Efficient Estimation With Panel Data: An Empirical Comparison of Instrumental Variables Estimators

Christopher Cornwell, Peter Rupert

Journal of Applied Econometrics, Volume 3, Issue 2 (Apr., 1988), 149-155.

Stable URL:
http://links.jstor.org/sici?sici=0883-7252%28198804%293%3A2%3C149%3AEEWPDA%3E2.0.CO%3B2-O

JOURNAL OF APPLIED ECONOMETRICS, VOL. 3, 149–155 (1988)

EFFICIENT ESTIMATION WITH PANEL DATA: AN EMPIRICAL COMPARISON OF INSTRUMENTAL VARIABLES ESTIMATORS

CHRISTOPHER CORNWELL AND PETER RUPERT
Department of Economics, West Virginia University, PO Box 6025, Morgantown, WV26505–6025 U.S.A.

SUMMARY

Hausman and Taylor (1981) proposed an efficient instrumental variables estimator for panel data regression models where the individual effects may be correlated with some of the regressors. Amemiya and MacCurdy (1986) and Breusch, Mizon, and Schmidt (1986) have suggested instrumental variables estimators that are potentially more efficient than the Hausman and Taylor estimator. We address the empirical question of how large these efficiency gains might be. In our returns to schooling example noticeable efficiency gains are limited to the coefficients of time-invariant endogenous variables. This is important, however, since education does not vary over time in our sample.

1. INTRODUCTION

It is well established that with panel data omitted or unobservable individual-specific effects can be controlled for in the specification of an econometric model. Since the effects may be correlated with some of the explanatory variables in the model, the choice of instrumental variables (IV) estimation techniques is natural. The traditional IV estimator is the 'within' estimator from analysis of covariance. It is consistent under the weakest set of assumptions, and it is also simple to compute: just transform the data into deviations from individual means and perform least-squares. However, the within estimator suffers from two significant defects. First, all time-invariant variables are eliminated by the data transformation, so their coefficients cannot be estimated. Second, the within estimator is not fully efficient since it ignores variation across individuals.

Hausman and Taylor (1981)—hereafter HT—propose an IV estimator with neither of these defects. The HT estimator exploits assumptions about which explanatory variables are uncorrelated with the individual effects. The extent to which their estimator represents an improvement over the within estimator depends on the number of exogeneity restrictions one is willing to impose. In general, if there are more time-varying exogenous variables than time-invariant endogenous variables, the HT estimator is consistent and more efficient than the within estimator.

Recently, improvements on the HT estimator have seen suggested. Amemiya and MaCurdy (1986)—hereafter AM—propose an IV estimator that, if consistent, is no less efficient than the HT estimator. Potential efficiency gains are derived from the use of each exogenous time-varying explanatory variable as $(T+1)$ instruments: as deviations from means and separately for each of the T available time periods. The HT estimator uses each such variable as only two instruments: as means and deviations from means.

0883–7252/88/020149–07$05.00

Received June 1987
Revised December 1987

Bruesch, Mizon, and Schmidt (1987)—hereafter BMS—clarify the relationship between HT and AM. In addition, they extend the AM reasoning to derive an even more efficient IV estimator. Implicitly, both HT and AM estimators use the deviations from means of the time-varying exogenous variables as instruments. The BMS estimator uses as additional instruments the $(T-1)$ linearly independent values of these deviations from means.

In this paper we address the empirical question of how large might be the efficiency gain to AM and BMS. Given that the computations involved in these procedures are not trivial, we feel that this is an important question. The framework for our analysis is the returns to schooling problem. This framework provides a special motivation for comparing these procedures, since in previous applications (e.g. Chowdhury and Nickell, 1985) the HT estimator has not yielded very precise schooling coefficient estimates.

We begin in the next section with a brief review of these IV procedures. In section 3 we apply these estimators to a standard wage equation, focusing on the returns to schooling. Our results show the efficiency gain to AMS and BMS is generally very small. An exception, however, is the schooling coefficient, which is estimated much more preceisely by AM and BMS.

2. MODEL AND ESTIMATORS

We consider models of the form

$$y_{it} = X_{it}'\beta + Z_i'\gamma + \alpha_i + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{2.1}$$

where X_{it} is a $K \times 1$ vector of time-varying explanatory variables, Z_i is a $G \times 1$ vector of time-invariant explanatory variables, and β and γ are conformably dimensioned parameter vectors. We assume the disturbances ε_{it} are *iid* $N(0, \sigma_\varepsilon^2)$ and the individual effects α_i are *iid* $N(0, \sigma_\alpha^2)$. The ε_{it} are assumed to be uncorrelated with both the explanatory variables and the effects, while the α_i may be correlated with parts of X and Z.

Combining all NT observations we can write (2.1) as follows:

$$y = X\beta + Z\gamma + V\alpha + \varepsilon, \tag{2.2}$$

where y and ε are $\mathrm{NT} \times 1$, X is $\mathrm{NT} \times K$, Z is $\mathrm{NT} \times G$, and V is an $\mathrm{NT} \times N$ matrix of individual-specific dummy variables. We follow HT and order the observations in N groups of length T. Our notation for projections also follows HT. For any matrix A, we define $P_A = A(A'A)^{-1}A'$ as the projection onto the column space of A. Then, $Q_A = I - P_A$ is defined as the projection onto the null space of A.

The most common IV estimator for models like (2.2) is the within estimator. It is calculated by projecting (2.2) onto the null space of V and performing least-squares. Since $Q_V Z = 0$, only β is estimated, so we have

$$\hat{\beta}_W = (X'Q_V X)^{-1} X'Q_V y, \tag{2.3}$$

The within estimator is consistent (as N or $T \to \infty$) whether or not the effects are correlated with the explanatory variables.

However, if we are willing to assume that parts of X and Z are uncorrelated with the effects, potentially more efficient IV procedures are available. Following HT, we partition X and Z:

$$X = (X_1, X_2),\ Z = (Z_1, Z_2), \tag{2.4}$$

and assume X_2 and Z_2 are correlated with the effects (e.g. plim $(\mathrm{NT})^{-1}\ X_2 V\alpha \neq 0$), while X_1

and Z_1 are not correlated with the effects. Note that X_1 has k_1 columns, X_2 has k_2 columns, and $k_1 + k_2 = K$; Z_1 has g_1 columns and Z_2 has g_2 columns, and $g_1 + g_2 = G$.

The efficient estimators of HT, AM, and BMS are calculated in a similar fashion. First, (2.2) is transformed so that the error term will have a scalar covariance matrix. Defining $\Omega \equiv \text{cov}(V\alpha + \varepsilon)$, the transformed model is

$$\Omega^{-1/2}y = \Omega^{-1/2}X\beta + \Omega^{-1/2}Z\gamma + \Omega^{-1/2}(V\alpha + \varepsilon). \tag{2.5}$$

where $\Omega^{-1/2} = Q_V + \theta P_V$ and $\theta^2 = \sigma_\varepsilon^2(\sigma_\varepsilon^2 + T\sigma_\alpha^2)^{-1}$. Then, with an instrument set A based on (2.4), IV is performed on (2.5). This yields estimators of the form

$$\begin{pmatrix}\hat{\beta}\\ \hat{\gamma}\end{pmatrix} = [(X, Z)'\Omega^{-1/2}P_A\Omega^{-1/2}(X, Z)]^{-1}(X, Z)'\Omega^{-1/2}P_A\Omega^{-1/2}y. \tag{2.6}$$

The HT estimator uses the instrument set[1]

$$A_1 = (Q_VX_1, Q_VX_2, P_VX_1, Z_1). \tag{2.7}$$

Note that each variable in X_1 provides two instruments since the means (P_VX_1) and deviations from means (Q_V, X_1) are used separately. The order condition for the HT estimator to exist is $K + k_1 + g_1 \geqslant K + G$, or $k_1 \geqslant g_2$.

To define the instrument set used by AM, let X_1^* be an NT × TK matrix where each column contains values of X_{1it} for a *single time period*. For example, the tth column of X_1^* is given by $X_{1t}^* = (X_{11t}, \dots, X_{11t}, \dots, X_{1Nt}, \dots, X_{1Nt})$. So, the AM estimator is IV on (2.5) using the set of instruments

$$A_2 = (Q_VX_1, Q_VX_2, X_1^*, Z_1). \tag{2.8}$$

While HT use each X_1 variable as two instruments, AM use each of these variables as $(T + 1)$ instruments (Q_VX_1 and X_1^*). The AM order condition for existence is $Tk_1 \geqslant g_2$.

The AM estimator, if consistent, is no less efficient than the HT estimator. However, in this case, consistency depends on a stronger exogeneity assumption. Since plim $(\text{NT})^{-1}X_1'V\alpha = 0$ implies plim $(\text{N})^{-1}\sum_{i-1}^{N}\bar{X}_{1i}'\alpha_i = 0$, HT only require the means of the X_1 variables to be uncorrelated with the effects. For the AM estimator to be consistent we need plim $(N)^{-1}\sum_{i=1}^{N} X_{1it}'\alpha_i = 0$ $(t = 1, \dots, T)$, or uncorrelatedness at each point in time. As AM and BMS suggest, it is difficult to imagine a situation in which the HT assumption is true, but the AM assumption is not.

BMS derive a potentially more efficient AM-like estimator. Noting that the AM estimator is equivalent to IV of (2.5) on

$$A_2 = (Q_VX_1, Q_VX_2, P_VX_1, (Q_VX_1)^*, Z_1), \tag{2.9}$$

where $(Q_VX_1)^*$ is defined in the same way as X_1^* (i.e. each column contains the deviations from means of the X_1 variables for a *single time period*), BMS extend the AM treatment of the X_1 variables to the X_2 variables. The BMS estimator uses the instrument set

$$A_3 = (Q_VX_1, Q_VX_2, P_VX_1, (Q_VX_1)^*, (Q_VX_2)^*, Z_1), \tag{2.10}$$

with $(Q_VX_2)^*$ defined exactly like $(Q_VX_1)^*$. So, the BMS estimator exists when $Tk_1 + (T - 1)k_2 \geqslant g_2$.

[1]We follow BMS in our definition of the instrument set used in the HT procedure. See BMS for the details regarding the relationship between the HT, AM, and BMS estimators.

The potential efficiency gain from the BMS procedure depends on whether the $(Q_V X_2)^*$ are legitimate instruments. The $(Q_V X_2)^*$ are valid instruments if the variables in X_2 are correlated with the effects only through a time-invariant component. If this were true, $Q_V X_2$ would not contain this component, and using the deviations sepearately for each time period would be legitimate.

3. ESTIMATOR COMPARISON: MEASURING THE RETURNS TO SCHOOLING

When the parameters of the model are overidentified the HT estimator represents a distinct improvement over the within estimator, since more efficient estimates of β and consistent estimates of γ are possible. Potentially, even more efficient estimates of β and γ can be obtained with the AM and BMS procedures. How large the efficiency gains will be is an empirical question. Both AM and BMS anticipate small gains in many applications.

Here we investigate the potential efficiency gains from using these more refined IV procedures to estimate the returns to schooling. Recent additions to the returns to schooling literature (e.g. Griliches, 1977; Lillard and Willis, 1978; HT, and Chowdhury and Nickell, 1985) have been concerned with the potential correlation between individual ability and education. Typically, ability is not observed, leading to the natural conclusion that the individual effects are correlated with education.[2] One argument states that education is positively correlated with the effects (ability). In this case, the OLS (or GLS) estimate is biased upward. However, as Griliches (1977) and Griliches, Hall, and Hausman (1978) have shown, when the amount of schooling is made endogenous, a negative correlation between the effects and education can arise. These results are based on traditional IV procedures which define a reduced form for education in terms of excluded exogenous variables like family background characteristics. Higher returns to schooling estimates are also reported by HT. We look for further evidence in our results of a negative correlation between education and the effects.

The data for our analysis are drawn from years 1976–1982 of the non-Survey of Economic Opportunity portion of the Panel Study of Income Dynamics (PSID). The individuals in our sample are 595 heads of household between the ages of 18 and 65 in 1976, who report a positive wage in some private, non-farm employment for all 7 years. So, for each individual, we have seven annual observations on the following wage-determining characteristics: years of education (ED), years of full-time work experience (EXP), weeks worked (WKS), occupation (OCC = 1, if the individual has blue-collar occupation), industry (IND = 1, if the individual works in a manufacturing industry), residence (SOUTH = 1, SMSA = 1 if the individual resides in the south, or in a standard metropolitan statistical area), marital status (MS = 1, if the individual is married), union coverage (UNION = 1, if the individual's wage is set by a union contract), sex and race (FEM = 1, BLK = 1, if the individual is female or black).[3]

Before we apply the HT, AM, or BMS estimator to our wage equation, we consider the results from conventional procedures, which are presented in the first two columns of Table I.[4]

---

[2]Some datasets—for example, the National Longitudinal Survey's (NLS) young men cohort—contain an IQ measure. However, using such variables can create sample selection problems (see Griliches, Hall, and Hausman, 1978). In this paper we use the Panel Study of Income Dynamics, which does not contain an IQ measure.

[3]Our regressions also contain year dummies to capture productivity and price level effects.

[4]Note that we do not treat serial correlation apart from the individual effects as others have (e.g., Lillard and Willis, 1978 and Chowdhury and Nickell, 1985). Since we are not interested in wage dynamics, we feel little is lost in considering the simpler case.

Table I. Dependent variable: log wage

	GLS	Within	HT	AM	BMS
Constant	0·2207		0·1492	0·1761	0·2093
	(0·0115)		(0·0500)	(0·0376)	(0·0203)
EXP	0·0565	0·0605	0·0605	0·0602	0·0601
	(0·0028)	(0·0028)	(0·0029)	(0·0029)	(0·0029)
EXP^2	−0·00045	−0·00042	−0·00042	−0·00042	−0·00042
	(0·00005)	(0·00005)	(0·00005)	(0·00005)	(0·00005)
WKS	0·00088	0·00086	0·00089	0·00089	0·00088
	(0·00058)	(0·00058)	(0·00059)	(0·00059)	(0·00058)
OCC	−0·0308)	−0·0291	−0·0291	−0·0293	−0·0295
	(0·0133)	(0·0132)	(0·0134)	(0·0134)	(0·0133)
IND	0·0225	0·0199	0·0200	0·0200	0·0211
	(0·0148)	(0·0148)	(0·0150)	(0·0149)	(0·0149)
SOUTH	0·0227	0·0263	0·0272	0·0244	0·0208
	(0·0320)	(0·0333)	(0·0327)	(0·0324)	(0·0321)
SMSA	−0·0321	−0·0359	−0·0356	−0·0374	−0·0315
	(0·0185)	(0·0186)	(0·0189)	(0·0187)	(0·0185)
MS	−0·0436	−0·0483	−0·0487	−0·0487	−0·0488
	(0·0182)	(0·0181)	(0·0183)	(0·0183)	(0·0182)
UNION	0·0197	0·0141	0·0134	0·0140	0·0156
	(0·0143)	(0·0143)	(0·0145)	(0·0144)	(0·0143)
ED	0·0941		0·2004	0·1585	0·1067
	(0·0170)		(0·0783)	(0·0586)	(0·0314)
FEM	−0·3482		−0·3559	−0·3484	−0·3385
	(0·1515)		(0·1531)	(0·1523)	(0·1518)
BLK	−0·1936		−0·0660	−0·1211	−0·1898
	(0·1850)		(0·2122)	(0·2000)	(0·1885)
S.E.R.	0·148	0·146	0·149	0·148	0·148
		$\chi^2_9 = 69{\cdot}47$	$\chi^2_3 = 1{\cdot}36$	$\chi^2_{13} = 3{\cdot}76$	$\chi^2_{13} = 5{\cdot}99$

X_1 = (WKS, SOUTH, SMSA, MS)'
Z_1 = (FEM, BLK)

In the first column we give the coefficient estimates from our GLS regression.[5] These estimates appear to be within the range of results from other GLS wage regressions on PSID or other data. In particular, the coefficient on education is estimated to be 0·094, suggesting an additional year of schooling produces a 9·4 per cent wage gain. With the exception of the coefficients on WKS, IND, SOUTH, and UNION, the parameter estimates seem reasonably precise. But GLS assumes that the effects are uncorrelated with the explanatory variables. If this assumption is at variance with the data, then the GLS estimates are biased and inconsistent.

The within estimates, given in the second column of table 1, are unbiased and consistent whether or not the effects are correlated with the explanatory variables. Note that all time-invariant variables have been eliminated by the data transformation. The remaining coefficient estimates differ somewhat from the corresponding GLS estimates; e.g. the coefficient on EXP rises by 7 per cent, the coefficient on SOUTH rises by 15 per cent, and the coefficient on UNION

[5]GLS is easily calculated by performing least-squares on (2.5) since the $\Omega^{-1/2}$ transformation is just a '$(1-\theta)$ differencing' of the data. Consistent estimates of the variance components are $\hat{\sigma}^2_\varepsilon = 0{\cdot}025$, $\hat{\sigma}^2_\alpha = 1{\cdot}453$, which implies $\hat{\theta} = 0{\cdot}049$. The variance component associated with ε is calculated from the sums of squares of the within residuals. The individual variance component is derived from the residuals incorporating the within estimate of β and the HT simple consistent estimate of γ (see HT, p. 1384).

declines by 33 per cent.[6] The differences suggest that the assumption of no correlation between the α_i and (X, Z) in the GLS regression may be incorrect. The null hypothesis, that the effects are uncorrelated with the explanatory variables, is decisively rejected in a Hausman test of the difference between the GLS and within estimates. The test statistic, which is a χ_9^2 random variable under H_0, is 69·47.

As stated, although the within estimator is generally consistent, it has other drawbacks. In the present case, that γ cannot be estimated is an important problem since our focus is on the returns to schooling. In addition, the within estimates may not be fully efficient. The HT, AM, and BMS estimators address both of these problems. Which estimator to use should be determined on efficiency grounds. If their instrument sets are valid, AM and BMS can be no less efficient than HT. So, we apply HT, AM, and BMS to our wage equation and examine their relative efficiencies. Since previous attempts at applying HT to similar wage equations have yielded imprecise schooling coefficient estimates, evidence of a sizeable efficiency gain to AM and/or BMS will be of some interest.

Our basic HT, AM, and BMS regressions are presented in columns 3, 4, and 5 of Table I. In these regressions we let X_1 = (WKS, SOUTH, SMSA, MS), X_2 = (EXP, EXP^2, OCC, IND, UNION), Z_1 = (FEM, BLK), and Z_2 = (ED). Consider the HT estimates given in the third column. Two things are immediately apparent. First, the coefficeint on education is estimated to be 0·2, a 75 per cent increase over the GLS estimate result. Second, the other estimated coefficients generally reflect the values of the within estimates, having moved away from the GLS estimates. The closeness of the HT and within estimates suggests that our initial instrument set, noted at the bottom of Table I, is legitimate. This is supported by a Hausman test of the difference between the within and HT estimates. The test-statistic is 1·36 and is distributed as χ_3^2 under the null hypothesis, so we cannot reject X_1 and Z_1 as valid instruments.

HT use the means and deviations from means of the X_1 variables as instruments. In addition to their deviations from means, AM use the T separate realizations of the X_1 variables as instruments. The AM results are given in the fourth column of Table 1. In general, the AM estimates resemble the within and HT estimates. Standard errors are very close to those produced by the HT estimator. An exception is the schooling coefficient estimate which has dropped 23 per cent from the HT estimate to 0·158, with a considerably lower standard error. The additional exogeneity restrictions employed by AM are not rejected in a Hausman test comparing AM to HT. The test statistic is 3·76 and is distributed as χ_{13}^2 under H_0.

BMS add to the instrument set of HT and AM the $(T-1)$ linearly independent values of the deviations from means of the X_2 variables.[7] The BMS estimates, presented in column 5, follow the pattern of the AM estimates. Like AM, the BMS results are generally similar to HT (and within). Again, the schooling coefficient is the notable exception with its estimate falling to 0·107, 63 per cent lower than the HT estimate. Furthermore, its standard error has dropped from 0·078 in the HT regression to 0·031 in the BMS regression. These changes imply a rise in the t-statistic from 2·56 to 3·40. It appears that, if the variables in X_2 are correlated with the effects, they are only correlated through a time-invariant component. A Hausman test comparing the BMS and AM estimates suggests we cannot reject $(Q_V X_2)^*$ as valid instruments. The test statistic, which is distributed as χ_{13}^2 under H_0, is 5·99.

The impact of the AM and BMS procedures on the returns to schooling can be explained as follows. The extra instruments employed by AM and BMS, X_2^* and $(Q_V X_2)^*$, are time-

[6]Percentage changes are calculated by differences in natural logarithms.

[7]Since using the deviations from means of EXP and EXP^2 cause singularity in the first stage regression, they are omitted.

invariant in the sense that they are orthogonal to Q_V. Therefore we should expect to see their effect primarily on time-invariant endogenous variables. Education is the only such variable in our wage equation, and its estimated coefficient and the estimate's standard error are sharply affected. However, for the remaining coefficient estimates there is little change, and the efficiency gain to AM or BMS is quite small.

A final comment contrasts the results here with those presented in the HT returns to schooling example. HT report that, for all specifications to which their estimator is applied, the estimated schooling coefficient increases over any estimate that does not control for correlation between the explanatory variables and the effects. The same is true in our example; the HT, AM, and BMS regression each yield estimated schooling coefficients that exceed the GLS estimate. This is at variance with conventional stories about ability bias, and represents further evidence of a negative correlation between education and ability when education is allowed to be endogenous.

4. SUMMARY

In this paper we have made an empirical comparison of the IV estimators proposed by HT, AM, and BMS. The framework for our comparison has been the measurement of the returns to schooling. Using data drawn from the years 1976–1982 of the non-Survey of Economic Opportunity portion of the Panel Study of Income Dynamics, we estimate a standard wage equation, focusing on the schooling coefficient. We find that noticeable efficiency gains to the methods put forward by AM and BMS are limited to the estimated coefficients on the time-invariant variables. However, in the present case this is meaningful, since education does not vary over time in our sample. Indeed, the standard errors of the schooling coefficient estimates change markedly—by as much as 90 per cent—when AM or BMS is applied. The corresponding t-statistics rise accordingly. The impact of the AM and BMS estimators falls primarily on time-invariant endogenous variables (like education) because the extra instruments employed by these methods are time-invariant.

ACKNOWLEDGEMENTS

We gratefully acknowledge the helpful comments of Peter Schmidt, Adrian Pagan, and an anonymous referee.

REFERENCES

Amemiya, T. and T. E. MacCurdy, (1986), 'Instrumental variables estimation of an error components model', *Econometrica,* **54,** 869–880.

Breusch, T. S., G. E. Mizon and P. Schmidt, (1986), 'Some results on panel data'. Econometrics Workshop Paper no. 8608. Michigan State University.

Chowdhury, G. and S. Nickell (1985), 'Hourly earnings in the U.S.: another look at unionization, schooling, sickness, and unemployment using PSID data," *Journal of Labor Economics,* **3,** 38–69.

Griliches, Z. (1977), 'Estimating the returns to schooling: some econometric problems', *Econometrica,* **45,** 1–22.

Griliches, Z., B. Hall and J. A. Hausman (1978), 'Missing data and self-selection in large panels', *Annales de l'Insee,* **30/31,** 137–176.

Hausman, J. A. (1978), 'Specification test in econometrics', *Econometrica,* **46,** 1251–1272.

Hausman, J. A. and W. Taylor, (1981), 'Panel data and unobservable individual effects', *Econometrica,* **49,** 1377–1399.

Lillard, L. and R. Willis, (1978), 'Dynamic aspects of earnings mobility', *Econometrica,* **46,** 985–1012.