ntER
for
omic Research

Discussion paper

[redacted]

TILBURG UNIVERSITY

CentER
for
Economic Research

No. 9351

ESTIMATING AND INTERPRETING MODELS WITH ENDOGENOUS TREATMENT EFFECTS: THE RELATIONSHIP BETWEEN COMPETING ESTIMATORS OF THE UNION IMPACT ON WAGES

by Francis Vella and Marno Verbeek

August 1993

ISSN 0924-7815

Estimating and Interpreting Models with Endogenous Treatment Effects: The Relationship Between Competing Estimators of the Union Impact On Wages

Francis Vella
Department of Economics
Rice University
Houston, Texas
and

Marno Verbeek
Department of Econometrics
Tilburg University
Tilburg, The Netherlands

November 1992
This version July 1993

Abstract

This paper examines the estimation and interpretation of models with endogenous treatment effects in labor economics. We argue that the two conventional methods of estimation, control functions and instrumental variables, are closely related and thus many of the issues raised with respect to the control function procedure are equally relevant to instrumental variables estimation. We also argue that the economic significance of arbitrary econometric specifications in these models is much greater than currently acknowledged. We also show that the severe restrictions regarding sorting behavior imposed by instrumental restrictions can be easily relaxed in estimation. Finally, we show that the various tests of endogeneity associated with these estimators are based on the same population moment.

We are grateful to Jim Brown, Whitney Newey, Trevor Breusch, Adrian Pagan and Bertrand Melenberg for helpful discussions. This paper was partially written while Vella was a visitor in the Department of Statistics, The Faculties and the Department of Economics, Research School of Social Sciences, Australian National University. We alone are responsible for any remaining errors.

1. Introduction

Empirical implementations of the Roy (1951) sorting model have proliferated the labor economics literature since the seminal work of Heckman (1979) on estimating models with sample selection bias. Recently this research has extended to the identification of "treatment effects" although many of the estimation issues remain unchanged (see, for example, Heckman and Robb 1986 and Heckman and Hotz 1989). A major feature of this literature is the ability to draw insightful economic inferences regarding sorting behavior from the estimated parameters of the model.

This paper addresses several issues relating to the estimation and interpretation of these models. First, we argue that the two popular methods of estimating these models, instrumental variables and control function estimators, are more closely linked than commonly treated. This is an important issue as many of the criticisms of the control function procedures are not levied at the instrumental variable approach although they are equally relevant. With respect to interpretation we claim that despite the instrumental variable estimator's weaker distributional assumptions, it imposes several restrictive constraints on the nature of the implicit sorting behavior. Furthermore, we argue that many of the arbitrary exclusion restrictions employed in these procedures for identification purposes have more economic significance than acknowledged. Finally, we show that the tests of selection bias and union endogeneity are based on different sample estimates of the same population moment.

Throughout the paper we employ the example of the union impact on wages as a vehicle of illustration although many of our points have wider applicability. We do this for two reasons. First, several papers (see Robinson 1989a for a recent survey), have examined the union impact on wages via these estimation procedures. Second, despite these papers it appears that many of the important implicit estimation and economic issues remain misunderstood. The following section describes the model and the estimators. In section 3 we outline the relationship between the control function and instrumental variables estimators. Section 4 discusses the economics implicit in the estimation procedures. We also propose a simple test for the validity of the instrumental variables estimator based on instrumental variables estimation. In section 5 we outline the relationship of the exogeneity tests. Section 6 concludes. Throughout the paper we assume the stochastic components of the model are normally distributed. Our primary aim is to show the relationship between two "competing" estimators and discuss the implicit economic issues rather than examine the relative merits of competing estimation procedures under different parametric circumstances.

2. A Simultaneous Model of Wages and Union Membership

It is common practice to account for the endogeneity of union membership when evaluating the impact of union membership on wages. A recent paper, Robinson (1989a), provides an excellent survey of this literature and a coherent discussion of the relationship between the estimators and the economics of union membership. We argue that the two conventional ways of estimating the union impact from cross sectional data, (instrumental variables (IV) and control functions (CF)), estimate the same measure of union impact although they are constructed to produce different estimates. Consider the following simple model of union membership and wage determination proposed by Robinson (1989a, 1989b) and others

(1) $W_{Ui}=X_i\gamma_U+\epsilon_{Ui}$ $\qquad i=1,\ldots,n$

(2) $W_{Ni}=X_i\gamma_N+\epsilon_{Ni}$ $\qquad i=1,\ldots,n$

where W_i is the log of individual i's wage; X_i is a vector of exogenous variables; γ_U and γ_N are parameters to be estimated; ϵ_{Ui} and ϵ_{Ni} are random disturbances drawn from a normal distribution with non zero covariance; and the U and N subscripts denote union and non-union sector, respectively. Following Robinson, assume that the union effect operates through different intercept terms and combine equations (1) and (2) in the following manner

(3) $W_i=\gamma_N+(\gamma_U-\gamma_N)U_i+X_i\gamma+V_i$ $\qquad i=1,\ldots,n$

(4) $V_i=\epsilon_{Ni}+(\epsilon_{Ui}-\epsilon_{Ni})U_i$ $\qquad i=1,\ldots,n$

where U_i is an indicator dummy denoting union membership. Also introduce equations capturing the cost of union membership and the index determining union membership

(5) $C_i=S_i\delta+\epsilon_{Ci}$ $\qquad i=1,\ldots,n$

(6) $I_i=W_{Ui}-W_{Ni}-C_i\ =Z_i\pi+\epsilon_i;\ \epsilon_i=\epsilon_{Ui}-\epsilon_{Ni}-\epsilon_{Ci}$ $\qquad i=1,\ldots,n$

where C_i can be interpreted as the cost of union membership; S_i is a vector of exogenous variables; I_i is a latent variable reflecting gain from union membership; $U_i=1$ if $I_i>0$ and $U_i=0$ otherwise; δ and π are vectors of parameters; and ϵ_{Ci} is a normally distributed error term potentially correlated with ϵ_{Ui} and ϵ_{Ni}. Z_i is the vector of *all* exogenous variables in the system.

The OLS estimate of $\gamma_P=(\gamma_U-\gamma_N)$ from (3) is biased if U_i and V_i, conditional on X_i, are correlated. A consistent estimate of γ_P is the generalized IV estimator $\hat{\gamma}_{PGIV}$ obtained by OLS from

$$(8) \quad W_i=\gamma_N + \hat{U}_i\gamma_P + X_i\gamma + \eta_i$$

where $\hat{U}_i$ is the linear prediction[1] of U_i given Z_i and η_i is a zero mean error term. γ_P has the interpretation of the return to union membership to an individual who is *randomly* assigned to a union. Identification of γ_P and estimation of (8) requires at least one regressor in Z_i which is not contained in X_i. An alternative consistent estimate is obtained through the control function methods of Heckman (1979) and Lee (1978) which exploit the assumption of joint normality of the error terms. The conditional expectation[2] of an individual's wage is

$$(9) \quad E[W_i|U_i]=\gamma_N+ X_i\gamma + U_i\gamma_P + \lambda_iU_i\theta_{U\epsilon} + \lambda_i(1-U_i)\theta_{N\epsilon}$$

where $\theta_{U\epsilon}$ is equal to $\sigma_{U\epsilon}/\sigma^2_\epsilon$ and $\theta_{N\epsilon}$ is equal to $\sigma_{N\epsilon}/\sigma^2_\epsilon$ where $\sigma_{N\epsilon}$ and $\sigma_{U\epsilon}$ are the covariances between ϵ_i and ϵ_{Ni}, and ϵ_i and ϵ_{Ui} respectively, and σ^2_ϵ is the variance of ϵ_i; and $\lambda_i=E[\epsilon_i|U_i]$. To estimate (9), we require an estimate of λ_i and this is given by the inverse mills ratio from (6). Note that λ_i is also known as the generalized residual for the probit model (see Gourieroux et.al 1987 and Vella 1993), given by

$$(10) \quad E[\epsilon_i|U_i]=\lambda_i=[1-U_i]\{-\phi(-Z_i\pi)/\Phi(-Z_i\pi)]\}+U_i[\phi(-Z_i\pi)/\{1-\Phi(-Z_i\pi)\}]$$

where $\phi(.)$ and $\Phi(.)$ represent the probability density and cumulative distribution functions of the standard normal distribution evaluated at the values inside the parentheses[3]. The estimable form of (9) can be written as

---

[1] This instrumental variables estimator is employed for simplicity. A more general IV estimator is discussed in Heckman and Robb (1986), where X_i and a known transformation of a variable in S_i are used as instruments.

[2] The conditional expectations in the sequel are also conditional upon Z_i.

[3] The log likelihood function for the union equation in (6) is

$$L = \sum_{i=1}^{n} [\ U_i\log \Phi(Z_i\pi)+ (1-U_i) \log \{1-\Phi(Z_i\pi)\}\ ]$$

and the estimates for π are defined by the first order conditions

$$(11)\quad W_i=\gamma_N + U_i\gamma_P + X_i\gamma + \lambda_i U_i\theta_{U\epsilon}+\lambda_i(1-U_i)\theta_{N\epsilon} + \nu_i$$

where ν is a zero mean error term uncorrelated with the regressors. To implement this procedure we estimate (6) over the n observations by probit maximum likelihood to obtain $\tilde{\pi}$, generate (10) at the probit estimates and insert the generated values into (11) and estimate by least squares. Unlike the instrumental variables estimator it is not necessary to include something in Z_i which is not in X_i as the mapping from Z_i to λ_i is non-linear. γ_P reflects the return to random assignment to union membership. The θ's, combined with λ_i, capture the individual random effects which vary across individual.

3. The Relationship Between Instrumental Variables and Control Function Estimators

IV and CF procedures generally produce estimates of the union differential that are comparable (see Robinson 1989a). However, under certain parametric assumptions the CF and IV estimates are identical by construction. Consider the intuition behind this result. It is well known that the endogeneity of regressors can be accounted for by replacing the endogenous regressor in the primary equation with its predicted value from an auxiliary regression, or including both the endogenous variable and the residual from this auxiliary regression. As the inverse mills ratio is the "residual" from the probit model the link between the control function and instrumental variable methodologies is immediately apparent.

However as this result requires the residuals to be a linear function of the regressors this exact relationship does not hold for λ_i. To show the exact relationship consider a variant of the Heckman two step approach which is referred to as a linear control function. Substitute the OLS residuals from regressing U_i on Z_i, $\hat{\epsilon}_i=U_i-\hat{U}_i$, in place of $\hat{\lambda}_i$.[4] This gives

---

$$\frac{\partial L}{\partial \pi} = \sum_{i=1}^{n}[\ U_iZ_i'\{\ \phi(-Z_i\pi)/(1-\Phi(-Z_i\pi))\}+(1-U_i)Z_i'\{-\phi(-Z_i\pi)/(\Phi(-Z_i\pi))\ \}\ ] = 0.$$

Using the definition of Gourieroux et. al (1987) that the generalized residual for the probit model is the derivative of the log likelihood function with respect to the intercept evaluated at the maximum likelihood estimates it is straightforward to establish that the inverse mills ratio is the generalized residual for the probit model.

[4] This procedure is suggested by Olsen (1980) which illustrates that under certain parametric assumptions this procedure generates consistent estimates of γ_P. Heckman and Hotz (1989) refer to this procedure as a linear control function as the first

$$(12) \quad W_i = \gamma_N + U_i\gamma_P + X_i\gamma + (U_i - \hat{U}_i)\theta + \eta_i$$

and where OLS on (12) reproduces $\hat{\gamma}_{PGIV}$ as an estimate of γ_P. Now substitute the least squares residual with the inverse mills ratio[5]

$$(13) \quad W_i = \gamma_N + U_i\gamma_P + X_i\gamma + \hat{\lambda}_i\theta + \nu_i.$$

Estimating (13) by instrumental variables eliminates $\hat{\lambda}_i$, since $\sum_{i=1}^{n} \hat{\lambda}_i Z_i = 0$ from the first order conditions of the probit model, and re-produces the instrumental variables estimator $\hat{\gamma}_{PGIV}$. Estimating (13) by OLS however does not eliminate $\hat{\lambda}_i$ and produces the estimate $\hat{\gamma}_{PCF}$. This estimate is more efficient than $\hat{\gamma}_{PGIV}$ due to the normality assumption, and the non-linear mapping from Z_i to $\hat{\lambda}_i$, employed in the CF estimate but ignored in the IV (see appendix). It is obvious from (12) and (13) that the IV and CF procedures produce identical estimates if $\hat{\lambda}_i = \hat{\epsilon}_i$. This condition, however, will never be satisfied under normality. However, for many applications the empirical correlation coefficient between $\hat{\lambda}_i$ and $\hat{\epsilon}_i$ will be very high[6] and this explains the similarity obtained from the IV and CF approaches (see appendix). This similarity will also hold when the true underlying distributions are not normal.

It is useful to consider the "economic" implications of the similarity between the CF and IV procedures. First, the IV requirement for at least one additional regressor in Z_i indicates that while the CF approach can be internally consistent with individuals sorting **only** on the basis of wages the IV approach cannot. That is, if costs and tastes are truly irrelevant then only the variables which appear in X_i should appear in Z_i. This implies that the IV estimator would not be identified. Second, while a tremendous deal of scrutiny is paid to exclusion restrictions in the first step probit, Z_i is typically specified in an arbitrary manner when employing IV. However, the similarity in the two methods indicates that Z_i should be chosen equally carefully whether CF or IV is employed. In practice, however, IV is often preferred over CF on the basis that it is easier to specify the instruments and there is less concern about economic justification

step is a linear probability model. Note that in this case it is necessary that some regressor wich appears in Z_i is not in X_i.

[5] We set $\theta_{U\epsilon} = \theta_{N\epsilon}$ in (13) for simplicity. We discuss the implication of this restriction below.

[6] In our experience, empirical correlation coefficients using real life data are usually in the range 0.99 - 0.9999.

for inclusion in Z_i. It is clear that both methods are equally sensitive to the choice of Z_i.

Although IV and CF both provide consistent estimates in this model it is generally true that IV in "typical" selectivity circumstances does not. The conventional selection bias model has wages only observed for the subset corresponding to $U_i=1$ or $U_i=0$ although the vector Z_i and U_i is observed for everyone. Suppose wages are only observed for the union members. Thus the model to be estimated is

$$\text{(1a)}\quad W_{Ui} = X_i\gamma_U + \epsilon_{iU} \qquad i=1\ldots n_1$$
$$\text{(6)}\quad I_i = Z_i\pi + \epsilon_i \qquad i=1\ldots n$$

where $n_1 < n$. The control function procedure produces consistent estimates of γ_U from OLS over the sample of union members after including the inverse mills ratio in (1a). However the instrumental variables estimator is no longer unbiased or consistent for γ_U. The reason is that Z_i is no longer uncorrelated with λ_i for non-random subsamples. The bias, see appendix, is directly related to $\sum_{i=1}^{n_1} \lambda_i Z_i$. By noting that $\sum_{i=1}^{n_1} \lambda_i Z_i$ are precisely the first order conditions from the probit model for union member observations (see footnote 3) it follows that the bias of the IV estimate of γ_U is a weighted function of the "distance" by which the first order conditions of the probit model, explaining union membership, are violated over the sub-sample. Accordingly, this term will be zero when evaluated over the entire n observations. However, in general they will be non zero for non-random sub samples.

4. The Economics of Union Membership

Non zero values for $\theta_{U\epsilon}$ and $\theta_{N\epsilon}$ indicate that the same unobserved influences upon an individuals union status also have an impact on that individuals wage. The θ's, and λ_i, also capture the economics of union membership. First consider the case where individuals sort solely on wages, i.e. costs are irrelevant. When $\sigma_{UN}<0$, the case of a comparative advantage structure, positive selection will produce a positive $\theta_{U\epsilon}$ and a negative $\theta_{N\epsilon}$. A hierarchical structure, characterized by $\sigma_{UN}>0$, suggests the better workers, as measured by their endowments of ϵ_U and ϵ_N, will locate in the higher variance sector and positive selection produces a positive value for at least one of the θ's[7]. These implications are far less clear when σ_{NC} and/or σ_{UC} are non zero and this is

[7] Both θ's will be positive when $\sigma_U^2>\sigma_{UN}>\sigma_N^2$.

discussed below.

An important point raised by Robinson (1989a, 1989b) is that the restriction required for IV estimation, shown in (14) below, rules out the possibility of positive (or negative) selection in both markets. While the assumption is restrictive it is incorrect to conclude that it does not allow positive sorting. The conditional expectation of the error term V_i from (3) has the following form

$$(14)\quad E[V_i] = E[\epsilon_{Ni}|U_i=0]Pr(U_i=0) + E[\epsilon_{Ui}|U_i=1]Pr[U_i=1] = 0.$$

Robinson argues that as both probabilities are strictly non-negative this imposes opposite signs on the conditional error means and thus rules out positive (or negative) selection in both markets. The first part of this statement is true and is confirmed by equation (10). The error terms for the respective sectors will have opposite signs by construction. However, the sorting process does not operate purely through the signs of ϵ_{Ui} and ϵ_{Ni}. As is shown in equation (9) the sorting phenomena is also captured by $\theta_{U\epsilon}$ and $\theta_{N\epsilon}$[8]. The statement by Robinson is true if $\sigma_{UN}<0$ and $\sigma_{UC}=\sigma_{NC}=0$, that is, if costs are irrelevant in the sorting decision and the skills in the two sectors are negatively correlated. However, in general the statement is not true. Following Robinson rewrite equation (14) as

$$(15)\quad E[V_i] = \theta_{N\epsilon}\frac{-\phi(-Z_i\pi)}{\Phi(-Z_i\pi)}\Phi(-Z_i\pi) + \theta_{U\epsilon}\frac{\phi(-Z_i\pi)}{1-\Phi(-Z_i\pi)}(1-\Phi(-Z_i\pi)) = 0$$

and simplify to get

$$(16)\quad E[V_i] = (-\theta_{N\epsilon}+\theta_{U\epsilon})\ \phi(-Z_i\pi)=0.$$

Condition (16) is necessary for IV to be consistent. That is, $\theta_{N\epsilon}=\theta_{U\epsilon}$ implying $\sigma_U^2-\sigma_{UN}-\sigma_{UC}=\sigma_{UN}-\sigma_N^2-\sigma_{NC}$. Consider where costs are irrelevant implying $\sigma_U^2+\sigma_N^2-2\sigma_{UN}=0$. For this to be satisfied a necessary condition is $\sigma_{UN}>0$. That is, the error structure must have a hierarchical structure implying the better union workers would also make the better non-union workers[9]. This imposes positive estimates for both coefficients. More

---

[8] It is possible to infer the sign of the θ's from ϵ_{Ui} and ϵ_{Ni} whenever the sign of ϵ_i is unique to the sector. However, while this is satisfied in the 2 sector model it clearly cannot be for 3 or more sectors.

[9] Robinson (1989a) refers to Robinson and Tomes (1984), where the coefficients on the

importantly, it also implies that ϵ_{Ui} and ϵ_{Ni} are perfectly correlated. This is much stronger than what is required for hierarchical sorting. It indicates that the performance in one sector is a sufficient statistic for performance in the other. It is useful to note that the CF procedure also imposes this constraint when the θ's are set equal in estimation.

The restrictive nature of IV is due to the implicit restriction it imposes. While the CF approach generally estimates both θ's separately IV implicitly imposes their equality. Robinson suggests one solution to this is to allow the value in (14) to equal some constant k rather than zero. This appears to be a satisfactory solution only if we are interested in γ_P. An alternative approach is the following. Consider equation (17) where we do not restrict $\theta_{U\epsilon}$ and $\theta_{N\epsilon}$ to be equal. This gives

$$\text{(17)} \quad W_i = \gamma_N + U_i\gamma_{P*} + X_i\gamma + \hat{\epsilon}_i\theta^* + (1\text{-}U)\hat{\epsilon}_i\theta^*_{N\epsilon} + \nu$$

where θ^* is the coefficient for the common random effect and $\theta^*_{N\epsilon}$ is the coefficient for the additional random effect for union sector. It is straightforward to show that the estimate of $\theta^*_{N\epsilon}$ is obtained from the regression

$$\text{(18)} \quad W_i = \gamma_N + \hat{U}_i\gamma_P - (1\text{-}U_i)\hat{U}_i\theta^*_{N\epsilon} + \hat{\epsilon}_i\theta^* + \nu$$

where a component of the random effect is captured by $\theta^*_{N\epsilon}$. That is, the union effect for a union worker can differ from that of a non-union worker. γ_P continues to represent the return to random assignment to union membership while holding the random effects constant. More importantly, the t-test on $\theta^*_{N\epsilon}$ is a test of the equality of the $\theta_{U\epsilon}$ and $\theta_{N\epsilon}$ and is thus a test of the instrumental variables approach. This test is more convenient than that proposed by Robinson (1989a) as it is conducted within the IV framework[10].

Now consider where costs are relevant in the sorting decision and are correlated with either and/or both union and non-union productivity. Consistency of the instrumental

---

selection terms in the respective union and non-union samples are both positive, as evidence that the IV estimator is inappropriate. However, the IV estimator requires both coefficients to possess the same sign which is inconsistent with positive sorting in the comparative advantage structure. The positive coefficients reported in Robinson and Tomes do not refute the use of the IV estimator as suggested by Robinson.

[10] Robinson proposes testing the IV restriction by estimating the θ's separately via the CF framework and testing their equality.

variables estimator requires, after minor rearrangement, $\sigma_U^2+\sigma_N^2-2\sigma_{UN}+\sigma_{NC}-\sigma_{UC}=0$. When $\sigma_{UN}<0$ a necessary condition for (16) is $\sigma_{UC}>\sigma_{NC}$. This is similar to a condition discussed in Robinson (1989b). However note that the difference in these later covariances must dominate the sum of the variances and covariances of the wage errors. In economic terms this appears to be a condition which is very unlikely to be satisfied. Thus the comparative advantage model appears to be inconsistent with the IV approach. Note, however, that as $\sigma_U^2+\sigma_N^2-2\sigma_{UN}$ is strictly non-negative even the hierarchical structure model requires $\sigma_{UC}>\sigma_{NC}$.

It is useful to extend this discussion of interpretation of the θ's to the control function approach which provides unrestricted estimates of the respective θ's. When the costs are irrelevant to the model or are uncorrelated with the respective productivity in the two sectors it is possible to assign the interpretation of various types of sorting to the model. However, if we recall that $\sigma_{U\epsilon}=\sigma_U^2-\sigma_{UN}-\sigma_{UC}$ and $\sigma_{N\epsilon}=\sigma_{UN}-\sigma_N^2-\sigma_{NC}$ it is immediately apparent that unless the σ_{UC} and σ_{NC} are non consequential the signs of the θ's say nothing about the sorting in the model as any pair of signs is possible and consistent with any type of sorting. It is important to note that it is common practice to specify the elements of Z_i such that it includes at least one, if not several, variables not found in X_i. This implies that the role of non-wage considerations is important in determining union membership and that the signs of the θ's say nothing about sorting.

While this later point may appear obvious we feel it is one which has been overlooked in this literature. Consistent with the discussion related to exclusion restrictions the interpretation of the θ's must be in accord with the variables included in Z_i. Attempts to identify the model through exclusion restrictions have subsequent implications for economic interpretation.

5. Exogeneity Tests

One final issue is the testing of the endogeneity of union membership to wages. It is well known that such a test is based on the parameters $\theta_{U\epsilon}$ and $\theta_{N\epsilon}$ and, accordingly, is a t-test on the coefficient for the inverse mills ratio. However, what is less well known is this is a Hausman test. To show the relationship between these tests we employ the testing strategy of Newey (1985) and examine the implied conditional moments being tested[11].

---

[11] To employ the results from Newey (1985) it is necessary to assume that all estimation is performed by maximum likelihood. This does not change any of the

Return to our original model and, for simplicity, set $\theta_{U\epsilon}=\theta_{N\epsilon}$. The Hausman test is based on testing whether $\hat{U}_i$ has a statistically significant coefficient when included in equation (3). The sample estimate of this is $\tau_{GIV}= n^{-1}\Sigma\ \hat{V}_i\hat{U}_i$ where $\hat{V}_i$ is the least squares residual from (3). However, it is well known that an equivalent test is based on $\tau_{GIV}=n^{-1}\Sigma\ \hat{V}_i\hat{\epsilon}_i$. That is, the Hausman test can be performed by including either the predicted value of the endogenous regressor or the residuals from the reduced form as an additional regressor in the primary equation. It is valuable to note that the corresponding population moment is the covariance between the errors from the two equations and this can be written as their product. That is, $\tau_{Hausman}= E(V_i\epsilon_i)$.

A test of selection bias in the CF framework is also a test of a non-zero covariance between V_i and ϵ_i and is also given by the conditional moment $E[\epsilon_i V_i]=0$. However, given that the union equation has a censored dependent variable we condition on the form of the censoring and the moment of interest becomes $\tau_{Heckman}= E\{E(V_i|W_i)*E(\epsilon_i|U_i)\}$ where we generate the error on the basis of the observability of the dependent variable. The sample estimate of this quantity is given by the product of the sample estimates of the errors (see Melino 1982, Pagan and Vella 1989, and Vella 1992). This is given by $\tau_{CF}= n^{-1}\Sigma\ \hat{V}_i\hat{\lambda}_i$ recalling that $E(\epsilon_i|U_i)=\lambda_i$. As $\hat{\lambda}_i$ and $\hat{\epsilon}_i$ are highly correlated the tests should perform similarly when distributional assumptions are satisfied. What is of more interest, however, is that the tests are based on the same population moment although the sample estimates differ due to the censoring.

6. Conclusion

This paper examines several issues in the estimation and interpretation of models with endogenous treatment effects. We argue that the two conventional methods of estimating these models, control functions and instrumental variables, are more closely related than commonly treated. We argue that many of the issues raised with respect to the specification of the control function procedure are equally relevant to instrumental variables estimation. We also find that the instrumental variables procedure implicitly imposes restrictions on the nature of the sorting in the models. We find, however, that these restrictions can be easily relaxed through the estimation of an additional parameter within the instrumental variables framework. Finally we show that the various tests of endogeneity associated with these estimators are based on the same population moment.

---

relationships that follow although it would be not possible to compute the test in the conditional moment framework unless maximum likelihood is employed.

References

Duncan, G., and D. Leigh "The Endogeneity of Union Status; An Empirical Test" Journal of Labor Economics 3 (1985), 385-402.

Gourieroux, C., A. Monfort , E. Renault and A. Trognon, "Generalized Residuals," Journal of Econometrics 34 (1987), 5-32.

Hausman, J., "Specification Tests in Econometrics," Econometrica 46 (1978), 1251-1271.

Heckman, J.,"Sample Selection Bias as a Specification Error," Econometrica 47 (1979), 153-161.

Heckman, J., and R. Robb, "Alternative Identifying Assumptions in Econometric Models of Selection Bias," in Advances in Econometrics: Innovations in Quantitative Economics, Essays in Honor of Robert L. Basmann (Volume 5) ed D. Slottje, Greenwich CT: JAI Press (1986) 243-287.

Heckman, J., and V. Joseph Hotz, "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," Journal of the American Statistical Association 84 (1989), 862-874.

Lee, L., "Unionism and Wage Rates: A Simultaneous Equation Model with Qualitative and Limited Dependent Variables," International Economic Review 19 (1978), 415-433.

Melino, A., "Testing for Sample Selection Bias," Review of Economic Studies 49 (1982) 151-153.

Newey, W., "Maximum Likelihood Specification Testing and Conditional Moment Tests," Econometrica 53 (1985), 1047-1070.

Pagan, A., and F. Vella, "Diagnostic Tests for Models Based on Individual Data: A Survey," Journal of Applied Econometrics 4 (1989), S29-S60.

Olsen, R.,"A Least Squares Correction for Selectivity Bias", Econometrica 48, (1980) 1815-1820.

Robinson, C., "The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models," Journal of Political Economy 97, (1989a), 639-667.

Robinson, C., "Union Endogeneity and Self-Selection," Journal of Labor Economics 7, (1989b) 106-112.

Robinson, C., and N. Tomes (1984) "Union Wage Differentials in the Public and Private Sectors: A Simultaneous Equations Approach," Journal of Labor Economics 2, (1984) 106-127.

Roy, A., "Some Thoughts on the Distribution of Earnings," Oxford Economic Papers 3, (1951) 135-146.

Vella, F., "Simple Tests for Sample Selection Bias in Censored and Discrete Choice Models," Journal of Applied Econometrics 7, (1992) 413-421.

Vella, F., "A Simple Estimator for Simultaneous Equation Models with Censored Endogenous Regressors" International Economic Review 34, (1993), 441-457.

Appendix:

In matrix form the model is given as

(A1) $W=U\gamma_P+X\gamma+V$

(A2) $I=Z\pi+\epsilon$

Let $P_z=Z(Z'Z)^{-1}Z'$, where Z is the nxk matrix of exogenous variables (including an intercept term).

Proposition 1: The OLS estimator of (13) is more efficient than the instrumental variable estimator. In matrix form equation (13) has the form

(A3) $W= U\gamma_P + X\gamma + \hat{\lambda}\theta + \nu.$

Decompose (A3) into two orthogonal regressions by premultiplying by P_z and its orthogonal complement $M_z = I - P_z$. This gives two equations containing the same information as (A3).

(A4) $P_zW = X\gamma + P_zU\gamma_P + P_z\lambda\theta + P_z\eta = X\gamma + P_zU\gamma_P + P_z\eta.$

(A5) $M_zW = M_zU\gamma_P + M_Z\lambda\theta + M_z\eta.$

OLS on (A3) produces the control function estimator $\hat{\gamma}_{PCF}$, while OLS on (A4) reproduces the IV estimator $\hat{\gamma}_{PGIV}$. If λ is not proportional to M_zU the OLS estimate of γ_P from (A5) will differ from γ_{PGIV}. Consequently, if the assumption of normality is correct, (A5) will contain additional information on γ_P improving the efficiency of the CF estimator relative to the IV.

Proposition 2: The exact relationship between the estimators, for the model shown in equations (A1) and (A2), can be obtained using straightforward algebra from the following results

(A6) $\hat{\gamma}_{PGIV}= (U'M_XU - U'M_ZU)^{-1}(U'M_XW - U'M_ZW)$

(A7) $\hat{\gamma}_{PCF}= (U'M_XU - U'P_\lambda U)^{-1}(U'M_XW - U'P_\lambda W).$

where $P_X=X(X'X)^{-1}X'$, $M_X = I - P_X$, and X denotes the exogenous variables in the wage equation and recalling that X is a subset of Z. P_λ denotes the projection onto λ. If

$P_\lambda U = M_Z U$ both estimators are identical. As $P_\lambda U = P_\lambda M_Z U = \kappa\lambda$ for some constant κ, the estimators are different to the extent that $\kappa\lambda$ differs from $M_Z U$. As the inverse Mill's ratio is highly correlated with the least squares residual, the similarity of the estimates follows.

Proposition 3: The bias of the IV estimator in the conventional selectivity model is proportional to the distances by which the first order conditions for the accompanying selection equation are violated. Suppose wages are only observed for the union members. Thus the model to be estimated is

$$\text{(A8)} \quad W_U = X\gamma_U + \epsilon_U \qquad i=1\ldots n_1$$

$$\text{(A9)} \quad I = Z\pi + \epsilon \qquad i=1\ldots n$$

The GIV estimate of γ_U from (A8) is given by $(X'P_zX)^{-1}X'P_zW$. Thus $E[\gamma_U|U=1] = (X'P_zX)^{-1}X'P_zE[W_U|U=1]$, where the condition $U=1$ reflects we only include observations for which $U_i=1$. Substitution gives $E[\gamma_U|U=1]=\gamma_U+(X'P_zX)^{-1}X'P_zE[\epsilon_U|U=1]$ noting that $E[\epsilon_U|U=1]= \lambda$ where λ is the vector of inverse mills ratios for the union members with elements $\phi(-Z_i\pi)/(1-\Phi(-Z_i\pi))$. Thus the bias of the IV estimate of γ_U is

$$\text{(A10)} \quad (X'P_zX)^{-1}X'Z(Z'Z)^{-1}Z'\lambda$$

noting that the terms $Z'\lambda$ are precisely the first order conditions from the probit model implied by (6) evaluated for the union member observations.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9215	W. Güth and K. Ritzberger	On Durable Goods Monopolies and the (Anti-) Coase-Conjecture
9216	A. Simonovits	Indexation of Pensions in Hungary: A Simple Cohort Model
9217	J.-L. Ferreira, I. Gilboa and M. Maschler	Credible Equilibria in Games with Utilities Changing During the Play
9218	P. Borm, H. Keiding, R. Mclean, S. Oortwijn and S. Tijs	The Compromise Value for NTU-Games
9219	J.L. Horowitz and W. Härdle	Testing a Parametric Model against a Semiparametric Alternative
9220	A.L. Bovenberg	Investment-Promoting Policies in Open Economies: The Importance of Intergenerational and International Distributional Effects
9221	S. Smulders and Th. van de Klundert	Monopolistic Competition, Product Variety and Growth: Chamberlin vs. Schumpeter
9222	H. Bester and E. Petrakis	Price Competition and Advertising in Oligopoly
9223	A. van den Nouweland, M. Maschler and S. Tijs	Monotonic Games are Spanning Network Games
9224	H. Suehiro	A "Mistaken Theories" Refinement
9225	H. Suehiro	Robust Selection of Equilibria
9226	D. Friedman	Economically Applicable Evolutionary Games
9227	E. Bomhoff	Four Econometric Fashions and the Kalman Filter Alternative - A Simulation Study
9228	P. Borm, G.-J. Otten and H. Peters	Core Implementation in Modified Strong and Coalition Proof Nash Equilibria
9229	H.G. Bloemen and A. Kapteyn	The Joint Estimation of a Non-Linear Labour Supply Function and a Wage Equation Using Simulated Response Probabilities
9230	R. Beetsma and F. van der Ploeg	Does Inequality Cause Inflation? - The Political Economy of Inflation, Taxation and Government Debt

No.	Author(s)	Title
9231	G. Almekinders and S. Eijffinger	Daily Bundesbank and Federal Reserve Interventions - Do they Affect the Level and Unexpected Volatility of the DM/$-Rate?
9232	F. Vella and M. Verbeek	Estimating the Impact of Endogenous Union Choice on Wages Using Panel Data
9233	P. de Bijl and S. Goyal	Technological Change in Markets with Network Externalities
9234	J. Angrist and G. Imbens	Average Causal Response with Variable Treatment Intensity
9235	L. Meijdam, M. van de Ven and H. Verbon	Strategic Decision Making and the Dynamics of Government Debt
9236	H. Houba and A. de Zeeuw	Strategic Bargaining for the Control of a Dynamic System in State-Space Form
9237	A. Cameron and P. Trivedi	Tests of Independence in Parametric Models: With Applications and Illustrations
9238	J.-S. Pischke	Individual Income, Incomplete Information, and Aggregate Consumption
9239	H. Bloemen	A Model of Labour Supply with Job Offer Restrictions
9240	F. Drost and Th. Nijman	Temporal Aggregation of GARCH Processes
9241	R. Gilles, P. Ruys and J. Shou	Coalition Formation in Large Network Economies
9242	P. Kort	The Effects of Marketable Pollution Permits on the Firm's Optimal Investment Policies
9243	A.L. Bovenberg and F. van der Ploeg	Environmental Policy, Public Finance and the Labour Market in a Second-Best World
9244	W.G. Gale and J.K. Scholz	IRAs and Household Saving
9245	A. Bera and P. Ng	Robust Tests for Heteroskedasticity and Autocorrelation Using Score Function
9246	R.T. Baillie, C.F. Chung and M.A. Tieslau	The Long Memory and Variability of Inflation: A Reappraisal of the Friedman Hypothesis
9247	M.A. Tieslau, P. Schmidt and R.T. Baillie	A Generalized Method of Moments Estimator for Long-Memory Processes
9248	K. Wärneryd	Partisanship as Information

No.	Author(s)	Title
9249	H. Huizinga	The Welfare Effects of Individual Retirement Accounts
9250	H.G. Bloemen	Job Search Theory, Labour Supply and Unemployment Duration
9251	S. Eijffinger and E. Schaling	Central Bank Independence: Searching for the Philosophers' Stone
9252	A.L. Bovenberg and R.A. de Mooij	Environmental Taxation and Labor-Market Distortions
9253	A. Lusardi	Permanent Income, Current Income and Consumption: Evidence from Panel Data
9254	R. Beetsma	Imperfect Credibility of the Band and Risk Premia in the European Monetary System
9301	N. Kahana and S. Nitzan	Credibility and Duration of Political Contests and the Extent of Rent Dissipation
9302	W. Güth and S. Nitzan	Are Moral Objections to Free Riding Evolutionarily Stable?
9303	D. Karotkin and S. Nitzan	Some Peculiarities of Group Decision Making in Teams
9304	A. Lusardi	Euler Equations in Micro Data: Merging Data from Two Samples
9305	W. Güth	A Simple Justification of Quantity Competition and the Cournot-Oligopoly Solution
9306	B. Peleg and S. Tijs	The Consistency Principle For Games in Strategic Form
9307	G. Imbens and A. Lancaster	Case Control Studies with Contaminated Controls
9308	T. Ellingsen and K. Wärneryd	Foreign Direct Investment and the Political Economy of Protection
9309	H. Bester	Price Commitment in Search Markets
9310	T. Callan and A. van Soest	Female Labour Supply in Farm Households: Farm and Off-Farm Participation
9311	M. Pradhan and A. van Soest	Formal and Informal Sector Employment in Urban Areas of Bolivia
9312	Th. Nijman and E. Sentana	Marginalization and Contemporaneous Aggregation in Multivariate GARCH Processes
9313	K. Wärneryd	Communication, Complexity, and Evolutionary Stability

No.	Author(s)	Title
9314	O.P.Attanasio and M. Browning	Consumption over the Life Cycle and over the Business Cycle
9315	F. C. Drost and B. J. M. Werker	A Note on Robinson's Test of Independence
9316	H. Hamers, P. Borm and S. Tijs	On Games Corresponding to Sequencing Situations with Ready Times
9317	W. Güth	On Ultimatum Bargaining Experiments - A Personal Review
9318	M.J.G. van Eijs	On the Determination of the Control Parameters of the Optimal Can-order Policy
9319	S. Hurkens	Multi-sided Pre-play Communication by Burning Money
9320	J.J.G. Lemmen and S.C.W. Eijffinger	The Quantity Approach to Financial Integration: The Feldstein-Horioka Criterion Revisited
9321	A.L. Bovenberg and S. Smulders	Environmental Quality and Pollution-saving Technological Change in a Two-sector Endogenous Growth Model
9322	K.-E. Wärneryd	The Will to Save Money: an Essay on Economic Psychology
9323	D. Talman, Y. Yamamoto and Z. Yang	The $(2^{n+m+1} - 2)$-Ray Algorithm: A New Variable Dimension Simplicial Algorithm For Computing Economic Equilibria on $S^n \times R^m_+$
9324	H. Huizinga	The Financing and Taxation of U.S. Direct Investment Abroad
9325	S.C.W. Eijffinger and E. Schaling	Central Bank Independence: Theory and Evidence
9326	T.C. To	Infant Industry Protection with Learning-by-Doing
9327	J.P.J.F. Scheepens	Bankruptcy Litigation and Optimal Debt Contracts
9328	T.C. To	Tariffs, Rent Extraction and Manipulation of Competition
9329	F. de Jong, T. Nijman and A. Röell	A Comparison of the Cost of Trading French Shares on the Paris Bourse and on SEAQ International
9330	H. Huizinga	The Welfare Effects of Individual Retirement Accounts
9331	H. Huizinga	Time Preference and International Tax Competition
9332	V. Feltkamp, A. Koster, A. van den Nouweland, P. Borm and S. Tijs	Linear Production with Transport of Products, Resources and Technology
9333	B. Lauterbach and U. Ben-Zion	Panic Behavior and the Performance of Circuit Breakers: Empirical Evidence

No.	Author(s)	Title
9334	B. Melenberg and A. van Soest	Semi-parametric Estimation of the Sample Selection Model
9335	A.L. Bovenberg and F. van der Ploeg	Green Policies and Public Finance in a Small Open Economy
9336	E. Schaling	On the Economic Independence of the Central Bank and the Persistence of Inflation
9337	G.-J. Otten	Characterizations of a Game Theoretical Cost Allocation Method
9338	M. Gradstein	Provision of Public Goods With Incomplete Information: Decentralization vs. Central Planning
9339	W. Güth and H. Kliemt	Competition or Co-operation
9340	T.C. To	Export Subsidies and Oligopoly with Switching Costs
9341	A. Demirgüç-Kunt and H. Huizinga	Barriers to Portfolio Investments in Emerging Stock Markets
9342	G.J. Almekinders	Theories on the Scope for Foreign Exchange Market Intervention
9343	E.R. van Dam and W.H. Haemers	Eigenvalues and the Diameter of Graphs
9344	H. Carlsson and S. Dasgupta	Noise-Proof Equilibria in Signaling Games
9345	F. van der Ploeg and A.L. Bovenberg	Environmental Policy, Public Goods and the Marginal Cost of Public Funds
9346	J.P.C. Blanc and R.D. van der Mei	The Power-series Algorithm Applied to Polling Systems with a Dormant Server
9347	J.P.C. Blanc	Performance Analysis and Optimization with the Power-series Algorithm
9348	R.M.W.J. Beetsma and F. van der Ploeg	Intramarginal Interventions, Bands and the Pattern of EMS Exchange Rate Distributions
9349	A. Simonovits	Intercohort Heterogeneity and Optimal Social Insurance Systems
9350	R.C. Douven and J.C. Engwerda	Is There Room for Convergence in the E.C.?
9351	F. Vella and M. Verbeek	Estimating and Interpreting Models with Endogenous Treatment Effects: The Relationship Between Competing Estimators of the Union Impact on Wages

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS