}) - \ln(w))$ for working females and job-seekers per age bracket.

Table 2.8. Average $\sigma_j u_i$ of the subjective belief distribution of $\ln(w_{max})$ - $\ln(w)$ for working and job-seeking females.
(number of respondents in parentheses)

	Average $(\ln(w_{max})$ - $\ln(w)) = \sigma_j u_i$ workers		job-seekers	
Under 25	0.13	(380)	-0.84	(23)
25 to 29	0.04	(668)	-0.40	(42)
30 to 34	0.05	(424)	-0.49	(37)
35 to 44	-0.00	(744)	-0.55	(47)
45 to 54	-0.17	(250)	-0.38	(14)
55 and older	-0.01	(33)	-0.56	(2)
Whole sample	0.03	(2499)	-0.57	(165)

We note that there is a change of sign at 35 years of age for working females, indicating that older females perceive their position on the labor market as "well-paid".

Summarizing we assume the following log-linear model

$$(2.1) \qquad \ln(h) = \alpha_0 + \alpha_1 \ln(age-16) + \alpha_2(\ln(age-16))^2 + \alpha_3 \ln(fs) + \alpha_4 D_{ch4} + \alpha_5 D_{ch12} + \alpha_6 E_2 + \alpha_7 E_3 + \alpha_8 E_4 + \alpha_9 E_5 + \alpha_{10} \ln(w_m) + \alpha_{11} \ln(h_m) + \varepsilon_1$$

$$\begin{aligned}(2.2)\qquad \ln(w_{min}) &= \beta_0 + \beta_1\ln(age-16) + \beta_2(\ln(age-16))^2 + \beta_3\ln(fs) + \\ &\quad \beta_4 D_{ch4} + \beta_5 D_{ch12} + \beta_6 E_2 + \beta_7 E_3 + \beta_8 E_4 + \beta_9 E_5 + \\ &\quad \beta_{10}\ln(h) + \beta_{11}(\ln(h))^2 + \beta_{12}\ln(w_m) + \varepsilon_2\end{aligned}$$

$$\begin{aligned}(2.3)\qquad \ln(w) &= \gamma_0 + \gamma_1\ln(age-16) + \gamma_2(\ln(age-16))^2 + \gamma_3 D_{ch4} + \gamma_4 D_{ch12} \\ &\quad + \gamma_5 E_2 + \gamma_6 E_3 + \gamma_7 E_4 + \gamma_8 E_5 + \varepsilon_3\end{aligned}$$

$$\begin{aligned}(2.4)\qquad \ln(w_{max}) &= \delta_0 + \delta_1\ln(age-16) + \delta_2(\ln(age-16))^2 + \delta_3\ln(fs) + \\ &\quad \delta_4 D_{ch4} + \delta_5 D_{ch12} + \delta_6 E_2 + \delta_7 E_3 + \delta_8 E_4 + \delta_9 E_5 + \\ &\quad \delta_{10}\ln(h) + \delta_{11}(\ln(h))^2 + \varepsilon_4\end{aligned}$$

where:
- age = age in calendar years
- fs = family size = number of children + 2
- D_{ch4} = dummy for the presence of a child below 4 years of age
- D_{ch12} = dummy for the presence of a youngest child between 4 and 12 years old
- E_i = dummy for education level; $E_i = 1$ if the highest education level equals i (i = 2, ... ,5)
- w_m = wage of the main breadwinner
- h_m = working hours of the main breadwinner
- ε = the error vector.

The four equations (2.1) to (2.4) will be estimated simultaneously, where we assume that the errors ε follow a joint normal distribution $N(0,\Sigma)$.

We call this structural system the "exact" model, describing the structure of the exact observations (h, w_{min}, w, w_{max}). We do not have any reason to assume the covariance matrix Σ to be diagonal.[7] As ε will catch individual characteristics not observed by the economist we can take it for granted that a worker with a more than average performance will elicit a higher w_{max}. As she is probably aware of her position at the labor market, her w_{min} will be accordingly higher as well. To a lesser extent, as the institutional wage is rather inflexible, w will also reflect this tendency.

---

[7] This assumption of non-diagonality of the covariance matrix is one of the crucial differences between the model developed by Van Soest and Kapteyn (1989) and the model in this paper. As they assume diagonality, the likelihood can be expressed in well-known easily computable integrals.

3. The SEM-algorithm.[8]

The estimation of this model is not simple. If the data ($\boldsymbol{h}$, $\boldsymbol{w}_{min}$, $\boldsymbol{w}$, $\boldsymbol{w}_{max}$) were exactly observable for every respondent, the estimation of the exact model, described in section 2, would be a standard problem under the assumption that it is identified. It could be done by 2SLS, 3SLS or FIML. However, the data are only *exactly* observed for a subsample of the working women.

If an individual is not-exactly observed, generally we know that the vector $\boldsymbol{z}_t = (\boldsymbol{h}_t, \boldsymbol{w}_{min,t}, \boldsymbol{w}_t, \boldsymbol{w}_{max,t})$ is in a specific proper subset $A_t \subset \mathbb{R}^4$. For the set of unemployed we know that they are voluntarily unemployed ($\boldsymbol{w}_{min} > \boldsymbol{w}$) or involuntarily unemployed ($\boldsymbol{w}_{max} < \boldsymbol{w}$) or that both inequalities hold simultaneously. Thus we observe the random event $\boldsymbol{z} \in \mathbf{A}$, where $\mathbf{A} = \{z | \boldsymbol{w}_{min} > \boldsymbol{w} \text{ or } \boldsymbol{w}_{max} < \boldsymbol{w}\}$; we speak of a random *set-valued* observation $\mathbf{A}_t$. As $\boldsymbol{z}_t$ is assumed to be normally distributed over $\mathbb{R}^4$, it follows that in such a case $\boldsymbol{z}_t \in A_t$ is an informative (i.e. non-trivial) observation ($P(\boldsymbol{z}_t \in A_t) < 1$), although it is clearly less informative than an exact (pointwise) observation. For a worker we do not have an exact observation, if she did not fill in the questions on w_{min} or w_{max}. For her we observe h and w and we assume in that case that w_{min} and w_{max} are such that $\boldsymbol{z}_t \in \{z | \boldsymbol{w}_{min} < \boldsymbol{w} < \boldsymbol{w}_{max}\} = \bar{\mathbf{A}}_t$. So, next to the categorical non-response of the unemployed, also individual workers may cause set-valued observations.

In this section we outline the SEM-algorithm to derive ML-estimators for set-valued observations, where we shall use a setting borrowed from Van Praag and Hop (1987). Basically, we replace the need for computing complex integrals by an iterative simulation procedure, being possible only in this computer era. First we consider ML-estimation on complete data and ML-estimation on incomplete or set-valued data. Then we outline the EM-algorithm as a way to compute ML-estimators for set-valued observations. Finally we come to the SEM-algorithm as a further modification of EM. In Ruud (1991) a SEM-algorithm is proposed which is similar to our approach; the conditional expectations are replaced by approximations by means of simulations. Ruud suggests to simulate either by means of an acceptance-rejection method or by using one sample of untruncated

---

8 In the remainder of this paper random variables and sets will be printed in bold. Readers who are not interested in the estimation technique may skip sections 3 and 4.

simulations. Following his own evaluation it may be argued that the former method is inefficient as many drawings may be needed to obtain one suitable value whereas the second method introduces correlation among observations. Our method, as described below, uses direct simulations from the appropriate conditional distributions without the need for rejecting drawings.

Our model for the basic event will be the following. The basic event is a random point $(\boldsymbol{x}, \boldsymbol{z}) \in (\mathbb{R}_x, \mathbb{R}_z)$, where $\boldsymbol{x}$ is a vector of exogenous variables and $\boldsymbol{z}$ is a vector of endogenous variables. In our case $\boldsymbol{z} = (\boldsymbol{h}, \boldsymbol{w}_{min}, \boldsymbol{w}, \boldsymbol{w}_{max})$, while $\boldsymbol{x}$ stands for the other variables in the model equations (2.1) - (2.4). The data will be assumed i.i.d. with a probability density element

$$(3.1) \qquad g(\mathrm{x},\ \mathrm{z};\ \theta) = p(\mathrm{x})\, f(\mathrm{z}\,|\,\mathrm{x};\ \theta)$$

Thus, we assume that the exogenous vector $\boldsymbol{x}$ is random with a marginal density $p(.)$. The second factor in (3.1) is the *conditional* density of $\boldsymbol{z}$, given $\boldsymbol{x}$. The functional specification of $f(.\,|\,.;.)$ is known except for the value θ_0 of the parameter vector θ. The density $g(\mathrm{x}, \mathrm{z}, \theta)$ defines a probability measure $\mathrm{P}(\theta) = \mathrm{P}_x \times \mathrm{P}_{z|x}(\theta)$ on the σ-algebra $\mathcal{C} = \mathcal{B}(\mathcal{C}_x \times \mathcal{C}_z)$ of Borel sets in $\mathbb{R}_x \times \mathbb{R}_z$, generated by the arguments.

Moreover we assume the usual *regularity* conditions (cf. Wilks (1962) p.345 a.f., Amemiya (1985) Ch.4) :

a. The parameter space Θ is a compact subset of $\mathbb{R}^k$ and θ_0, the true value of θ, is an interior point of Θ
b. The function $\ln\, g(\mathrm{x}, \mathrm{z}; \theta)$ is a $\mathrm{P}(\theta_0)$-measurable function $\forall\theta \in \Theta$ and dominated by a $\mathrm{P}(\theta_0)$-integrable function $\mathrm{h}(\mathrm{x}, \mathrm{z})$, i.e., $|\ln\, g(\mathrm{x}, \mathrm{z}; \theta)| \le \mathrm{h}(\mathrm{x}, \mathrm{z})\ \forall\theta \in \Theta$
c. The function $\ln\, g(\mathrm{x}, \mathrm{z}; \theta)$ is differentiable with respect to θ
d. The score function $l(\mathrm{x}, \mathrm{z}; \theta) = \frac{\partial}{\partial\theta} \ln\, g(\mathrm{x}, \mathrm{z}; \theta)$ is almost everywhere continuous in θ and dominated by a $\mathrm{P}(\theta_0)$-integrable function
e. The first-order derivatives of $l(\mathrm{x}, \mathrm{z}; \theta)$ with respect to θ exist and are dominated as well.

These assumptions ensure that we may interchange the limiting operations of integration and differentiation (see e.g. Billingsley (1986), Sect. 16 or Amemiya (1985), Ch.1).

If we assume that there is a one-to-one relationship between θ and $\mathrm{P}(\theta)$, the parameter vector θ is called *identifiable* on the basis of observations (x,z). Another characterization of identifiability can be given in terms of the population log-likelihood[9]

$$\mathcal{L}(\theta{:}\theta_0) = \int \ln(g(x,\ z;\ \theta))\ g(x,\ z;\ \theta_0)\ dxdz \tag{3.2}$$

$$= \mathrm{E}_{\theta_0} \ln\ g(\boldsymbol{x},\ \boldsymbol{z};\ \theta)$$

The expression $\mathcal{L}(\theta{:}\theta_0)$ may be interpreted as a "directed divergence" which measures the divergence between $\mathrm{P}(\theta)$ and $\mathrm{P}(\theta_0)$ (Kullback (1959)). $\mathcal{L}(\theta{:}\theta_0)$ reaches its maximum for $\theta = \theta_0$. We call this property the *maximum likelihood (ML) property.* If there is a unique maximizing θ_0 the model is uniquely identified. In that case we call $\mathrm{P}(\theta)$ identifiable. If there are several maximizing values $\theta \in \Theta_0$, all $\theta \in \Theta_0$ define the same measure P on $\mathcal{C}$.
A consistent estimate of (3.2) may be assessed from the sample of basic events $\{(\boldsymbol{x}_t,\ \boldsymbol{z}_t)\}_{t=1}^{T}$, viz., the sample log-likelihood

$$\mathcal{L}_T(\theta{:}\theta_0) = \frac{1}{T} \sum_{t=1}^{T} \ln\ g(\boldsymbol{x}_t,\ \boldsymbol{z}_t;\ \theta) \tag{3.3}$$

Its maximizing value $\hat{\theta}_0$ is the consistent ML-estimator of θ_0. The identifiability of an economic model is ensured by the usual conditions.

Consider now the case of set-valued observations $(\boldsymbol{x},\mathrm{A})$ where the sets **A** generate a coarser σ-algebra $\mathcal{A} \subset \mathcal{C}$ on which the corresponding probability measure $\mathrm{P}_{\mathcal{A}}(\theta)$ is defined as the restriction of $\mathrm{P}(\theta)$ to $\mathcal{A}$. It is not always true that this mapping $\mathrm{P} \to \mathrm{P}_{\mathcal{A}}$ is one-to-one. Consider the probit situation where $\mathcal{A}$ consists of two atoms $\mathrm{A}_1 = (-\infty,0]$ and $\mathrm{A}_2 = (0,\infty)$ respectively. $\mathcal{A}$ defines a partition of the real line $\mathbb{R}$. If the underlying probability distribution is $\mathrm{N}(\mu,1)$ we have

$$\mathrm{p}_1(\mu) = \mathrm{N}(0;\ \mu,\ 1) \text{ and } \mathrm{p}_2(\mu) = 1 - \mathrm{p}_1(\mu)$$

If we know $\mathrm{p}_1(\mu)$ we may solve for μ. However, if the underlying distribution is $\mathrm{N}(\mu, \sigma^2)$ we can not retrieve μ *and* σ^2 from this two-set partition. If $\mathcal{A}$ contains at least *three* atoms, μ and σ^2 are uniquely identifiable. If there are more $(\mathrm{k} > 3)$ atoms, we have $\mathrm{k} - 2$ dependent equations. This suggests the

---

9 Contrary to convention we introduce the arguments θ and θ_0. This will be useful in the subsequent analysis.

number condition that the number of atoms should be larger than k, the number of independent parameters.[10] We shall assume that $\mathcal{A}$ is *fully informative*, i.e. θ is uniquely identifiable from $P_{\mathcal{A}}(\theta)$.

The probability density on $\mathcal{A}$ will be denoted by $g(x, A; \theta)$ and we have[11]

$$(3.4) \qquad g(x, A; \theta) = p(x) \int_A f(z|x; \theta)\, dz$$

where the integral is either the probability P(A) or a line- or surface integral if $P(A) = 0$.

Analogous to (3.2) we may define the population log-likelihood based on the set-valued observations as

$$(3.5) \qquad \mathcal{L}_{\mathcal{A}}(\theta{:}\theta_0) = \iint \ln(g(x, A; \theta))\, g(x, A; \theta_0)\, dA dx$$

$$= E_{\mathcal{A},\theta_0}[\ln g(\boldsymbol{x}, A; \theta)]$$

Now we define identifiability of θ from $\mathcal{A}$ more formally by the requirement that $\mathcal{L}_{\mathcal{A}}(\theta{:}\theta_0)$ has one global maximum at θ_0.

The empirical estimation lies at hand by maximizing

$$(3.6) \qquad \mathcal{L}_{\mathcal{A},T}(\theta{:}\theta_0) = \frac{1}{T} \sum_{t=1}^{T} \ln g(\boldsymbol{x}_t, A_t; \theta)$$

yielding $\hat{\theta}_{\mathcal{A}}$.

The technical problem is that in a case like the one considered in this paper the integrals (3.4) cannot be assessed, at least not at such a large scale as necessary. Therefore we choose the circumventing way opened by the EM-algorithm (Dempster, Laird and Rubin (1977)). We have

---

[10] Strictly speaking the number condition is necessary but not sufficient. Consider the case where $(X,Y) = N(\mu,\Sigma)$ and where the partition $\mathcal{A}$ consists only of parallel sets $\{(x,y)|x \in A, -\infty<y<\infty\}$. In such a case only the marginal parameters μ_x and Σ_{xx} can be estimated. A sufficient and necessary condition is that there is a set of at least k independent equations $P(A_i|\theta) = p_i$ (i=1, ... ,k), where k is the number of independent parameters, such that a unique $\hat{\theta}$ can be solved from it. If k is larger than the number of parameters, we have a case of over-identification and the parameter values are found by minimizing a least-squares criterion.

[11] Then $g(x,A;\theta)$ is the Radon-Nikodym derivative with respect to $P_{\mathcal{A}}$.

$$g(x,\ z;\ \theta) = g(x,\ A;\ \theta)\ f(z|z \in A,\ x;\ \theta)$$

where the last factor is the conditional density of z given $z \in A$.
Then we have[12]

$$\mathcal{L}(\theta{:}\theta_0) = \iiint \ln(g(x,\ z;\ \theta))\ g(x,\ A;\ \theta_0)\ f(z|z \in A,\ x;\ \theta_0)\ dzdAdx \tag{3.7}$$

$$= \iint \ln(g(x,\ A;\ \theta))\ g(x,\ A;\ \theta_0)\ dAdx +$$

$$\iint \left(\int_A \ln(f(z|z{\in}A,x;\theta))\ f(z|z{\in}A,x;\theta_0)\ dz\right)\ g(x,\ A;\ \theta_0)\ dAdx$$

$$= \mathcal{L}_{\mathcal{A}}(\theta{:}\theta_0) + E_{\mathcal{A},\theta_0}\left(\mathcal{L}(\theta{:}\theta_0|A)\right).$$

where $\mathcal{L}(\theta{:}\theta_0|A)$ is the population log-likelihood of the *conditional* density of z given $z \in A$, i.e,

$$\mathcal{L}(\theta{:}\theta_0|A) = \int \ln(f(z|z \in A,\ x;\ \theta))\ f(z|z \in A,\ x;\ \theta_0)\ dz$$

$$= E_{\theta_0}\left(\ln f(z|z \in A,\ x;\ \theta)\ |z \in A\right)$$

We notice that the three terms $\mathcal{L}(.)$, $\mathcal{L}_{\mathcal{A}}(.)$ and $E_{\mathcal{A},\theta_0}$ in (3.7) reach their joint maximum at θ_0 due to the ML-property which holds for the three terms separately. This property is exploited by the EM-algorithm. Each incomplete observation A_t is "completed" by using the conditional expectation $\bar{z}_t^{(n)} = E(z_t|z_t \in A_t;\ \theta^{(n)})$ where $\theta^{(n)}$ is a previous estimate of θ_0. Then maximization of $\mathcal{L}_{\mathcal{A}}(\theta{:}\theta_0)$ is replaced by maximization of $\mathcal{L}(\theta{:}\theta_0)$ where the latent z_t, corresponding to A_t, is replaced by its conditional expectation $\bar{z}_t^{(n)}$.
Thus the EM-algorithm maximizes

$$\bar{\mathcal{L}}_T(\theta;\ \theta_0) = \frac{1}{T}\sum_{t=1}^{T} \ln\ g(x_t,\ \bar{z}_t^{(n)};\ \theta^{(n)}) \tag{3.8}$$

with respect to θ, yielding a $(n+1)^{st}$ estimate $\theta^{(n+1)}$.
Dempster, Laird and Rubin (1977) show that this iteration method yields a sequence $\{\theta^{(n)}\}$ converging to θ_0. Actually the EM-algorithm is only a

---

12 In (3.7) we use integration over and expectation with respect to the atoms A of $\mathcal{A}$, although this type of integral is not formally defined in this paper.

computational device to compute the ML estimator $\hat{\theta}$. It has the same variance-covariance matrix as the traditional ML-estimator, viz.,

$$\frac{1}{T}\left[E \left[\frac{\partial}{\partial\theta} \ln g(x_t, A_t; \theta_0)\right]\left[\frac{\partial}{\partial\theta'}\ln g(x_t, A_t; \theta_0)\right]\right]^{-1}$$

In many cases it is impossible to evaluate $\bar{z}_t^{(n)}$, as it also involves weird integrals. However, it may be possible to simulate from the conditional distribution $f(z|z \in A_t, x; \theta^{(n)})$. Let us denote a simulated observation from $f(z|z \in A_t, x; \theta^{(1)})$ by $z_t^{(1)}$ where $\theta^{(1)}$ is a first conjecture of θ_0. Then the conjectured sample log-likelihood based on the completed sample $\{(x_t, z_t^{(1)})\}_{t=1}^{T}$ is given by

$$\text{(3.9)} \qquad \mathcal{L}_T^{(1)}(\theta:\theta_0) = \frac{1}{T}\sum_{t=1}^{T} \ln g(x_t, z_t^{(1)}; \theta)$$

The attractivity of (3.9) is that it is a likelihood on (pseudo-) exact observations which can be easily assessed. Similar to the case of the EM-algorithm, with similar assumptions, it can be shown for the SEM-algorithm that the sequence $\{\theta^{(n)}\}_{n=0}^{\infty}$ converges to θ_0.

The SEM-algorithm

As (3.9) is a consistent estimator of (3.8) the sample analog of the SEM-algorithm is clear.

1 Select an initial estimate $\hat{\theta}^{(1)}$;

2 Draw $z_t^{(1)}$ from A_t according to the conditional density $f(z|z \in A_t, x_t; \hat{\theta}^{(1)})$;

3 Calculate (3.9), i.e. $\mathcal{L}_T^{(1)}(\theta:\theta_0)$;

4 Maximize (3.9) with respect to θ, yielding $\hat{\theta}^{(2)}$;

5 Go to 2 and repeat the loop until convergence.

If it converges we have the consistent estimate $\theta_{\mathcal{A}}^{(\infty)}$.

Statistical properties of $\theta_{\mathcal{A}}^{(\infty)}$

By this procedure we have found a consistent estimator $\theta_{\mathcal{A}}^{(\infty)}$ of θ_0 and a consistent estimator of the population log-likelihood

$$\text{(3.12)} \qquad \mathcal{L}_{\mathcal{A}}(\theta:\theta_0) = \mathcal{L}(\theta:\theta_0) - E_{\mathcal{A},\theta_0}\left(\mathcal{L}(\theta:\theta_0|A)\right)$$

Consider the simulated score

$$(3.13)\qquad \boldsymbol{u}_t = \frac{\partial}{\partial\theta} \ln g(\boldsymbol{x}_t, \boldsymbol{z}_t^{(\infty)}; \theta_0)$$

We have used this score in lieu of the traditional score $\frac{\partial}{\partial\theta} \ln g(\boldsymbol{x}_t, \mathrm{A}_t; \theta_0)$ corresponding to (3.6). Both have expectation zero but the simulated score has a larger variance, as we have

$$\frac{\partial}{\partial\theta} \ln g(\boldsymbol{x}_t, \boldsymbol{z}_t; \theta_0) = \frac{\partial}{\partial\theta} \ln g(\boldsymbol{x}_t, \mathrm{A}_t; \theta_0) + \frac{\partial}{\partial\theta} \ln f(\boldsymbol{z}_t | \boldsymbol{z}_t \in \mathrm{A}_t, \boldsymbol{x}_t; \theta_0)$$

Notice that the expectations of the three terms are zero, as they are derivatives of the three terms in (3.7) which all reach their maximum at $\theta = \theta_0$. The third term may be interpreted as an additional independent error. We have the following decomposition of the score covariance matrix

$$\begin{aligned}(3.14)\qquad \mathrm{E}\boldsymbol{u}_t\boldsymbol{u}_t' &= \Sigma \\ &= \mathrm{var}\left[\frac{\partial}{\partial\theta} \ln g(\boldsymbol{x}_t, \mathrm{A}_t; \theta_0)\right] + \mathrm{E}\ \mathrm{var}\left[\frac{\partial}{\partial\theta} \ln f(\boldsymbol{z}_t|\mathrm{A}_t, \boldsymbol{x}_t; \theta_0)\right] \\ &= \Sigma_{between} + \Sigma_{within}\end{aligned}$$

where the covariance matrices $\Sigma_{between}$ and Σ_{within} are implicitly defined. Applying a well-known result (see e.g. Amemiya (1985), p. 111) we get

$$(3.15)\qquad \mathrm{var}\ (\hat{\theta}_{\mathcal{A}}^{(\infty)}) = \frac{1}{\mathrm{T}} \left[\frac{\partial^2}{\partial\theta\partial\theta'}\mathcal{L}_{\mathcal{A}}(\theta_0; \theta_0)\right]^{-1} \mathrm{var}(\boldsymbol{u}_t) \left[\frac{\partial^2}{\partial\theta\partial\theta'}\mathcal{L}_{\mathcal{A}}(\theta_0; \theta_0)\right]^{-1}$$

According to ML-theory

$$\begin{aligned}(3.16)\qquad \frac{\partial^2}{\partial\theta\partial\theta'}\mathcal{L}_{\mathcal{A}}(\theta_0; \theta_0) &= -\mathrm{E}\left[\frac{\partial}{\partial\theta} \ln g(\mathrm{X}_t, \mathrm{A}_t; \theta_0)\right]\left[\frac{\partial}{\partial\theta'}\ln g(\mathrm{X}_t, \mathrm{A}_t; \theta_0)\right] \\ &= -\Sigma_{between}\end{aligned}$$

and hence

$$(3.17)\qquad \mathrm{var}\ (\hat{\theta}_{\mathcal{A}}^{(\infty)}) = \frac{1}{\mathrm{T}} (\Sigma_{between})^{-1}(\Sigma_{between} + \Sigma_{within})(\Sigma_{between})^{-1}$$

As the orthodox ML estimator $\hat{\theta}_{\mathcal{A}}$ has $\mathrm{var}(\hat{\theta}_{\mathcal{A}}) = \frac{1}{\mathrm{T}} (\Sigma_{between})^{-1}$ we see that θ_0

estimated by the simulated score estimator ($\hat{\theta}_{\mathcal{A}}^{(\infty)}$) has a *larger* variance than $\hat{\theta}_{\mathcal{A}}$. The additional variance component caused by replacing the conditional expectation by one simulation is

$$\frac{1}{T}\,(\Sigma_{between})^{-1}\Sigma_{within}(\Sigma_{between})^{-1}$$

A way, which lies at hand now, is to simulate z_t from A_t not once for each observation t but K times, yielding simulated observations $\{(x_t,\ z_{tk})\}_{k=1}^{K}$ for each t. In that case we use the average $\frac{1}{K}\,\Sigma_{k=1}^{K} u_{tk}$ with mean covariance matrix

$$\text{(3.18)} \qquad \text{var}\,(u_{t.}) = (\Sigma_{between} + \frac{1}{K}\,\Sigma_{within})$$

It follows that for K-fold simulation we have

$$\text{(3.19)} \qquad \text{var}_K(\hat{\theta}_{\mathcal{A}}^{(\infty)}) = \frac{1}{T}\left[(\Sigma_{between})^{-1} + \frac{1}{K}\,(\Sigma_{between})^{-1}\Sigma_{within}(\Sigma_{between})^{-1}\right]$$

If $K \to \infty$, the SEM-algorithm becomes the EM-algorithm and we get $\lim_{K\to\infty} \text{var}_K(\hat{\theta}_{\mathcal{A}}^{(\infty)}) = \text{var}(\hat{\theta}_{\mathcal{A}})$, as intuition predicts.

Obviously the matrices $\Sigma_{between}$ and Σ_{within} have to be estimated by their sample counterparts.The technique may be easily generalized to the case where a part of the observations, say a fraction p_1, is observed in one mode, say the $\mathcal{A}_1$-mode and a part (say p_2) is observed in the $\mathcal{A}_2$-mode, where $\mathcal{A}_1$ and $\mathcal{A}_2$ are different (informative) σ-algebras. In that case in (3.8) $\mathcal{L}_{\mathcal{A}}(\theta:\theta_0)$ is replaced by $p_1\,\mathcal{L}_{\mathcal{A}_1}(\theta:\theta_0) + p_2\mathcal{L}_{\mathcal{A}_2}(\theta:\theta_0)$ and similarly $E_{\mathcal{A},\theta_0}\left(\mathcal{L}(\theta:\theta^{(1)}|A)\right)$ is replaced by the corresponding sum over the two σ-algebras. The method does nor change except that observations A_1 in the $\mathcal{A}_1$-mode are completed by drawing from $A_1 \in \mathcal{A}_1$ while observations $A_2 \in \mathcal{A}_2$ are completed by drawing from A_2. A trivial generalization to $\mathcal{A}_1, \dots, \mathcal{A}_r$ lies at hand. A specific case arises if $\mathcal{A}_1$ is fully informative, i.e. $\mathcal{L}_{\mathcal{A}_1}(\theta:\theta_0)$ has a unique global maximum θ_0, while $\mathcal{L}_{\mathcal{A}_2}(\theta:\theta_0)$ is not fully informative, i.e. $\mathcal{L}_{\mathcal{A}_2}(\theta:\theta_0)$ attains its maximum for a proper subset Θ_0 of the parameter space with the true $\theta_0 \in \Theta_0$. Also in that case $p_1\,\mathcal{L}_{\mathcal{A}_1}(\theta:\theta_0) + p_2\mathcal{L}_{\mathcal{A}_2}(\theta:\theta_0)$ attains a unique maximum at θ_0.
In the next section we apply the technique for the specific case of the female labor market.

4. The estimation method in detail.

The model in Section 2 may be written more formally as

$$\text{(4.1)} \qquad \begin{aligned} h &= \qquad\quad B_1 x + \varepsilon_1 \\ W &= \Gamma(h) + B_{-1} x + \varepsilon_{-1} \end{aligned}$$

where W stands for (w_{min}, w, w_{max}), $\Gamma(h) = \Gamma^{(1)}h + \Gamma^{(2)}h^2$, x stands for the other explanatory variables in the model, $\varepsilon_{-1} = (\varepsilon_2, \varepsilon_3, \varepsilon_4)$, and where the components of $z = (h, w_{min}, w, w_{max})$ are logarithms of the original variables. The covariance matrix of ε is Σ. We assume (4.1) to be identifiable.[13]

We reformulate the system (4.1) by subtracting the w-equation from the w_{min}-equation and the w_{max}-equation as we observe the differences rather than the variables w_{min} and w_{max} themselves. If $w_{min} - w = \ln(1-\alpha)$, then $1-\alpha$ is the tolerable reduction rate of the net wage, which we know for workers and active job-searchers. The difference $w_{max} - w$ is measured as described in section 2. Moreover we expect the correlation of w_{min} and w_{max} with the institutional wage w to be considerable; by using differences this problem can be mitigated.

The system becomes

$$\text{(4.2)} \qquad \begin{aligned} h &= \qquad\quad \tilde{B}_1 x + \varepsilon_1 \\ \tilde{W} &= \Gamma(h) + \tilde{B}_{-1} x + \tilde{\varepsilon}_{-1} \end{aligned}$$

where $\tilde{W}$ stands for $(\tilde{w}_{min}, w, \tilde{w}_{max})$ with $\tilde{w}_{min} = w_{min} - w$ and $\tilde{w}_{max} = w_{max} - w$ and $\tilde{\varepsilon}_{-1} = (\varepsilon_2-\varepsilon_3, \varepsilon_3, \varepsilon_4-\varepsilon_3)$. We denote $(h, \tilde{W}) = \tilde{z}$. The error covariance matrix $\tilde{\Sigma}$ of $\tilde{\varepsilon}$ is found by application of this linear difference transformation to Σ.

We start the iteration procedure as follows. We take the subset of observations on hours supply (the workers), estimate the h-equation by OLS, and calculate the prediction $\hat{h}$ for all individuals in the sample. This prediction is used as an explanatory variable in the other three equations,

---

[13] The model (4.1) would be identifiable when exactly observed. The fact that part of the observations are incomplete calls for additional conditions, viz., the number condition and the exclusion of parallel sets as described in footnote 10. In our case both conditions are satisfied.

which are estimated by separate OLS estimations on the corresponding sets of available observations. The matrix with the variances of the residuals on its diagonal is the first estimate of Σ. We denote this first estimate, consisting of B_1, B_{-1}, Γ and Σ, by $\theta^{(1)}$. On the basis of $\theta^{(1)}$ we simulate observations $\tilde{z}^{(1)}$ for non-workers and complete the observations for active job-searchers. In practice there are also workers who did not fill in all questions. Also for those individuals the observations are completed by simulating from the appropriate set.

For example, for an unemployed individual, we calculate the structural part of h and draw ε_1 from $(-\infty,\infty)$. Then we calculate the structural part of $\tilde{W}$, including $\hat{h}$, and draw $\tilde{\varepsilon}_{-1}$. The vector $\tilde{\varepsilon}_{-1} = (\varepsilon_2-\varepsilon_3, \varepsilon_3, \varepsilon_4-\varepsilon_3)$ is conditionally normally distributed, the condition being the previously drawn value of ε_1. Moreover we condition for being unemployed, i.e., the vector $\tilde{\varepsilon}_{-1}$ is drawn such that $\tilde{w}_{min} > 0$ and/or $\tilde{w}_{max} < 0$. This implies that we draw $\boldsymbol{z}$ such that $\boldsymbol{z} \in A = \{z|\boldsymbol{w}_{min} > \boldsymbol{w}$ and/or $\boldsymbol{w}_{max} < \boldsymbol{w}\}$.

Now we have a completed set of observations $\{\tilde{z}^{(1)}\}$ for all individuals in the sample. Then the equations (4.2) are estimated simultaneously yielding a second parameter estimate $\theta^{(2)}$ with a non-diagonal covariance matrix $\Sigma^{(2)}$. Again we complete the set of observations for non-workers yielding $\{\tilde{z}^{(2)}\}$ and re-estimation of (4.2) yields $\theta^{(3)}$. The procedure is repeated until convergence (see section 3).

The problem is now how to simulate from $\mathbf{A}_t$. In Van Praag and Hop (1987) it was shown how to draw them from the corresponding conditional distributions. (see also Gouriéroux and Monfort (1989)).

We reduce the simulation to fourfold single drawings. We notice that

$$(4.3) \qquad n(\tilde{z}) = n(h)\ n(w|h)\ n(\tilde{w}_{min}|h, w)\ n(\tilde{w}_{max}|h, w, \tilde{w}_{min})$$

where the first factor is a marginal distribution and the other factors stand for densities of conditional normal distributions. Given the parameters of n(.) those conditional densities can be calculated.[14] This gives the clue for

---

[14] It is well-known that if $X = (X_1, X_2)$ is $N(\mu,\Sigma)$ distributed, then X_1, given $X_2 = x_2$ is also normally distributed with conditional expectation

$$E(X_1 | X_2=x_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$$

and covariance matrix

$$\mathrm{cov}(X_1 | X_2=x_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

efficient simulation from a distribution on A_t, in the sense that no rejection of simulated values outside A_t is necessary.
First let us consider the case of a working female, who did not fill in w_{min} and w_{max}. For her we assume, lacking exact observation, that

$$\boldsymbol{w}_{min,t} \leq \boldsymbol{w}_t \leq \boldsymbol{w}_{max,t}$$

As we know $\boldsymbol{h}_t$ and $\boldsymbol{w}_t$, say $\boldsymbol{h}_t = \mathrm{h}$ and $\boldsymbol{w}_t = \mathrm{b}$, we may calculate the conditional density $\mathrm{n}(\tilde{w}_{min,t}|\boldsymbol{h}_t = \mathrm{h}, \boldsymbol{w}_t = \mathrm{b})$. As $\tilde{w}_{min,t} = w_{min,t} - \boldsymbol{w}_t$ has to be drawn from the interval $(-\infty,0)$ the conditional distribution function is

$$(4.4) \qquad \mathrm{P}\,(\tilde{\boldsymbol{w}}_{min,t} \leq \mathrm{a} \mid \tilde{\boldsymbol{w}}_{min,t} \leq 0,\ \boldsymbol{h}_t{=}\mathrm{h},\ \boldsymbol{w}_t{=}\mathrm{b}) = \frac{\mathrm{N}_{\tilde{w}_{min,t}}(\mathrm{a} \mid \boldsymbol{w}_t{=}\mathrm{b},\ \boldsymbol{h}_t{=}\mathrm{h})}{\mathrm{N}_{\tilde{w}_{min,t}}(0 \mid \boldsymbol{w}_t{=}\mathrm{b},\ \boldsymbol{h}_t{=}\mathrm{h})}$$

if $\mathrm{a} \leq 0$ and 0 otherwise.

Then the value $\tilde{w}_{min,t}$ is simulated by drawing $\boldsymbol{v}$ from a uniform distribution on $[0,1]$ and inverting the conditional distribution function (4.4). We denote the conditional value momentarily by a. In a similar way we subsequently draw $\tilde{w}_{max,t}$ from the interval $(0,\infty)$ by inverting the truncated distribution function

$$(4.5) \qquad \mathrm{P}\,(\tilde{\boldsymbol{w}}_{max,t} \leq \mathrm{c} \mid \tilde{\boldsymbol{w}}_{max,t} \geq 0,\ \boldsymbol{h}_t{=}\mathrm{h},\ \boldsymbol{w}_t{=}\mathrm{b},\ \tilde{\boldsymbol{w}}_{min,t}{=}\mathrm{a})$$

$$= \frac{\mathrm{N}_{\tilde{w}_{max,t}}(\mathrm{c} \mid \boldsymbol{h}_t{=}\mathrm{h},\ \boldsymbol{w}_t{=}\mathrm{b},\ \tilde{\boldsymbol{w}}_{min,t}{=}\mathrm{a}) - \mathrm{N}_{\tilde{w}_{max,t}}(0 \mid \boldsymbol{h}_t{=}\mathrm{h},\ \boldsymbol{w}_t{=}\mathrm{b},\ \tilde{\boldsymbol{w}}_{min,t}{=}\mathrm{a})}{1 - \mathrm{N}_{\tilde{w}_{max,t}}(0 \mid \boldsymbol{h}_t{=}\mathrm{h},\ \boldsymbol{w}_t{=}\mathrm{b},\ \tilde{\boldsymbol{w}}_{min,t}{=}\mathrm{a})}$$

if $\mathrm{c} \geq 0$ and 0 otherwise

In a similar way we may proceed for non-participating women. Knowing nothing at all we start with drawing one of the variables, say $\boldsymbol{h}_t$ from $(-\infty,\infty)$ by drawing from the marginal distribution of $\boldsymbol{h}_t$. Then we draw from the conditional distribution on $(-\infty,\infty)$ of $\boldsymbol{w}_t$, given our value of $\boldsymbol{h}_t$ just drawn. Similarly we draw from the conditional distribution of $\tilde{\boldsymbol{w}}_{min,t}$, given the drawn values of $\boldsymbol{h}_t$ and $\boldsymbol{w}_t$. Now there are two possibilities. Either we find $\tilde{\boldsymbol{w}}_{min,t} \leq 0$ or $\tilde{\boldsymbol{w}}_{min,t} > 0$. In the first case there must hold $\tilde{\boldsymbol{w}}_{max,t} < 0$, for otherwise the female would be in case A (as listed in the introduction) and observed to participate. Therefore we conclude the simulation by drawing $\tilde{w}_{max,t}$ on $(-\infty,0)$ from its conditional distribution. In the second case the female requires more than the institutional wage which is a sufficient reason for being observed as non-working. In that case we draw $\tilde{w}_{max,t}$ according to its conditional distribution from the interval $(-\infty,\infty)$. For another observation we may start by simulating $\tilde{w}_t$ and $\tilde{w}_{max,t}$ and proceed in a similar way for $\tilde{w}_{min,t}$.

Obviously the procedure depends on the selection order. In the iteration procedure we change the selection order per simulation and we simulate $K\,(\geq 1)$ observations per respondent. In this way for each individual we simulate one, two, three or four components (or none at all) according to the number of unobserved variables.

A final word about our assumption $w \leq w_{max}$, which we assume to hold for workers. Our operational definition of w_{max} does not ensure that $w \leq w_{max}$ for all workers. It may be that $w > w_{max}$ for specific workers, who have answered that they would very improbably find a new job when fired. Actually the same phenomenon is found in the probit situation where a number of observations are incorrectly predicted to belong to class zero, while they are found to be in class one.

5. Empirical Results.

The data set has been collected in 1983 by Van Praag and Hagenaars among the readership of Dutch regional dailies.[15] For this study we consider only women with a working partner, most of whom are formally married. In the sample, which consists of 6352 observations, about 41% of the females participate in the labor market. In Section 2 we described the model and its variables. As mentioned the observed wages are *net* wages. The respondent has been presented with the possibility to specify the pay period as being a week, a month, four weeks or a year. Thus it was possible for the respondent to give an amount of salary as she perceives it without recalculations for a different reference period than she is used to. For a similar reason the net wage has been asked to avoid the problem of imputing tax and social premium reductions which are fairly complex in the Dutch tax system in 1983. It is evident that the employer pays much more than the net wage. However, there is a unique relationship (given exogenous variables) between the net wage accruing to the worker and the gross wage cost paid by the employer. So properly speaking the net wage must be seen as a wage cost indicator to the employer.

Hours are asked per week. Because this information has been asked in a structure which is adaptable to the respondent's frame of reference, we do not feel that there will be a large division bias (Borjas (1980)). For both working and non-working respondents we used the predicted number of hours as explanatory variable in the wage equations to avoid simultaneity bias.

In table 5.1^{A} we present the estimates of the reformulated model of section 4, i.e., the dependent variables are h and $\tilde{W}$. The number of simulations k starts at 2 and is gradually increased up to 9 at the final iteration. The corresponding standard deviations are given in parentheses. Notice that some effects, especially for education, are non-significant in the second and the fourth column. This means that education differences have approximately the same effect on w_{min} and w_{max} as on w. The estimates in table 5.1^{B} yield the results for the original forms (2.1) - (2.4), derived by inversion of the difference transformation. The latter estimates will be discussed in the remainder of this paper.

---

[15] The sample is fairly representative for the Dutch population (see Van Duin and Hagenaars (1984)). The observations have been reweighed to improve the representativity. This dataset is also used by Homan (1988) and by Renes and Hagenaars (1989).

Table 5.1[A]. Estimate of the model for the partner.

(standard deviations in parentheses)

(*: significant at the 5% level)

	$\ln(h)$	$\ln(w_{min})$-$\ln(w)$	$\ln(w)$	$\ln(w_{max})$-$\ln(w)$
intercept	2.3815* (0.3753)	2.0326* (0.6480)	1.4096* (0.1634)	0.7450* (0.3612)
$\ln(age-16)$	0.4100* (0.1982)	0.0006 (0.1005)	0.5009* (0.1275)	0.4021* (0.0487)
$\ln^2(age-16)$	-0.1217* (0.0362)	-0.0044 (0.0233)	-0.0739* (0.0237)	-0.0801* (0.0121)
$\ln(fs)$	-0.5176* (0.0397)	-0.0841 (0.0750)	-	-0.0536 (0.0436)
D_{ch4}	-0.6124* (0.0331)	-0.1255 (0.0987)	0.1019* (0.0220)	-0.1690* (0.0526)
D_{ch12}	-0.3483* (0.0305)	-0.0820 (0.0550)	-0.0523* (0.0169)	-0.0583 (0.0304)
E_2	0.1515* (0.0271)	-0.0307 (0.0262)	0.0939* (0.0167)	0.0131 (0.0146)
E_3	0.1982* (0.0329)	-0.0202 (0.0344)	0.1462* (0.0213)	0.0244 (0.0195)
E_4	0.2121* (0.0453)	-0.0367 (0.0404)	0.2581* (0.0281)	0.0371 (0.0232)
E_5	0.1221* (0.0310)	-0.0374 (0.0259)	0.3158* (0.0208)	0.0218 (0.0153)
$\overline{\ln(h)}$	-	-1.2604* (0.3011)	-	-0.8212* (0.1733)
$\overline{\ln^2(h)}$	-	0.1850* (0.0425)	-	0.1375* (0.0255)
$\ln(w_m)$	0.0669 (0.0382)	0.0432* (0.0173)	-	-
$\ln(h_m)$	0.2212* (0.0609)	-	-	-
Covariances	0.2670			
	-0.0024	0.0427		
	-0.0524	-0.0178	0.1067	
	-0.0002	0.0061	-0.0052	0.0237

Table 5.1[B]. Estimate of the model for the partner.
(standard deviations in parentheses)
(*: significant at the 5% level)

	$\ln(h)$	$\ln(w_{min})$	$\ln(w)$	$\ln(w_{max})$
intercept	2.3815*	3.4422*	1.4096*	2.1546*
	(0.3753)	(0.6572)	(0.1634)	(0.3956)
$\ln(age-16)$	0.4100*	0.5015*	0.5009*	0.9030*
	(0.1982)	(0.1345)	(0.1275)	(0.1302)
$\ln^2(age-16)$	-0.1217*	-0.0738*	-0.0739*	-0.1540*
	(0.0362)	(0.0284)	(0.0237)	(0.0250)
$\ln(fs)$	-0.5176*	-0.0841	-	-0.0536
	(0.0397)	(0.0750)		(0.0436)
D_{ch4}	-0.6124*	-0.0236	0.1019*	-0.0671
	(0.0331)	(0.0984)	(0.0220)	(0.0558)
D_{ch12}	-0.3483*	-0.1343*	-0.0523*	-0.1106*
	(0.0305)	(0.0558)	(0.0169)	(0.0337)
E_2	0.1515*	0.0632*	0.0939*	0.1070*
	(0.0271)	(0.0283)	(0.0167)	(0.0196)
E_3	0.1982*	0.1260*	0.1462*	0.1706*
	(0.0329)	(0.0380)	(0.0213)	(0.0284)
E_4	0.2121*	0.2214*	0.2581*	0.2952*
	(0.0453)	(0.0449)	(0.0281)	(0.0343)
E_5	0.1221*	0.2784*	0.3158*	0.3376*
	(0.0310)	(0.0293)	(0.0208)	(0.0244)
$\overline{\ln(h)}$	-	-1.2604*	-	-0.8212*
		(0.3011)		(0.1733)
$\overline{\ln^2(h)}$	-	0.1850*	-	0.1375*
		(0.0425)		(0.0255)
$\ln(w_m)$	0.0669	0.0432*	-	-
	(0.0382)	(0.0173)		
$\ln(h_m)$	0.2212*	-	-	-
	(0.0609)			
Covariances/ Correlations	0.2670	-0.3144	-0.3183	-0.2939
	-0.0548	0.1138	0.8068	0.8970
	-0.0524	0.0889	0.1067	0.7684
	-0.0526	0.0898	0.1015	0.1200

Covariances in the lower triangle, correlations in the upper triangle

The hours equation.

Not unexpectedly the relation of hours supply and (age–16), taken as a proxy of experience, is at first increasing with a top at an age of about 21. At 18 the age elasticity is about 0.33 and at 36 the elasticity is about 0.05. At 49 it is -0.02. The household size has a strong negative influence. On top of that very small children and children between 4 and 12 years old have an additional negative impact. Hours supply is higher for women with education above the base level, but there is no significant difference between the various levels above the base level. The wage of the husband has hardly any influence, but female hours supply is increasing with respect to the working hours of the husband.

The w_{min} equation.

Also this variable is quadratic in age/experience with a maximum at an age of 46. The household size has a non-significant effect. Contrary to intuition the presence of very small children decreases the minimum wage required, however, the effect is nonsignificant. The presence of older children also decreases the mother's minimum wage. This counter-intuitive behavior may be explained, when we realize that children at home cause a need for more income but also for more time at home to care. The first factor reduces the minimum wage required, while the second one increases the minimum wage. For older children the need for money becomes stronger, while the need for time to care for the children is reduced. As expected, education, especially at the highest levels, increases w_{min}. Changes in the main breadwinner's wage rate have a slight positive impact on the minimum wage level of the female. Finally we find a minimum in hours supply at ca. 30 hours a week.

The wage equation.

The institutional wage is increasing with age until the age of 46. The tax and family allowance system has the effect that the presence of small children in the household increases the net wage while the presence of older children has the reverse impact. Naturally the institutional wage increases with the level of education. The difference in net wage levels in Holland is rather small as a consequence of the rather progressive tax system. Notice however that we consider female labor; for male workers the differences would be much larger.

The w_{max} equation.

This is the most interesting equation in our opinion. It reflects marginal labor productivity. First we notice that the maximum wage is a quadratic function of the age variable with a maximum at 35. It reflects the fact that firms are very eager to hire young females, especially if they have some experience. We find the effect of family size to be small; the presence of children has a negative effect on the employer's maximum wage. This shows that firms are not eager to hire females with household chores. Finally, we find that firms do not easily hire part-time workers; the minimum of the quadratic relation with respect to hours lies at 20 hours a week.

In table 5.2 we consider a classification of the sample with respect to the employment situation and for the unemployed with respect to the cause of unemployment. When is the official wage too low or too high and for how many cases is the unemployment voluntary or caused by a refusal by the firm? Finally, when are both firm and female unwilling to match? This yields the six classes A to F given in the introduction. In table 5.2 the fractions of all respondents in each of the six classes are presented, as well as for the actually employed and unemployed separately. The fractions are found by simulating from the four dimensional distribution of the hours supply and wage rates as given in table 5.1[B] by counting the relative frequencies. Simultaneously we calculated the fractions of the (potential) hours supply corresponding to those individuals. This is also tabulated in table 5.2.

Table 5.2. A classification of employment situation.
(# is % respondents; h is % hours supply)

	A	B	C	D	E	F	N
# Total	23.9	30.1	2.3	13.6	3.7	26.3	(6352)
h	31.2	29.2	2.1	10.8	3.1	23.6	(131081)
# Employed	58.0	20.4	-	-	-	21.6	(2621)
h	60.5	20.6	-	-	-	18.9	(67671)
# Unemployed	-	36.9	4.0	23.2	6.3	29.6	(3731)
h	-	38.3	4.4	22.3	6.4	28.6	(63410)

A to F are the (un)employment classifications, described in section 1,
N is the total number of respondents and their hours labor supply.

We see by adding the percentages under B and C that for about 41% of the unemployed it would be possible to match supply and demand when the wage paid could differ from the institutional wage rate. This amounts to an increase of employment by a factor of 1.6. Similarly there holds that about 43% of the potential hours supply of the presently unemployed could be matched to demand, if the legal restriction on the institutional wage could be circumvented.
We notice that only 23.9% of the population would be re-employable although 41% is employed in the sample. Since we have asked the worker: "If you would lose your present job, do you expect to find another job within one year?", we take into account that many workers would not be hired anew, if they would lose their present job. This is shown in table 5.2 by the classifications B and F (w_{max} is too low) of the working females. The difference of 17% between the real employment and the predicted re-employment rates has to be ascribed to the insider-outsider effect.

The question is whether the estimated model, describing a world *without* insider-outsider effect, has predictive value for the "real" world *with* an insider-outsider effect. Let us consider the dummy D_E which equals one for actually employed and equals zero for actually unemployed. Two obvious explanatory variables are the predicted differences between the institutional wage and the minimum wage $(w_{min} - w)$ and the maximum wage $(w_{max} - w)$ respectively. The chance on being employed $P(D_E > 0)$ should increase if $(w_{min} - w)$ becomes more negative and, similarly, $P(D_E > 0)$ should increase with $(w_{max} - w)$ becoming more positive.
Hence we estimate the following probit equation

$$P(D_E > 0) = N(\alpha_0 + \alpha_1(w_{min} - w) + \alpha_2(w_{max} - w) > 0)$$

We find (standard errors in parentheses)

$$\underset{(0.0551)}{\alpha_0 = 0.4451} \quad \underset{(0.4350)}{\alpha_1 = -2.4355} \quad \underset{(0.2982)}{\alpha_2 = 11.8922} \qquad (N = 6352)$$

Indeed α_1 and α_2 have the expected signs and are highly significant. Using the estimates of the variances of $(w_{min} - w)$ and $(w_{max} - w)$ from Table 5.1[A] (0.0427 and 0.0237 respectively) we find standardized values for α_1 and α_2 of -0.5033 and 1.8308 respectively. It appears after standardization that labor *demand* is the more important factor. If $w_{min} - w = 0$ and $w_{max} - w = 0$ the original model of table 5.1[A] would predict 50% chance on being employed. However the actual chance on employment is predicted by $N(\alpha_0)=N(0.4451)\approx$ 67%. The difference is 17%. The α_0 reflects the insider-outsider effect. It is a premium given to the present worker; an advantage compared to those who have to enter a new job. Remaining in a present job is easier than entering a new one.

In table 5.3 we compare the average chances of being employed *with* (the observed employment rate) and *without* (the simulated employment rate) taking into account the insider-outsider effect for Dutch women at various ages and education levels.

Table 5.3. The insider advantage for females of various age brackets and education levels.

Age		Education level 1	2	3	4	5	
Under 25	P_{with}	46	73	72	76	77	66
	$P_{without}$	30	45	54	47	41	43
25 to 29	P_{with}	23	42	54	53	70	47
	$P_{without}$	18	24	29	44	35	26
30 to 34	P_{with}	22	30	37	29	44	30
	$P_{without}$	15	19	20	24	19	18
35 to 44	P_{with}	34	40	41	32	57	39
	$P_{without}$	21	22	22	18	26	22
45 to 54	P_{with}	24	27	32	58	47	29
	$P_{without}$	10	12	16	32	15	13
55 and older	P_{with}	7	20	0	25	28	13
	$P_{without}$	3	9	0	5	12	5
	P_{with}	30	41	52	51	59	41
	$P_{without}$	18	24	32	33	27	24

We see that the insider-outsider effect is 66%-43%=23% for ages under 25 and falls for older women to 8% with a dip (12%) for the ages 30-34. Not unexpectedly the insider advantage rises from 12% to 32% with the education level.

6. Exercises in socio-economic policy.

In this section we consider a number of applications of the above model with an eye on labor market policy for the Dutch female population.

An interesting point is the impact of policy measures on employment and its distribution over various subgroups. First we consider four types of *net* wage changes, which affect the worker but not the employer. This is realized in the model by adding to w a second concept w^{w}. The first remains our old w which is the index of relevant wage costs for the employer. The wage w^{w} is the net wage of the worker, which is found by adding to or subtracting from w a specific amount. These measures can be enforced by manipulating the income tax or social security premiums, paid by the worker. The first policy is an increase of $\ln(w)$ by α, that is, $w^{w}=\exp(\alpha)w$ (with $\alpha=5\%$ and 10% respectively). The second one is a rise by a constant amount, which is equivalent to a fraction $\exp(\alpha)$ of the median wage (with $\alpha=$Dfl.100 and Dfl.200 per month respectively). The results of those measures, both in terms of additional persons and in terms of additional hours, are presented in table 6.1 in the first column. In a similar way we may think of fiscal measures which reduce or increase the wage cost w^{f} for the firm. Then the remuneration of the worker is kept unchanged, but the wage costs corresponding to it are changed. This policy is presented in column II of table 6.1. Thus far we assumed fiscal measures aiming at either changing the net wage to the worker or the gross wage costs to the firm. However, this implies a change in fiscal revenue which has to be born by the state. It is also conceivable to avoid this impact on the treasury by shifting the burden directly from the worker to the firm. In that case an increase of the net wage for the worker implies also an increase of wage costs to the firm. This case where w^{w} and w^{f} are increased or reduced by the same amount is given in the third column.

We see that the measures concerning the worker only have a small positive effect on employment. However a decrease of the employer's cost by 5% increases the volume of employment relatively with about 30%. The third measure where the employer pays for the wage rise of the worker decreases the volume of employment substantially. These results follow from the fact that in most cases in the sample the unemployment is caused by unwillingness of the employer, as the classes B, D and F in table 5.2 are rather large.

Table 6.1. The impact of wage policy measures on participation and hours supply.
(# is % respondents; h is % hours supply)

		I	II	III
Base	#	23.9	23.9	23.9
	h	31.2	31.2	31.2
Proportional change of 5%	#	25.2	34.7	18.7
	h	32.3	43.3	24.1
Proportional change of 10%	#	26.3	46.4	12.4
	h	33.2	54.6	16.7
Absolute change of Dfl.100 p.m.	#	25.5	36.6	17.0
	h	32.7	45.7	21.9
Absolute change of Dfl 200 p.m.	#	26.9	48.5	10.5
	h	33.9	57.2	13.3

I increase the net wage of the worker without affecting the employer's cost
II decrease the cost of the employer without affecting the worker's net wage
III increase the net wage of the worker by increasing the employer's cost

For all measures the impact on hours supply is just the same as on the number of respondents. Absolute changes in the wage rates yield similar results which are presented in the last two rows of table 6.1.

In the Netherlands there is a statutory minimum wage that should be paid to workers over 23 years old. In 1983 this was a net amount of Dfl. 8.48 per hour. For some of the respondents the predicted institutional wage rate (w) falls below this level. For the individuals who are predicted to earn less than this statutory minimum, we set the institutional wage to this statutory minimum wage rate of Dfl. 8.48. Then we find that an additional 1.9% of the women in the sample is predicted to be unemployed (see table 6.2, column I and II). In a similar way we can calculate the impact of changes in the statutory minimum wage rate. Just as in the calculations for table 6.1 we introduce an institutional wage for the worker w^w and one for the firm w^f, which we change independently, reflecting changes in the net wage of the worker and the net cost of the firm respectively. The results of proportional increases and decreases with 5% and 10% are given in table 6.2, both in terms of additional persons and in terms of additional labor supply.

Table 6.2. The impact of changes in the statutory minimum wage on participation and hours supply.
(# is % respondents; h is % hours supply)

		I	II	III	IV	V	VI
Change of 5%	#	23.9	21.8	22.0	21.0	22.4	22.2
	h	31.2	28.7	28.8	27.7	29.3	29.2
Change of 10%	#	23.9	21.8	22.2	20.4	22.9	22.5
	h	31.2	28.7	29.1	26.8	29.8	29.6

I Base
II Introduction of a statutory minimum wage of Dfl 8.48 per hour
III Proportional increase (with 5 or 10%) of the statutory minimum wage without affecting the employer's cost
IV Proportional increase (of 5 or 10%) of the statutory minimum wage by increasing the employer's cost
V Proportional decrease (of 5 or 10%) of the statutory minimum wage without affecting the worker's wage
VI Proportional decrease (of 5 or 10%) of the statutory minimum wage by decreasing the worker's wage

We see that these measures have a limited effect on the participation rate of the women in the sample; the effects all are smaller than 2%. The introduction of the minimum wage rate (column I in both rows) decreases the participation with about 2%, as the employer has to bear the cost. Then an increase of this statutory minimum with 10% increases the participation with only 0.7% if the employer's cost are unaffected (column III) while the participation decreases by another 2% if the employer pays for this increase (column IV). Reductions of the statutory minimum have the reverse effect; the size of the effects is somewhat larger (column V and VI).

It is frequently suggested that one of the main factors hampering female labor participation is not wages, but rather the lack of childcare facilities for working females (see e.g. European Commission (1988)). This can also be found from our estimates. First we find that hours supply is reduced by a factor $\exp(-0.6124)=0.54$ when there are small children. Second, we find that employers are not very fond of hiring women with children at home; w_{max} decreases with the presence of children. Our last exercise is devoted to the impact on participation and labor supply if there were childcare facilities at no cost or reduced cost, supplied by the authorities. Therefore we calculate the participation rates if D_{ch4} or D_{ch12} or both are set at zero, reflecting perfect childcare at no cost. A similar calculation is made for a measure providing better and/or cheaper childcare, which is approximated by setting the value of the dummies from one to one half. The results are given in the next table, classified according to the level of education.

Table 6.3. The impact of childcare facilities on participation and labor supply.
(# is % respondents; h is % hours supply)

		Base	D_{ch4}=½	D_{ch12}=½	both ½	D_{ch4}=0	D_{ch12}=0	both 0
E_1	#	17.7	18.9	18.2	19.3	21.2	18.8	22.3
	h	21.5	21.7	21.8	22.0	23.5	22.3	24.1
E_2	#	23.9	25.8	24.6	26.4	29.1	25.4	30.6
	h	30.5	30.8	30.6	30.9	33.3	31.1	33.7
E_3	#	32.4	35.4	32.8	35.7	40.2	33.9	41.6
	h	41.7	42.2	41.3	41.9	45.8	42.0	46.6
E_4	#	33.4	35.3	33.9	35.8	37.5	35.3	39.4
	h	43.2	43.8	42.9	43.6	43.8	43.9	44.5
E_5	#	27.3	30.9	27.7	31.2	35.5	27.9	36.2
	h	34.0	35.5	33.8	35.4	39.6	33.5	39.0
Total	#	23.9	25.9	24.5	26.4	29.2	25.2	30.5
	h	31.2	31.7	31.2	31.7	33.7	31.7	34.5

E_i (i = 1, ... ,5) are the subgroups of the sample according to the level of education, Total gives the whole sample.

This table shows that easing the childcare chores may have an effect on female employment. For the lowest level of education employment would increase from 17.7% to 20.3%. At the highest level of education it would increase from 27.3% to 36.2%. It is obvious that childcare facilities seem to be a major clue to solve Dutch female unemployment. On the average the effect is 6.6% at most; that is a relative increase of about 28%.

7. Conclusion.

In this paper we describe a model of the labor market which is richer and more realistic than the traditional models. The traditional models focus almost exclusively on the supply behavior of the worker, while in this model we introduce a symmetric approach. We combine a demand and a supply equation. Undoubtedly, our empirical model would gain if we also had characteristics of the firms involved, because labor demand depends on specific characteristics of the firms as well, which we do not know from our data set. Likewise workers prefer a specific type of firm to another one. This is still an untapped source of heterogeneity. In addition it is unrealistic to assume that w_{max} does not depend on the number of already employed or rather on the size and the composition of the labor force of the specific firm. Those effects would be estimable if we had data on the employers and their firm as well. Nevertheless, it is surprising that such a plausible model could be found, given those limitations.

With respect to the estimation technique it shows that the SEM-algorithm is a very powerful method to estimate models, which can only be perceived through the troubled glasses of incomplete or set-valued observations. The method is based on the fundamental truth that we are not interested in the exact observations individually, but only in their distribution. We construct a pseudo sample with a distribution identical to that of the exact observations, which are only observed setwise. In this example we estimated up to 46 different parameters in addition to the covariance matrix.

We also see that the method allows us to consider the econometrics of limited dependent and discrete variables as nicely embedded in traditional econometrics instead of as a separate branch consisting of ad hoc methods for specific cases, see e.g. the admirable catalogue by Maddala (1983). The estimation method is actually viable because of modern computer technology, so the model can behave as much as possible in conformity with the data. Computationally the method utilizes only traditional econometric tools, viz., estimation methods applicable for continuous data and simulation from univariate normal densities. This makes it rather simple to program without need for sophisticated software. However we need a fast computer as large data blocks have to be used during many iterations.

One of the most striking outcomes with political relevance is that the household situation of the female does not only influence her willingness to supply her labor, but that this situation is also important for the employer in determining his willingness to employ the worker. Except for its methodological and theoretical interest the results are relevant for labor market policy as well. The indirectly observed wage rates w_{min} and w_{max} can be compared to the observed market wage rates. Then it will become clear that the market wage rate is sometimes lower than w_{min} indicating that there is a risk on voluntary unemployment and that the market rate is sometimes higher than w_{max} which yields involuntary unemployment.

It stands to reason that the same analysis can be applied to male workers. However, as the great majority of male workers in population is employed, it is more difficult to estimate the underlying relations than in the case of females.

Appendix

It is instructive to look at the traditional specification where we ignore the w_{max}-equation. In that case we find table 1^A with completely different effects; table 1^B gives the results of this 3-equation model if we undo the difference transformation.

It is in our view impossible to compare the 3-equation model with the 4-equation model, as they describe different phenomena. The first attempts to explain (h, w_{min}, w) while the second explains (h, w_{min}, w, w_{max}). We observe that the equations for h and w do not differ much, but that the w_{min}-equations differ dramatically. The difference is caused by the fact that the w_{min}-concept in the 3-equations model has to serve two patrons; it has to explain voluntary *and* involuntary unemployment. In the 4-equations model the two roles are separated between w_{min} *and* w_{max}. For economic reasons we prefer the latter specification.

Table 1[A]. Estimate of the model for the partner.
(standard deviations in parentheses)
(*: significant at the 5% level)

	$\ln(h)$	$\ln(w_{min})$-$\ln(w)$	$\ln(w)$
intercept	2.3309*	-0.3026	1.4157*
	(0.3651)	(0.3767)	(0.1561)
$\ln(age-16)$	0.5310*	-0.4188*	0.4753*
	(0.1913)	(0.0691)	(0.1229)
$\ln^2(age-16)$	-0.1535*	0.0924*	-0.0658*
	(0.0350)	(0.0175)	(0.0231)
$\ln(fs)$	-0.4596*	0.0957*	–
	(0.0399)	(0.0462)	
D_{ch4}	-0.6217*	0.2515*	0.0548*
	(0.0332)	(0.0652)	(0.0227)
D_{ch12}	-0.3454*	0.1242*	-0.0987*
	(0.0304)	(0.0367)	(0.0169)
E_2	0.1402*	-0.0501*	0.0648*
	(0.0269)	(0.0164)	(0.0170)
E_3	0.1950*	-0.0766*	0.1407*
	(0.0330)	(0.0233)	(0.0221)
E_4	0.2238*	-0.0771*	0.2519*
	(0.0441)	(0.0261)	(0.0303)
E_5	0.1158*	-0.1170*	0.3119*
	(0.0309)	(0.0158)	(0.0215)
$\overline{\ln(h)}$	–	0.3972*	–
		(0.1668)	
$\overline{\ln^2(h)}$	–	-0.0557*	–
		(0.0223)	
$\ln(w_m)$	0.0718	-0.0087	–
	(0.0369)	(0.0108)	
$\ln(h_m)$	0.1962*	–	–
	(0.0599)		
Covariances	0.2654		
	0.0031	0.0303	
	-0.0576	-0.0083	0.1187

Table 1[B]. Estimate of the model for the partner.
(standard deviations in parentheses)
(*: significant at the 5% level)

	$\ln(h)$	$\ln(w_{min})$	$\ln(w)$
intercept	2.3309*	1.1131*	1.4157*
	(0.3651)	(0.4072)	(0.1561)
$\ln(age-16)$	0.5310*	0.0565	0.4753*
	(0.1913)	(0.1342)	(0.1229)
$\ln^2(age-16)$	-0.1535*	0.0266	-0.0658*
	(0.0350)	(0.0278)	(0.0231)
$\ln(fs)$	-0.4596*	0.0957*	–
	(0.0399)	(0.0462)	
D_{ch4}	-0.6217*	0.3063*	0.0548*
	(0.0332)	(0.0677)	(0.0227)
D_{ch12}	-0.3454*	0.0255	-0.0987*
	(0.0304)	(0.0401)	(0.0169)
E_2	0.1402*	0.0147	0.0648*
	(0.0269)	(0.0236)	(0.0170)
E_3	0.1950*	0.0641*	0.1407*
	(0.0330)	(0.0309)	(0.0221)
E_4	0.2238*	0.1748*	0.2519*
	(0.0441)	(0.0402)	(0.0303)
E_5	0.1158*	0.1949*	0.3119*
	(0.0309)	(0.0262)	(0.0215)
$\overline{\ln(h)}$	–	0.3972*	–
		(0.1668)	
$\overline{\ln^2(h)}$	–	-0.0557*	–
		(0.0223)	
$\ln(w_m)$	0.0718	-0.0087	–
	(0.0369)	(0.0108)	
$\ln(h_m)$	0.1962*	–	–
	(0.0599)		
Covariances/	0.2654	-0.2917	-0.3245
Correlations	-0.0545	0.1315	0.8764
	-0.0576	0.1095	0.1187

Covariances in the lower triangle, correlations in the upper triangle

REFERENCES:

Amemiya, T. (1985) *Advanced Econometrics*, Harvard University Press, Cambridge (Mass).

Becker, G.S. (1965) "A Theory of the Allocation of Time", *The Economic Journal*, 493-517.

Billingsley, P. (1986) *Probability and Measure*, ***second edition***, Wiley & Sons, New York.

Blundell, R.J. Ham and C. Meghir (1987) "Unemployment and Female Labour Supply", *The Economic Journal*, 44-64.

Borjas, G.J. (1980) "The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias", *Journal of Human Resources*, 409-423.

Cogan, J. (1981) "Fixed Costs and Labor Supply", *Econometrica*, 945-964.

Dempster, A.P., N.M. Laird and D.B. Rubin (1977) "Maximum Likelihood via the EM Algorithm", *Journal of the Royal Statistical Society*, *B-Series*, 1-38.

European Commission (1988) "Childcare and the Equality of Opportunity", Consolidated Report to the European Commission.

Gouriéroux, C., A. Monfort, E. Renault and A. Trognon (1987) "Simulated Residuals", *Journal of Econometrics*, 201-252.

Gouriéroux, C., and A. Monfort (1989) "Simulation Based Inference in Models with Heterogeneity", mimeo, report 8905, INSEE, Paris.

Hajivassiliou, V.A. and D. McFadden (1987) "External Debt Crises: Estimation by the Method of Simulated Moments", mimeo.

Hausman, J.A. (1985) "The Econometrics of Nonlinear Budget Sets", *Econometrica*, 1255-1282.

Heckman, J.J. (1974) "Shadow Prices, Market Wages, and Labor Supply", *Econometrica*, 679-694.

Heckman, J. and T.E. MaCurdy (1986) "Labor Econometrics", in Z. Griliches and M.D. Intrilligator (eds.), *Handbook of Econometrics*, Volume III, North-Holland, Amsterdam.

Homan, M.E. (1988) *The Allocation of Time and Money in One-earner and Two-earner Families*; *An Economic Analysis*, Erasmus University Rotterdam.

Keane, M.P. (1989) "A Computationally Practical Simulation Estimator for Panel Data, with Applications to Labor Supply and Real Wage Movement over the Business Cycle", mimeo, May 1989, University of Minnesota.

Killingsworth, M.R. and J. Heckman (1986) "Female Labor Supply: A Survey", in O. Ashenfelter and R. Layard, (eds.), *Handbook of Labor Economics*, Volume I, North-Holland, Amsterdam.

Kullback, S. (1959) *Information Theory and Statistics*, Wiley & Sons, New York.

Laroque, G. and B. Salanié (1989) "Estimation of Multi-Market Fix-Price Models: An Application of Pseudo Maximum Likelihood Methods", *Econometrica*, 831-860.

Lindbeck, A. and D.J. Snower (1988) *The Insider-Outsider Theory of Employment and Unemployment*, MIT Press, Cambridge (Mass.).

Maddala, G.S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press., Cambridge (Mass.).

McFadden, D. (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration", *Econometrica*, 995-1026.

Mroz, Th.A. (1987) "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions", *Econometrica*, 749-764.

Pakes, A. and D. Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, 1027-1058.

Pencavel, J. (1986) "Labor Supply of Men: A Survey", in O. Ashenfelter and R. Layard, (eds.), *Handbook of Labor Economics*, Volume I, North-Holland, Amsterdam.

Renes, G. and A.J.M. Hagenaars (1989) "The perceived distribution of Job Characteristics", *The European Economic Review*, 845-862.

Ruud, P.A. (1991) "Extensions of estimation methods using the EM algorithm", *Journal of Econometrics*, 305-341.

Simon, H.A. (1979) *Models of Thought*, Yale University Press, New Haven.

Van Duin, E.P. and A.J.M. Hagenaars (1984) "Representativity of a Newspaper Survey", (in Dutch), Report 84.23, COEPS, Leyden University.

Van Praag, B.M.S. (1991) "Ordinal and Cardinal Utility: An Integration of the Two Dimensions of the Welfare Concept", *Journal of Econometrics*, (forthcoming).

Van Praag, B.M.S. and J.P. Hop (1987) "Estimation of Continuous Models on the Basis of Set-Valued Observations", paper presented at the Econometric Society European Meeting, Copenhagen 1987.

Van Soest, A. and A. Kapteyn (1989) "The Impact of Minimum Wage Regulations on Employment and the Wage Rate Distribution", CentER Discussion paper 8947, Tilburg University.

Wilks, S.S. (1962) *Mathematical Statistics*, Wiley & Sons, New York.