Estimating the impact of endogenous union choice on wages using panel data (Revised version)

Vella, F.; Verbeek, M.J.C.M.

Publication date:
1992

CentER — for Economic Research

Discussion paper

[redacted]

No. 9232

ESTIMATING THE IMPACT OF ENDOGENOUS UNION CHOICE ON WAGES USING PANEL DATA

by Francis Vella
and Marno Verbeek

October 1992

ISSN 0924-7815

Estimating the Impact of Endogenous Union Choice on Wages Using Panel Data*

Francis VELLA
Department of Economics
Rice University
and
Marno VERBEEK
Department of Econometrics
Tilburg University

August 1991
Revised October 1992

*This paper was partially written while Vella was a visitor at the CentER for Economic Research at Tilburg University. We have benefited from useful discussions with John Antel, Donald Deere, Richard Freeman, Bertrand Melenberg, Whitney Newey and Adrian Pagan. We are also grateful to Peter Hartley, Jörn-Steffen Pischke and Theo Nijman for comments on an earlier version and to Craig Strain for preparing the data. Vella acknowledges the financial assistance from the Netherlands Organization for Scientific Research (N.W.O.). We alone are responsible for any remaining errors.

Abstract

The primary focus of this paper is the estimation of the union premium for young males in the U.S during a period of declining unionization and increasing income inequality (1980-87). We estimate the premium through a procedure developed by Vella and Verbeek (1992a), which is an extension to panel data of the control function approach based on the method of Heckman (1979). This approach enables the identification of several sources of endogeneity such as unobserved heterogeneity and business cycle effects. For the sample period examined the union premium is of the order of 30 percent. The results suggest that unionism and collective bargaining appear to protect relatively lower productivity workers from low wages. The parameters capturing the selection process indicate that the relatively less productive workers join the unions as a means of increasing their wages. Moreover, negative shocks in the economy which induce higher rates of unionism simultaneously reduce the general level of wages. The results also indicate that the decline in unionization has not significantly contributed to the observed increase in income inequality.

1 Introduction

Recent growth in the union premium (see Linneman and Wachter (1986)) and substantial increases in the degree of income inequality, generated by movements in the returns to education and the skill premium (see Juhn, Murphy and Pierce (1989)), have been accompanied by a decline in the unionization rate in the United States. One may infer that the decrease in unionization is causally related to the change in the income distribution. That is, it is possible that the changes in the union and skill premia have combined with the reductions in unionization to jointly alter the income distribution. Freeman (1991) presents evidence indicating that the decline in union membership has had only a minor role to play in the increasing degree of income inequality. Card (1991), however, concludes that twenty percent of the increase in wage inequality during the period 1973-1987 is due to the decline in unionization. To shed further light on this question it is useful to: 1) examine the factors determining union status during the later half of this period and the impact of union status on wages; 2) determine the extent to which the decline in unionization has contributed to the increase in inequality. To analyze these questions it is particularly useful to employ panel data to examine the determinants of union status, and its subsequent impact on wages, within the same group of individuals over the period of interest.

In evaluating the wage differential attributable to union membership it is common practice to account for the endogeneity of union status (for a recent survey see Robinson (1989)). The argument in support of this potential endogeneity is that individuals possess unobservable characteristics which influence their productivity in union versus non-union employment and, hence, both their ultimate wage and their choice of union or non-union employment. If union membership is endogenous the error term in least squares wage regressions is correlated with union status. In the absence of appropriate adjustments this will produce inconsistent estimates of the effect of union membership on wages. For observations on individuals several simple methods of adjusting for such endogeneity are available and involve the use of instrumental variables, control functions or, in the case of longitudinal data, the use of differencing from time or individual means.

Methods based on panel data typically assume the error attached to each unit of observation has a fixed component, correlated with union status, and a random component which is uncorrelated with union status (for a recent example see Jakubson (1991)). In that case, simple individual differencing, or treating the data as departures from individual means, will eliminate the fixed component and purge the error term of the component correlated with the union status regressor. However, these procedures, known as fixed effects estimators, are quite restrictive in the manner in which they treat the unobserved heterogeneity. In particular, it is likely that the economy will experience time specific shocks to the labor market which will also influence the union membership decision and, perhaps, the impact of union membership on wages. A preferable approach would allow a flexible decomposition of the unobserved wage equation error into an individual specific effect, a time period effect common to all individuals, and an individual/time specific effect with each of these potentially correlated with the unobserved influences determining union status. Another interesting issue is the time variation of the union effect. If, as the evidence suggests, unions compress wages then the union premium is likely to be counter-cyclical. However, as the composition of the union may vary over the business cycle it is possible that accounting for the endogeneity of membership may eliminate this effect.

In this paper we examine the impact of union status on wages while allowing a flexible composition of the terms capturing the heterogeneity and endogeneity. We examine which components of the error term generate the observed endogeneity thereby providing greater insight into the economic mechanism generating union membership. We also examine how the relative bargaining strength of unions fluctuates, if at all, with the level of economic activity and the business cycle.

We employ a new procedure proposed by Vella and Verbeek (1992a) to estimate parameters from panel data models with censored endogenous regressors. The estimator is a generalization to panel data of the control function procedure popularized by Heckman (1979). We examine data on young males taken from the National Longitudinal Survey (Youth Sample) for the period 1980 to 1987. We conclude that, contrary to earlier work, the selection bias observed in union models is not solely driven by fixed individual specific effects. Furthermore, attempts to correct for the selection effects through focusing on the individual specific effects produces underestimates of the union effect. We provide evidence that incorporating a more flexible structure into the error term produces an increased estimate of the union effect. We also find that the union effect does not vary over the sample period despite significant changes in the level of activity in the economy and the pro-cyclicality of the real wage level. Our results also suggest that the decline in unionization has contributed significantly to the increase in income inequality. As the method of estimation is an important issue we discuss the biases arising from various estimators employed in examining union effects in Section 2. WE also discuss the conventional methods of estimating union wage effects from cross sectional data. In Section 3 we introduce the empirical model and the the estimation procedure, while the empirical results are presented in Section 4. In section 5 we perform some counter factual "simulations" to assess the impact of the decline in unionization on the increase in income inequality. Section 6 contains some concluding comments.

2 Estimating Union Effects from Cross Sectional and Panel Data

Characterize the determination of an individuals wage rate in the following simple manner

$$w_{it} = \beta' X_{it} + \delta U_{it} + e_{it} \quad t = 1...T; i = 1...N \tag{1}$$

where w_{it} represents the wage of individual i in time period t; X_{it} is a vector of characteristics determining the wage level; U_{it} is a dummy variable taking the value 1 if individual i is a union member at time t; β is a vector of parameters; δ is a parameter; and e_{it} is a zero mean error term.

It is commonly assumed that the error term can be decomposed into a fixed individual effect, denoted α_i, and a time and individual specific random effect ε_{it}. However, to enable greater flexibility assume the error term also contains a time specific effect μ_t common to all individuals. It follows from our introductory comments that α_i may be correlated with U_{it} and least squares estimation of (1) will produce inconsistent estimates of β and δ. To overcome this, the fixed effects estimator is used which employs deviations from individual mean behavior annihilating the α_i term. The differencing does not, however, eliminate the μ_t or ε_{it} and failure to account for these terms, and their potential correlation with U_{it}, may mean that the bias in the estimated effect of union membership on wages has not been eliminated.

The parameters in (1) can also be estimated using "control function" procedures based on the methods developed by Heckman (1979) and Lee

(1978). This requires the specification of a union status equation and the joint distribution of the error terms in the union and wage equations. These procedures are usually confined to cross sectional studies as they are more difficult to implement in longitudinal data[1].

Control function procedures are often referred to as "alternatives" to the instrumental variable procedure. However, given the parametric assumptions, it is generally true that the estimates obtained using control function methods and the instrumental variable estimates are similar by construction. Vella and Verbeek (1992b) illustrate that although the control function and instrumental variable procedures generally identify different effects, they are vey closely related. To illustrate this consider, in addition to (1), an equation explaining union membership and a censoring mechanism, ignoring for the time being the time aspect of the data by focusing on one cross section

$$w_i = \beta' X_i + \delta U_i + e_i \tag{2}$$

$$U_i^* = \gamma' Z_i + \eta_i \tag{3}$$

$$U_i = \mathrm{I}(U_i^*), \tag{4}$$

where U_i^* is a latent variable denoting some propensity to join the union; U_i is the dummy variable generated by the censoring function I; Z_i are exogenous variables; γ is a vector of parameters to be estimated; and η_i is a normally distributed mean zero error term which is potentially correlated with e_i.

Consider the conventional methods of estimating the parameters from (2). The first is based on instrumental variable methods and projects U_i onto the exogenous variables in the system to obtain $\hat{U}_i$. Equation (1) is then estimated by OLS after replacing U_i by $\hat{U}_i$. This produces consistent estimates as the correlation between $\hat{U}_i$ and e_i is purged through the first step.

An alternative procedure employs the control function approach. This requires first step estimation of equation (3) by probit. Having obtained the estimates of γ we construct the inverse of Mill's ratio, denoted λ_i, and estimate equation (1) by OLS using λ_i as an additional regressor. This

[1] For example, while the simple two step estimators can be easily employed in cross sectional data it is inappropriate to argue that the observations are independent in the longitudinal context as is required to compute the conditional expectation of the error from the simple probit model.

methodoloy also produces consistent estimates as the inclusion of λ_i accounts for the correlation between U_i and ϵ_i.

Vella and Verbeek (1992b) argue that the control function and instrumental variables approach are essentially the same thereby explaining the similarity in estimates (see Robinson (1989)). The intuition behind this relationship is the following. As shown by Hausman (1978), the endogeneity of regressors can be accounted for by: a) replacing the endogenous regressor in the primary equation with its predicted value from an auxiliary regression; or b) including both the endogenous variable and the residuals from the auxiliary equation in the primary equation. Having acknowledged that the inverse of Mill's ratio is the generalized residual from the probit model (see Gourieroux et al. (1987) and Vella (1992)), the link between the control function and instrumental variable methodologies is immediately apparent[2].

Our discussion has highlighted the possible forms of bias in estimating (1) when U_{it} is measured without error. Freeman (1984), Chowdhury and Nickell (1985) and Card (1991), however, suggest that the mismeasurement of union status may be a common problem in the estimation of union effects from longitudinal data. Consider the impact of such measurement error. First, consider ordinary least squares estimation. As unions increase wages misclassification will *reduce* the estimated union effect. Second, consider the impact of measurement error on the fixed effect estimates of the union wage premium. Assume individuals who report frequent changes in union status are those most likely to misclassify while those who do not change status are classified correctly. The fixed effects estimator "eliminates" those observations which do not change status as the regressors appear in deviation from individual mean behavior form. Thus the fixed effect estimator places greater weight on the mismeasured observations. The influence of measurement error in the fixed effect model is to further bias *downwards* the estimate of the union effect. While the above mentioned studies propose methods to mitigate the problems of measurement error the method of estimation we now develop is not capable of incorporating these features. However, our method of estimation has other advantages which make it desirable. Having noted this, however, it is useful to consider methods which may make our proposed

[2]Another common method of accounting for endogeneity is to substitute U_i with $\Pr(U_i = 1)$. Vella and Verbeek (1992b) also discuss this procedures relationship with IV and control functions.

estimator robust to such mismeasurement.

3 Empirical model and estimation procedure

Let us consider the following three equation model of wage determination and union status

$$w_{it} = \beta' X_{it} + \delta_t U_{it} + \alpha_i + \mu_t + \varepsilon_{it} \qquad t = 1...T;\ i = 1...N \tag{5}$$

$$U^*_{it} = \gamma' Z_{it} + \theta_i + \varphi_t + \eta_{it} \qquad t = 1...T;\ i = 1...N \tag{6}$$

$$U_{it} = \mathrm{I}(U^*_{it}) \qquad t = 1...T;\ i = 1...N \tag{7}$$

where w_{it} is the log of the hourly real wage rate of individual i in time period t; U_{it} is a dummy variable denoting whether individual i's wage in period t was determined through collective bargaining; X and Z are vectors of exogenous variables; U^*_{it} is a latent variable from which U_{it} is derived through the indicator function I denoting a positive gain from union membership for individual i; β, δ and γ are parameters to be estimated; α_i and θ_i are individual specific effects; μ_t and φ_t are time specific effects common to all individuals; and ε_{it} and η_{it} are random error terms due to individual time specific shocks. We assume that these individual and time specific effects and the error terms are distributed in the following manner: $(\alpha_i, \theta_i, \mu_t, \varphi_t, \varepsilon_{it}, \eta_{it})' \sim \mathrm{MVN}(0, \Sigma)$, where Σ has the variance of the various effects on the principal diagonal and the following covariances, $\sigma_{\alpha\theta}$, $\sigma_{\mu\varphi}$ and $\sigma_{\varepsilon\eta}$ are allowed to be non zero. All other covariances are set equal to zero. Consequently, we allow for correlation between the individual specific effects in the two equations, correlation between the aggregate shocks and correlation between the contemporaneous transitory shocks. Thus the correlation between the error terms in different time periods is due to (time constant) individual heterogeneity and all contemporaneous correlation between the errors of different individuals is due to aggregate macro shocks.

Consider the economic implications of this model. First, the δ coefficient is subscripted by t to indicate that the returns to collective bargaining are not constrained to be constant across time periods. This allows us to examine whether the union effect is constant over time and whether the time variation is an artifact of the changing composition of union membership. Second, the introduction of individual and time specific effects in both the union and wage equations enable a flexible selection process. The individual and time effects are allowed to be correlated across equations as are the error terms which are unique to individuals and time periods. Consequently, the structure of this model allows a rich relationship between the unobserved influences in the economy which are simultaneously determining union status, the use of the collective bargaining and the level of wages. This represents a substantial departure from previous empirical studies endogenizing union status in earnings equations. Moreover, the structure of our model enables the estimation and identification of all these separate time and individual effects.

To estimate the model we employ the methodology developed by Vella and Verbeek (1992a), which combines the procedures in Ridder (1990), Nijman and Verbeek (1992), Verbeek and Nijman (1992) and Vella (1992). The former papers present estimation procedures that correct for attrition bias in models estimated from panel data, while the latter one suggests an estimator for models in cross sectional studies where the regressors are censored and potentially endogenous.

To proceed, rewrite (5) conditional on the vector $\underline{U}$, of length NT, denoting the union status of each individual in each period

$$E[w_{it} \mid \underline{U}] = \beta' E[X_{it} \mid \underline{U}] + \delta_t E[U_{it} \mid \underline{U}] + E[\alpha_i \mid \underline{U}]$$
$$+E[\mu_t \mid \underline{U}] + E[\varepsilon_{it} \mid \underline{U}] \quad (8)$$

where $\underline{U}$ represents the vector of observed outcomes over the entire history of U_{it} for each individual i. As the X's are exogenous they are unaffected by the conditioning. Furthermore, in the presence of conditioning on observed union status, the dummy variables denoting union status are also unaffected. To estimate equation (8) we need to generate estimates of the conditional error terms. To obtain these terms it is necessary to first estimate equation (6).

Given the structure of the error term appearing in (6) an appropriate means of estimation is the random effects probit estimator. The endogeneity bias operates in (8) through the correlation across the errors in the different equations. Once we estimate the parameters from (6) we are able to adjust for this bias by obtaining estimates of the conditional error terms. We replace the error terms with their conditional expectations and obtain a form of (8) that can be estimated by least squares.

A feature of this model is that the economic behavior of agents is partly captured in the correlation of the error terms across equations. This correlation reflects whether union status is weakly exogenous to the wage level. The test of endogeneity is based on the coefficients associated with the respective error terms as this reflects the normalized elements of the covariance matrix. Each test is a separate t-test on each coefficient as each variable captures a different form of endogeneity.

To implement our method we first estimate γ from (6). Conditional on θ and φ, the likelihood function is given by

$$L = \prod_{i,t} \Phi\left(\frac{\gamma' Z_{it} + \theta_i + \varphi_t}{\sigma_\eta}\right)^{U_{it}} \left(1 - \Phi(\frac{\gamma' Z_{it} + \theta_i + \varphi_t}{\sigma_\eta})\right)^{1-U_{it}} \tag{9}$$

where Φ denotes the cumulative probability function of the standard normal distribution. Maximizing (9) requires optimization over $k+N+T$ parameters. However, it is possible to integrate out θ_i by employing our distributional assumptions. This gives

$$L = \prod_i \int \prod_t \Phi(\frac{\gamma' Z_{it} + \theta_i + \varphi_t}{\sigma_\eta})^{U_{it}} \times$$

$$\left(1 - \Phi(\frac{\gamma' Z_{it} + \theta_i + \varphi_t}{\sigma_\eta})\right)^{1-U_{it}} \frac{1}{\sigma_\theta}\phi(\frac{\theta_i}{\sigma_\theta})d\theta_i \tag{10}$$

and this produces consistent estimates of γ, σ_θ^2 and φ_t (given some normalization on σ_η^2). The integral in (10) has to be determined numerically, which is fairly straightforward using Gaussian quadrature procedures (cf. Butler and Moffitt (1982)). Given our distributional assumptions regarding φ_t it is also possible to integrate out these parameters. However, for computational simplicity we did not do so. With these estimates of γ and φ we obtain the conditional expectations of the residuals in the following manner[3]

$$E[\alpha_i \mid \underline{U}] = \sigma_{\alpha\theta}\left[\frac{T}{\sigma_\eta^2 + T\sigma_\theta^2} E[\bar{v}_{i.} \mid \underline{U}]\right] \tag{11}$$

[3]The derivation of these terms is provided in the appendix.

$$E[\mu_t \mid \underline{U}] = \sigma_{\mu\varphi}\left[\frac{N}{\sigma_\eta^2 + N\sigma_\varphi^2}E[\bar{v}_{.t} \mid \underline{U}]\right] \tag{12}$$

$$E[\varepsilon_{it} \mid \underline{U}] = \sigma_{\varepsilon\eta}\left[\sigma_\eta^{-2}E[v_{it} \mid \underline{U}] - \frac{T\sigma_\theta^2}{\sigma_\eta^2(\sigma_\eta^2 + T\sigma_\theta^2)}E[\bar{v}_{i.} \mid \underline{U}]\right.$$

$$\left. - \frac{N\sigma_\varphi^2}{\sigma_\eta^2(\sigma_\eta^2 + N\sigma_\varphi^2)}E[\bar{v}_{.t} \mid \underline{U}]\right] \tag{13}$$

where $v_{it} = \theta_i + \varphi_t + \eta_{it}$, $\bar{v}_{i.} = \frac{1}{T}\sum_{t=1}^{T} v_{it}$ and $\bar{v}_{.t} = \frac{1}{N}\sum_{i=1}^{N} v_{it}$.

Due to the normality assumption the conditional expectations are linear in the covariances. The remaining expressions in (11)-(13) are known functions of the parameters in the probit model (6) and can thus be estimated consistently once the probit model is estimated. The mathematical form for these functions is given in the appendix and involves one dimensional numerical integration. Thus the estimation procedure is straightforward. First we estimate the union status equation by employing the random effects probit estimator. With these estimates we then generate the conditional expectations of the error terms. We then estimate the primary log wage equation by least squares with the conditional error terms appearing as additional regressors. Our procedure is clearly an extension of conventional cross sectional control function procedures, although the time dimension afforded through the use of panel data enables greater insight into the form of the endogeneity of union choice and thus the economic behavior of the agents in this model.

A major limitation of the control function approach, or any method based on the use of conditional expectations to replace latent variables, is the reliance on strict distributional assumptions. Our approach is no different although we now employ some results from the semi-parametric and diagnostic testing literature to restrict the importance of our assumptions. First note however, that the estimation of (9) requires that we correctly specify the distribution of the latent effects. Although this assumption cannot be easily relaxed it is possible to test the normality assumption by employing the conditional moment framework of Newey (1985) as discussed with respect to normality tests in cross sectional probit models by Pagan and Vella (1989) and in panel data probit models by Vella and Verbeek (1992b). In our application we perform seperate normality tests for θ_i and η_{it} as well as a joint normality test for both effects, based upon the LM test against the Pearson family of distributions (cf. Ruud (1984)).

We estimate the random effects model under the assumption of normality and the distributional assumption is employed again in generating equations (11), (12) and (13). The normality assumption allows us to express the latent effects in the wage equation as linear functions of the first moments of the random latent effects in the union equation. However, using the results of Lee (1984), Gallant and Nychka (1987) and Pagan and Vella (1989) it is possible to capture sensible departures from normality in (5) by expressing the latent effects in the wage equations as higher order functions of the latent effects from the union equation. To capture these potential departures from normality we augment (8) with powered values of the latent effects. By employing this approach we not only obtain estimates which are consistent in the absence of normality but a test on the statistical significance of the higher order terms is a test of normality.

4 Empirical Results

To estimate the model we employ data taken from the National Longitudinal Survey (Youth Sample). We examine a sample of full time working males who have completed their schooling by 1980 and follow their paths over the period 1980 to 1987. We exclude individuals who fail to provide sufficient information to be included in every year over the eight year period, leaving a sample of 545 observations. The summary statistics for the total period are reported in Table 1. Our measure of union membership is based on the question reflecting whether or not the individual had his wage set in a collective bargaining agreement[4]. This measure of union status displays some year to year variation indicating movement in and out of union membership[5]. Overall, union members enjoy an unconditional wage premium of around 15 percent.

The random effects probit estimates of union membership are reported in Table 2. Several of the explanatory variables have a statistically significant impact on the probability of union membership. The negative coefficient on

[4]We will refer to those who responded yes to this question as being union members.

[5]For the period examined union membership in the private sector, in our sample, ranged from 21 % in 1986 to 26 % in 1982 and 1987.

the industry dummies reflects the sizable unionization rate in the control group which is the public sector. The time dummies display an increasingly negative pattern with respect to the control group of 1980, but only 1985 and 1986 are significant at the ten percent level. This is consistent with the aggregate data which indicates sizable decreases in unionization over this period. It should be noted however that the observations in our data display a weaker tendency to leave unions than is revealed by aggregate data.

The coefficients on the dummy variables denoting that the individual is black or hispanic are both positive and statistically significant. Furthermore, the coefficients are large in magnitude. This is consistent with earlier studies and also consistent with our introductory comments regarding the composition of unions. If these groups suffer from discrimination, they are likely to look to unions for protection in the labor market.

An important result is captured by the magnitude of the estimate of σ_{θ}^2. This estimate of 0.73 indicates that seventy three percent of the total variance is due to across individual variation[6]. This indicates a great degree of unobserved heterogeneity and highlights the importance of using the random effects probit model in preference to the conventional probit model[7]. This also highlights the inappropriateness in models from panel data of the use of the conventional inverse Mill's ratio as the additional regressor to account for the endogeneity.

The conditional moment tests for normality of the error components in (6), excluding the time effects, gave the following results. The separate tests against normality of θ_i and η_{it} resulted in values of 0.96 and 7.08, respectively. Under the null hypothesis the test statistics are Chi-squared distributed with 2 degrees of freedom and, consequently, we do not take our results as evidence against the null[8]. The joint test on normality of both components, which corresponds to a χ^2 with 4 degrees of freedom, yields the insignificant value of 8.59.

Focusing on estimation of the wage equation, we first estimate equation (5) by ordinary least squares and constrain the union effect to be time invari-

[6] The normalization used is $\sigma_{\theta}^2 + \sigma_{\eta}^2 = 1$.

[7] It should be noted that several of the estimates reported in Table 2 were significantly different from the estimates obtained from a conventional probit model estimated over the data set obtained by pooling the cross sections.

[8] Note that 7.08 is not significant at the 2.5 % level.

ant. This estimate of the union effect, along with the additional regressors, is shown in column (1) of Table 3. The point estimate of the union effect is of the order of 15 percent and, given the evidence in Robinson (1989), appears to be quite low. Recall, however, that this estimate is contaminated with the endogeneity generated by the various latent effects.

We now estimate equation (5) using the fixed effects estimator, recalling that this procedure accounts only for the individual specific endogeneity and suffers significantly in the presence of measurement error. These estimates are reported in column (2) of Table 3 and, consistent with previous results, the most recent being Angrist and Newey (1991) and Jakubson (1991), the estimate of the union effect falls markedly. The point estimate of 7.9% is approximately half of the already downward biased least squares estimate. It is difficult to assess the merit of this equation as the fixed effects estimator annihilates many of the variables. It is interesting to note however that the coefficient on the experience variable appears to be unusually high and the rural effect, while not significant at the five percent level, has a sign inconsistent with our a priori expectations.

In column (3) of this table we re-estimate the OLS equation and include a variable capturing the state of the economy, the unemployment rate, and an interaction of this cyclical variable with union status. Due to multicollinearity the union effect, as measured by the coefficient for the union variable, now is not statistically different from zero. Also note that the coefficient on the unemployment rate, included to capture cyclical influences, is insignificant. This does not support the results of Bils (1985) and Keane, Moffitt and Runkle (1988) who concluded that the real wage was pro-cyclical although the large standard errors attached to our estimates make it impossible to be precise. There is also no evidence that the union effect is affected by the level of unemployment.

Column (4) reports the estimates of the wage equation when we include the three correction terms discussed in the previous section. We only include the variables in their first moments so there is no scope for non-normality to be captured. The estimates indicate that the union effect increase dramatically to 46 percent. While this estimate is in the range of previous estimates it seems somewhat high. Furthermore, at this stage all three corrections terms are statistically significant and negative[9]. This indicates that not only fixed

[9]Standard errors in Table 3 are computed taking into account the covariance structure

ant. This estimate of the union effect, along with the additional regressors, is shown in column (1) of Table 3. The point estimate of the union effect is of the order of 15 percent and, given the evidence in Robinson (1989), appears to be quite low. Recall, however, that this estimate is contaminated with the endogeneity generated by the various latent effects.

We now estimate equation (5) using the fixed effects estimator, recalling that this procedure accounts only for the individual specific endogeneity and suffers significantly in the presence of measurement error. These estimates are reported in column (2) of Table 3 and, consistent with previous results, the most recent being Angrist and Newey (1991) and Jakubson (1991), the estimate of the union effect falls markedly. The point estimate of 7.9% is approximately half of the already downward biased least squares estimate. It is difficult to assess the merit of this equation as the fixed effects estimator annihilates many of the variables. It is interesting to note however that the coefficient on the experience variable appears to be unusually high and the rural effect, while not significant at the five percent level, has a sign inconsistent with our a priori expectations.

In column (3) of this table we re-estimate the OLS equation and include a variable capturing the state of the economy, unemployment. The coefficient on the union variable, now is not statistically different from zero. Also note that the coefficient on the unemployment rate, included to capture cyclical influences, is insignificant. This does not support the results of Bils (1985) and Keane, Moffitt and Runkle (1988) who concluded that the real wage was pro-cyclical although the large standard errors attached to our estimates make it impossible to be precise. There is also no evidence that the union effect is affected by the level of unemployment.

Column (4) reports the estimates of the wage equation when we include the three correction terms discussed in the previous section. We only include the variables in their first moments so there is no scope for non-normality to be captured. The estimates indicate that the union effect increase dramatically to 46 percent. While this estimate is in the range of previous estimates it seems somewhat high. Furthermore, at this stage all three corrections terms are statistically significant and negative[9]. This indicates that not only fixed

[9]Standard errors in Table 3 are computed taking into account the covariance structure

individual effects are important but there are also time and individual/time effects influencing the union effect.

Consistent with our discussion in section 3 we include powered up values of the residuals to capture the possibility of non-normal disturbances in equation (5). The order of the non-linear terms was chosen by cross-validation. The cross-validation criterium we use is the sum of squares of prediction errors from predicting one observation using coefficient estimates based on all other observations[10]. Table 4 gives the cross-validation criteria for values of k, the power of the highest order residual, between 0 and 6. The value of k which minimizes the CV criterium is 3 or 4. However, because the criterium value for $k = 2$ is close to the minimal value, and because the differences in point estimates for $k = 2$ and $k = 3$ are small, we prefer to choose $k = 2$. This choice preserves degrees of freedom and reduces collinearity in the model.

The results of the wage regression with linear and squared residuals included are shown in column (5) of column 3. They reveal that the coefficient on the union dummy variable is now 32 percent. The coefficients on the included correction terms are jointly statistically significant, revealing the operation of strong selection forces in the economy and the higher order terms indicate the presence of non-normality in the wage equation. While this union effect also appears high it is in the range of reported estimates of the union effect[11]. Furthermore, as the sample comprises only young workers and as unions flatten the age profile it is expected that their union premium is relatively higher. This is consistent with the results of Card (1991) who estimated that workers at the lower end of the wage distribution received a union premium in the order of 30%. Given that younger workers will comprise a majority of the workers in the lower tail of the distribution our estimates are remarkably similar to those reported by Card. Finally, we estimate (5) with all three correction terms and their second orders along with the unemployment variable and the interaction term. These results are reported in column

---

of the error terms and, for the two-step results, using the appropriate formulae given in, for example, Newey (1984) or Pagan (1986). Ignoring the fact that the included residuals are based on a first stage estimation, produces standard errors that underestimate the correct ones by a factor between 1 and 11 %.

[10] A recent discussion on the optimality of several cross-validation criteria is given in Andrews (1991).

[11] Robinson (1989) reports an estimate of 43 percent and Linnemann and Wachter (1986) present estimates in excess of 50 percent.

6. This results suggest a union effect of a similar magnitude to the previous column. The point estimate of 30% is significant at the 5 percent level, while the cyclical effect is not. There is no sign of any interactive effects.

Let us first focus on the interpretation of these corrections' coefficients in this preferred model. The evidence here, and in many other studies, indicates the union effect is downward biased by the presence of selection effects. The statistically significant and negative coefficients on the selection terms indicate that the workers who receive lower wages, after conditioning on their characteristics, and in the absence of unions, are those most likely to be in the union. The union appears to act as a protector for lower paid workers. This is consistent with the findings of Heywood (1990) that minorities displayed a greater tendency to queue for union jobs than whites. It is also consistent with Robinson (1989) who concluded that there was no support for the popular argument that better workers are chosen from a queue by the union. However, while this evidence does not support the queuing hypothesis it does not refute it. It is possible that the less productive workers queue to join the union and the union chooses the better workers from among these candidates.

Further support for our "protective" view of unions is captured by the coefficient on the time specific correction. This indicates that negative shocks to the economy which induce higher rates of collective bargaining also lead to lower levels of wages. Thus, in times when wages are negatively affected, more workers seek union membership. This result appears consistent with decreasing union membership during the eighties when economic growth was higher than in the previous decade.

It is worth noting that in column (4) under the strict assumption of normality it appeared that all three latent effects were having an impact on the wage level. When the departures from normality were accounted for, through the inclusion of the powered up residuals, this is still the case, while the residual capturing the individual specific effects has a significant higher order effect. Consequently, we have to reject the normality assumption of α_i in (5).

The coefficient on the unemployment rate is negative, though not significant. This result is consistent with previous work which found a pro-cyclical real wage following the selection bias adjustment. Although the adjustments in Keane et al. (1988) and Bils (1985) took a different form it is possible that the selection mechanisms share some component pertaining to the effect of the business cycle on real wages. It is also worth noting that the interaction term between union status and the unemployment rate is statistically insignificant in this preferred equation. This result, consistent with that found by Meghir and Whitehouse (1992) for the United Kingdom, indicates that the union differential does not appear to be influenced by cyclical factors.

We noted above that much of the controversy in the union wage differential literature has focused on the range of estimates of the union effect. However given that the effect is typically assessed using cross section, and more recently, panel data, this debate would only be valid if the time periods being examined were identical or the union effect was time invariant. We have noted that in addition to our results substantial empirical evidence has been provided, see for example, Bils (1985), Keane, Moffitt and Runkle (1988), that the real wage shows some relationship with the business cycle. It is not immediately apparent how the union differential should respond to the business cycle, if at all, and how the union wage should differ from the non-union wage although there is empirical evidence to suggest that the union premium is counter cyclical (see Wunnava and Honney (1991) for a recent discussion). To explore this possibility we estimate our preferred equation while including a dummy variable denoting union membership in a particular year. While there appeared to be some time variation there was no instance where the time varying union effect was significantly different from the average (time-constant) union effect.

It has also been established that union members often receive different returns to their various characteristics (see Perloff and Sickles (1987)) and, in particular, face a lower rate of return to education. To incorporate this possibility we re-estimated our preferred education and included an additional term capturing the interaction between education and union membership. The coefficient on this additional regressor was not statistically significant from zero. We also examined whether the union premium differed by industry. However, due to the relatively small sample size compared to the large number of industries this methodology only produced imprecise estimates of the union premium.

5 Unionization and Income Inequality

An important issue in labor economics is the substantial increase in the degree of income inequality which has occurred in the last decade (see for example Juhn, Murphy, and Pierce (1989)). It is possible that the decline in unionization has contributed to this increasing inequality and thus it is useful to employ our estimates to examine the extent to which this has occurred in our data.

In assessing wage inequality most studies have primarily focused on the standard error in wage equations. We, however, examine this issue by performing the following simple experiment. Using our random effects probit model estimates we predict union status for the observations over the eight years in our data period. Using this predicted union status and our wage equation estimates we then predict the wage of each individual for each year of the sample. We then compute the empirical wage distribution and the corresponding Gini coefficient for each year. Having established these measures of income distribution we now focus on a counter factual case of some interest. We examine how the income distributions would appear if there had been no decline in unionization. To obtain these counter factual distribution we predict union status for each individual in each period after we have set the time effects in the random effects probit model equal to zero. If we are willing to attribute the decline in unionization to these effects this is a reasonable way to generate the counter factual union status. We then predict the wage each individual would receive in each time period and compute the income distribution and Gini coefficient for each year.

We now have the Gini coefficients for the actual and counter factual cases and by generating their ratio we are able to examine the impact of declining union status on wage distributions. It should be noted that as predicted wage distributions will generally have less variance than actual wage distributions our estimates of inequality will be biased downward. However bearing this in mind it is still useful to make this comparisons across distributions.

In generating this distributions we treated union status in two ways. First we generated the predicted union status. That is, we generated the latent variable and then assigned values of one for union status to those with positive values and zero for union status to those with negative values. This ratio of actual/counter factual is in column (1) of Table 5. An alternative approach is to generate the predicted probability from the random effects model with and without the time dummies being set to zero and use the probability as the appropriate value for union status. This ratio is reported in column (2) of Table 5. A third alternative is based on predicting union status as a random drawing from a binomial distribution with the probability of success equal to the predicted probability based on the random effects probit model. This approach results in a fair amount of noise in the Gini coefficients without giving quantitatively different results. Results of this approach are therefore not presented.

The results in Table 5 reveal that the decline in unionization has not contributed to an increase in income inequality over the period 1980-1987. Independently of the way in which union status was predicted (either as a discrete value or as a probability), our measure of income inequality hardly differs between the actual and the counterfactual case. A possible explanation for this finding may be that in our sample of young males both the decline in unionization as well as the increase in income inequality are not as pronounced as in the entire economy.

6 Conclusions

The primary focus of this paper is the estimation of the union premium for young males during a period of declining unionization and increasing income inequality. We proceed by employing a new methodology for evaluating the impact of unions on wages using panel data. The estimator we employ is an extension of the commonly employed control function approach. The estimator controls for the fixed and specific time effects operating through the union membership decision. We also test for sources of endogeneity which present greater insight into the mechanism driving union membership.

Our empirical work identifies several important results. First, for the data period examined, the union premium is in the order of thirty percent. Second, the parameters capturing the selection process indicate that the less productive workers join unions as a means of increasing their wage. Moreover, the negative shocks in the economy which induce higher rates of unionism simultaneously reduce the level of wages in general. In this context, the union movement can be seen as a form of protection to the workers against a declining economy and negative personal shocks. The results also indicate that, for the sample period examined, there is no evidence of time variance in the union effect after accounting for the endogeneity of membership. Finally, a simple simulation indicates that the decrease in unionization has not contributed to the increase in income inequality.

Appendix

In this appendix we follow Vella and Verbeek (1992a) and sketch the estimation method and derive the appropriate correction terms. Represent the respective error terms in the following manner

$$e_{it} = \alpha_i + \mu_t + \varepsilon_{it}; \quad v_{it} = \theta_i + \varphi_t + \eta_{it} \tag{14}$$

We need to compute the conditional expectation of the elements of e_{it} given the NT vector $\underline{U}$ (i.e. given the inequality constraints on all elements of v_{it}). Employing our assumption of joint normality the conditional expectation of e_{it} given the vector v can be derived from the standard formulae for the conditional expectation of two normally distributed vectors. This results in

$$E[\alpha_i \mid v] = \sigma_{\alpha\theta}\left[\frac{T}{\sigma_\eta^2 + T\sigma_\theta^2}\bar{v}_{i.} - \frac{\sigma_\varphi^2 NT}{(\sigma_\eta^2 + T\sigma_\theta^2)(\sigma_\eta^2 + T\sigma_\theta^2 + N\sigma_\varphi^2)}\bar{v}_{..}\right] \tag{15}$$

$$E[\mu_t \mid v] = \sigma_{\mu\varphi}\left[\frac{N}{\sigma_\eta^2 + N\sigma_\varphi^2}\bar{v}_{.t} - \frac{\sigma_\theta^2 NT}{(\sigma_\eta^2 + T\sigma_\theta^2)(\sigma_\eta^2 + T\sigma_\theta^2 + N\sigma_\varphi^2)}\bar{v}_{..}\right] \tag{16}$$

and

$$E[\varepsilon_{it} \mid v] = \sigma_{\varepsilon\eta}\left[\frac{1}{\sigma_\eta^2}v_{it} - \frac{T\sigma_\theta^2}{\sigma_\eta^2(\sigma_\eta^2 + T\sigma_\theta^2)}\bar{v}_{i.} - \frac{N\sigma_\varphi^2}{\sigma_\eta^2(\sigma_\eta^2 + N\sigma_\varphi^2)}\bar{v}_{.t}\right.$$
$$\left. + \frac{T\sigma_\theta^2}{\sigma_\eta^2 + T\sigma_\theta^2}\frac{N\sigma_\varphi^2}{\sigma_\eta^2 + N\sigma_\varphi^2}\frac{2\sigma_\eta^2 + T\sigma_\theta^2 + N\sigma_\varphi^2}{\sigma_\eta^2(\sigma_\eta^2 + T\sigma_\theta^2 + N\sigma_\varphi^2)}\bar{v}_{..}\right] \tag{17}$$

where $\bar{v}_{..} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} v_{it}$. To obtain the conditional expectations given the vector $\underline{U}$ replace the v_{it}'s in (15)-(17) by their conditional expectations given the $\underline{U}$. These conditional expectations are complicated because $\underline{U}$ is determined by an NT-variate probit model. To simplify computations we condition on the time effects, φ_t, in addition to $\underline{U}$. Note that, conditional on φ_t, the error terms in the probit model are independent across individuals.

To derive the conditional expectations of v_{it} given $\underline{U}$ we first note that $E[v_{it} \mid \underline{U}, \varphi] = E[v_{it} \mid \underline{U}_i, \varphi]$, where $\underline{U}_i$ is the T vector of observed outcomes for individual i and $\varphi = (\varphi_1, ..., \varphi_T)'$ Next we use

$$E[\theta_i + \eta_{it} \mid \underline{U}_i, \varphi] = \int_{-\infty}^{\infty}[\theta_i + E[\eta_{it} \mid \underline{U}_i, \varphi, \theta_i]]\, f(\theta_i \mid \underline{U}_i, \varphi)d\theta_i, \tag{18}$$

where $E[\eta_{it} \mid \underline{U}_i, \varphi, \theta_i] = E[\eta_{it} \mid U_{it}, \varphi_t, \theta_i]$ is the usual generalized residual of the probit model given by

$$E[\eta_{it} \mid \underline{U}_i, \varphi, \theta_i] = (2U_{it} - 1)\sigma_\eta \frac{\phi(b_{it})}{\Phi(b_{it})}, \tag{19}$$

where $b_{it} = (2U_{it} - 1)(\gamma' Z_{it} + \theta_i + \varphi_t)/\sigma_\eta$. In (18) we integrate over the conditional distribution of θ_i given $\underline{U}_i$ and φ, which is given by

$$f(\theta_i \mid \underline{U}_i, \varphi) = \frac{\prod_{s=1}^{T} \Phi(b_{is})\sigma_\theta^{-1}\phi(\theta_i/\sigma_\theta)}{\int \prod_{s=1}^{T} \Phi(b_{is})\sigma_\theta^{-1}\phi(\theta_i/\sigma_\theta)d\theta_i}. \tag{20}$$

Consequently, given the parameter estimates for the probit model (including the variance components) the generalized residual for the random effects probit model can be computed from (18) using (19) and (20). This requires numerical integration over one dimension (both in (20) and (18)). By construction of the generalized residuals the average residual has mean zero. Consequently, in computing the correction terms from (15)-(17) the last terms (involving $\bar{v}_{..}$) are identically zero and can be deleted.

References

Andrews, D.W.K. (1991) "Asymptotic Optimality of Generalized C_L, Cross-Validiation, and Generalized Cross-Validation in Regression with Heteroskedastic Errors", *Journal of Econometrics*, 47, 359-377.

Angrist, J.D. and W.K. Newey (1991) "Over-Identification Tests in Earnings Functions with Fixed Effects", *Journal of Business and Economic Statistics*, 9, 317-323.

Bils, M. (1985) "Real Wages over the Business Cycle: Evidence from Panel Data", *Journal of Political Economy*, 93, 666-689.

Butler, J.S. and R. Moffitt (1982) "A computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model", *Econometrica*, 50, 761-764.

Card, D. (1991) "The Effects of Unions on the Distribution of Wages: Redistribution or Relabelling", working paper 287 Industrial Relations Section, Princeton University.

Chowdhury, G.,and S. Nickell (1985) "Hourly Earnings in the United States:Another Look at Unionization, Schooling, Sickness, and Unemployment Using The PSID Data", *Journal of Labor Economics*, 3, 38-69.

Freeman, R. (1984) "Longitudinal Analyses of the Effects of Trade Unions" *Journal of Labor Economics*, 2, 1-26.

Freeman, R. (1991) "How Much has De-Unionization Contributed to the Rise in Male Earnings Inequality", NBER working paper No. 3826.

Gallant, A. and D. Nychka (1987) "Semi-Nonparametric Maximum Likelihood", *Econometrica*, 55, 363-390.

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

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

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

Heywood, J.S. (1990) "Who Queues for a Union Job", *Industrial Relations*, 29, 99-127.

Jakubson, G. (1991) "Estimation and Testing of the Union Wage Effect Using Panel Data", *The Review of Economic Studies*, 58, 971-991.

Juhn, C., K. Murphy and B. Pierce (1989) "Wage Inequality and the Rise in the Returns to Skill", NBER conference on Labor.

Keane, M., R. Moffitt and D. Runkle (1988) "Real Wages over the Business Cycle: Estimating the Impact of Heterogeneity with Micro Data", *Journal of Political Economy,* 96, 1232-1265.

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

Lee, L.F. (1984) "Tests for the Bivariate Normal Distribution in Econometric Models with Selectivity", *Econometrica,* 52, 843-863.

Linneman, P., and M. Wachter (1986) "Rising Union Premiums and the Declining Boundaries Among Noncompeting Groups", *American Economic Review,* 76, 103-108.

Meghir, C. and E. Whitehouse (1992) "The Evolution of Wages in the United Kingdom: Evidence from Micro Data", mimeo, Institute for Fiscal Studies, London.

Newey, W.K. (1984) "A Method of Moments Interpretation of Sequential Estimators", *Economics Letters,* 14, 201-206.

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

Nijman, Th.E. and M. Verbeek (1992) "Nonresponse in Panel Data: the Impact on Estimates of a Life Cycle Consumption Function", *Journal of Applied Econometrics,* 7, 243-257.

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

Pagan, A. (1986) "Two Stage and Related Estimators and Their Applications", *Review of Economic Studies,* 53, 517-538.

Pagan, A. and F. Vella (1989) "Diagnostic Tests for Models based on Unit Record Data: A Survey", *Journal of Applied Econometrics,* 4, s29-s60.

Perloff, J.M. and R.C. Sickles (1987) "Union Wage Hours, and Earnings Differentials in the Construction Industry", *Journal of Labor Economics,* 5, 174-210.

Ridder, G. (1990) "Attrition in Multi-Wave Panel Data", in: J. Hartog, G. Ridder and J. Theeuwes, eds., *Panel Data and Labor Market Studies,* Elsevier, North-Holland.

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

Ruud, P.A. (1984) "Tests of Specification in Econometrics", *Econometric Reviews*, 3, 211-242.

Vella, F. (1992) "A Simple Estimator for Simultaneous Models with Censored Endogenous Regressors", Rice University, Houston; forthcoming in *International Economic Review.*

Vella, F. and M. Verbeek (1992a) "Estimating and Testing Simultaneous Equation Panel Data Models with Endogenous Explanatory Variables", mimeo, Rice University, Houston.

Vella, F. and M. Verbeek (1992b) "The Relationship Between Competing Estimators of the Union Wage Effect", mimeo, Rice University, Houston.

Verbeek, M., and T. Nijman (1992) "Testing for Selectivity Bias in Panel Data Models", *International Economic Review,* 33, 681-703.

Wunnava, P., and J.Honney (1991) "The Union-Non Union Wage Differential over the Business Cycle: Evidence from the PSID", *Economics Letters,* 37, 97-103.

variable	definition	mean	standard deviation
School	Years of schooling	11.76	1.75
Exper	Age-6-School	6.51	2.83
Exper2	Experience Squared	50.42	40.78
Union	Wage set by collective Bargaining	.24	.43
Mar	Married	.44	.50
Black	Black	.12	.32
Hisp	Hispanic	.16	.36
Health	Has health disability	.02	.13
Rural	Lives in rural area	.20	.40
NE	Lives in North East	.19	.39
NC	Lives in Northern Central	.26	.44
S	Lives in south	.35	.48
Wage	Log of hourly wage	1.65	.53
EWage	Hourly Wage ($)	5.91	3.20
Wdif	Union differential	0.87	
Industry Dummies			
AG	Agricultural	.03	
MIN	Mining	.02	
CON	Construction	.08	
TRAD	Trade	.27	
TRA	Transportation	.06	
FIN	Finance	.04	
BUS	Business & Repair Service	.08	
PER	Personal Service	.02	
ENT	Entertainment	.02	
MAN	Manufacturing	.28	
PRO	Professional & Related Service	.08	
PUB	Public Administration	.04	

Table 1: Descriptive statistics, 1980–1987

Variable	Estimate	Variable	Estimate
Constant	-.667	NE	.281*
	(.486)		(.113)
Exper	.079	S	-.003
	(.043)		(.108)
Exper2	-.004*	NC	.194*
	(.002)		(.088)
School	-.010	AG	-.466*
	(.034)		(.120)
Mar	.079	MIN	-.109
	(.045)		(.150)
Black	.505*	CON	-.406*
	(.142)		(.106)
Hisp	.334*	MAN	-.154*
	(.143)		(.077)
Rural	.023	TRA	-.041
	(.064)		(.086)
Health	-.156	TRAD	-.486*
	(.118)		(.080)
D81	-.081	FIN	-1.149*
	(.073)		(.201)
D82	-.107	BUS	-.648*
	(.092)		(.108)
D83	-.199	PER	-.505*
	(.124)		(.159)
D84	-.223	ENT	-.176
	(.156)		(.143)
D85	-.324	PRO	-.109
	(.187)		(.084)
D86	-.406	σ_θ^2	.736*
	(.219)		(.023)
D87	-.231		
	(.247)		
Log-Likelihood Value -1590.1			

Table 2: Random Effects Probit Estimates of Union Membership

Variable	(1) OLS	(2) FE	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Constant	.219	-	.293*	.230	.277	.406*
	(.128)		(.148)	(.166)	(.167)	(.201)
Union	.146*	.079*	.103	.461*	.317*	.299
	(.026)	(.018)	(.099)	(.126)	(.138)	(.169)
Unemployment Rate	-	-	-.008	-	-	-.011
			(.007)			(.008)
Unemployment Rate*Union	-	-	.005	-	-	.002
			(.012)			(.012)
School	.091*	-	.089*	.088*	.087*	.085*
	(.008)		(.009)	(.009)	(.009)	(.009)
Exper	.075*	.112*	.077*	.057*	.055*	.054*
	(.011)	(.008)	(.011)	(.016)	(.017)	(.017)
Exper2	-.002*	-.004*	-.002*	-.001	-.001	-.001
	(.0008)	(.0005)	(.0009)	(.0009)	(.0009)	(.0009)
Hisp	-.058	-	-.058*	-.086*	-.067	-.067
	(.041)		(.042)	(.045)	(.046)	(.046)
Black	-.154*	-	-.154*	-.203*	-.179*	-.178*
	(.044)		(.044)	(.050)	(.051)	(.051)
Rural	-.131*	.050	-.131*	-.133*	-.134*	-.133*
	(.031)	(.027)	(.031)	(.032)	(.032)	(.032)
Mar	.110*	.039*	.110*	.110*	.106*	.106*
	(.024)	(.017)	(.024)	(.024)	(.024)	(.024)
α	-	-	-	-.074*	-.062	-.062
				(.036)	(.037)	(.037)
μ	-	-	-	-.004*	-.004*	-.004*
				(.002)	(.002)	(.002)
ε	-	-	-	-.110*	-.069	-.069*
				(.038)	(.041)	(.041)
α^2					.036*	.036*
					(.015)	(.016)
μ^2					-.0002	-.0003
					(.0001)	(.0002)
ε^2					-.002	.001
					(.004)	(.004)
Adjusted R^2	.260	.186	.260	.264	.267	.267

Table 3: Wage Regressions with Union Effects

Note: All regressions include industry and region dummy variables.

Highest order k	Est. Union Effect	CV value
0	0.146	920.8
1	0.461	916.1
2	0.317	913.2
3	0.316	913.1
4	0.422	913.1
5	0.400	914.2
6	0.408	915.4

Table 4: Cross-Validation for Order of Residuals

Year	GiniA	GiniB
1980	1.000	1.000
1981	1.002	1.000
1982	1.007	1.001
1983	0.987	1.999
1984	0.991	1.004
1985	0.980	1.000
1986	0.995	0.999
1987	1.006	1.000

Table 5: Impact of Union Decline on Wage Inequality

Presented are ratios of Gini coefficients based on predicted wage distributions with and without time effects in the unionization process. GiniA is based on zero-one predictions of union membership while GiniB is based on predicted probabilities.

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

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9111	F. van der Ploeg and Th. van de Klundert	Political Trade-off between Growth and Government Consumption
9112	Th. Nijman, F. Palm and C. Wolff	Premia in Forward Foreign Exchange as Unobserved Components
9113	H. Bester	Bargaining vs. Price Competition in a Market with Quality Uncertainty
9114	R.P. Gilles, G. Owen and R. van den Brink	Games with Permission Structures: The Conjunctive Approach
9115	F. van der Ploeg	Unanticipated Inflation and Government Finance: The Case for an Independent Common Central Bank
9116	N. Rankin	Exchange Rate Risk and Imperfect Capital Mobility in an Optimising Model
9117	E. Bomhoff	Currency Convertibility: When and How? A Contribution to the Bulgarian Debate!
9118	E. Bomhoff	Stability of Velocity in the G-7 Countries: A Kalman Filter Approach
9119	J. Osiewalski and M. Steel	Bayesian Marginal Equivalence of Elliptical Regression Models
9120	S. Bhattacharya, J. Glazer and D. Sappington	Licensing and the Sharing of Knowledge in Joint Ventures
9121	J.W. Friedman and L. Samuelson	An Extension of the "Folk Theorem" with Continuous Reaction Functions
9122	S. Chib, J. Osiewalski and M. Steel	A Bayesian Note on Competing Correlation Structures in the Dynamic Linear Regression Model
9123	Th. van de Klundert and L. Meijdam	Endogenous Growth and Income Distribution
9124	S. Bhattacharya	Banking Theory: The Main Ideas
9125	J. Thomas	Non-Computable Rational Expectations Equilibria
9126	J. Thomas and T. Worrall	Foreign Direct Investment and the Risk of Expropriation
9127	T. Gao, A.J.J. Talman and Z. Wang	Modification of the Kojima-Nishino-Arima Algorithm and its Computational Complexity

No.	Author(s)	Title
9128	S. Altug and R.A. Miller	Human Capital, Aggregate Shocks and Panel Data Estimation
9129	H. Keuzenkamp and A.P. Barten	Rejection without Falsification - On the History of Testing the Homogeneity Condition in the Theory of Consumer Demand
9130	G. Mailath, L. Samuelson and J. Swinkels	Extensive Form Reasoning in Normal Form Games
9131	K. Binmore and L. Samuelson	Evolutionary Stability in Repeated Games Played by Finite Automata
9132	L. Samuelson and J. Zhang	Evolutionary Stability in Asymmetric Games
9133	J. Greenberg and S. Weber	Stable Coalition Structures with Uni-dimensional Set of Alternatives
9134	F. de Jong and F. van der Ploeg	Seigniorage, Taxes, Government Debt and the EMS
9135	E. Bomhoff	Between Price Reform and Privatization - Eastern Europe in Transition
9136	H. Bester and E. Petrakis	The Incentives for Cost Reduction in a Differentiated Industry
9137	L. Mirman, L. Samuelson and E. Schlee	Strategic Information Manipulation in Duopolies
9138	C. Dang	The D^*_2-Triangulation for Continuous Deformation Algorithms to Compute Solutions of Nonlinear Equations
9139	A. de Zeeuw	Comment on "Nash and Stackelberg Solutions in a Differential Game Model of Capitalism"
9140	B. Lockwood	Border Controls and Tax Competition in a Customs Union
9141	C. Fershtman and A. de Zeeuw	Capital Accumulation and Entry Deterrence: A Clarifying Note
9142	J.D. Angrist and G.W. Imbens	Sources of Identifying Information in Evaluation Models
9143	A.K. Bera and A. Ullah	Rao's Score Test in Econometrics
9144	B. Melenberg and A. van Soest	Parametric and Semi-Parametric Modelling of Vacation Expenditures
9145	G. Imbens and T. Lancaster	Efficient Estimation and Stratified Sampling

No.	Author(s)	Title
9146	Th. van de Klundert and S. Smulders	Reconstructing Growth Theory: A Survey
9147	J. Greenberg	On the Sensitivity of Von Neuman and Morgenstern Abstract Stable Sets: The Stable and the Individual Stable Bargaining Set
9148	S. van Wijnbergen	Trade Reform, Policy Uncertainty and the Current Account: A Non-Expected Utility Approach
9149	S. van Wijnbergen	Intertemporal Speculation, Shortages and the Political Economy of Price Reform
9150	G. Koop and M.F.J. Steel	A Decision Theoretic Analysis of the Unit Root Hypothesis Using Mixtures of Elliptical Models
9151	A.P. Barten	Consumer Allocation Models: Choice of Functional Form
9152	R.T. Baillie, T. Bollerslev and M.R. Redfearn	Bear Squeezes, Volatility Spillovers and Speculative Attacks in the Hyperinflation 1920s Foreign Exchange
9153	M.F.J. Steel	Bayesian Inference in Time Series
9154	A.K. Bera and S. Lee	Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis
9155	F. de Jong	A Univariate Analysis of EMS Exchange Rates Using a Target
9156	B. le Blanc	Economies in Transition
9157	A.J.J. Talman	Intersection Theorems on the Unit Simplex and the Simplotope
9158	H. Bester	A Model of Price Advertising and Sales
9159	A. Özcam, G. Judge, A. Bera and T. Yancey	The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model
9160	R.M.W.J. Beetsma	Bands and Statistical Properties of EMS Exchange Rates: A Monte Carlo Investigation of Three Target Zone Models Zone Model
9161	A.M. Lejour and H.A.A. Verbon	Centralized and Decentralized Decision Making on Social Insurance in an Integrated Market Multilateral Institutions
9162	S. Bhattacharya	Sovereign Debt, Creditor-Country Governments, and
9163	H. Bester, A. de Palma, W. Leininger, E.-L. von Thadden and J. Thomas	The Missing Equilibria in Hotelling's Location Game
9164	J. Greenberg	The Stable Value

No.	Author(s)	Title
9165	Q.H. Vuong and W. Wang	Selecting Estimated Models Using Chi-Square Statistics
9166	D.O. Stahl II	Evolution of Smart$_n$ Players
9167	D.O. Stahl II	Strategic Advertising and Pricing with Sequential Buyer Search
9168	T.E. Nijman and F.C. Palm	Recent Developments in Modeling Volatility in Financial Data
9169	G. Asheim	Individual and Collective Time Consistency
9170	H. Carlsson and E. van Damme	Equilibrium Selection in Stag Hunt Games
9201	M. Verbeek and Th. Nijman	Minimum MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross Sections
9202	E. Bomhoff	Monetary Policy and Inflation
9203	J. Quiggin and P. Wakker	The Axiomatic Basis of Anticipated Utility; A Clarification
9204	Th. van de Klundert and S. Smulders	Strategies for Growth in a Macroeconomic Setting
9205	E. Siandra	Money and Specialization in Production
9206	W. Härdle	Applied Nonparametric Models
9207	M. Verbeek and Th. Nijman	Incomplete Panels and Selection Bias: A Survey
9208	W. Härdle and A.B. Tsybakov	How Sensitive Are Average Derivatives?
9209	S. Albæk and P.B. Overgaard	Upstream Pricing and Advertising Signal Downstream Demand
9210	M. Cripps and J. Thomas	Reputation and Commitment in Two-Person Repeated Games
9211	S. Albæk	Endogenous Timing in a Game with Incomplete Information
9212	T.J.A. Storcken and P.H.M. Ruys	Extensions of Choice Behaviour
9213	R.M.W.J. Beetsma and F. van der Ploeg	Exchange Rate Bands and Optimal Monetary Accommodation under a Dirty Float
9214	A. van Soest	Discrete Choice Models of Family Labour Supply
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

No.	Author(s)	Title
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 Hans 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
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

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