NBER WORKING PAPER SERIES

WHY DO FIXED-EFFECTS MODELS PERFORM SO POORLY? THE CASE OF ACADEMIC SALARIES

Daniel S. Hamermesh

Working Paper No. 2135

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1987

I am indebted to Eric Weissend for very capable research assistance on this project, and to Ronald Ehrenberg and Peter Schmidt for helpful comments. The data used in this study are available on diskette from the author. The research reported here is part of the NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #2135
January 1987

Why Do Fixed-Effects Models Perform So Poorly?
The Case of Academic Salaries

ABSTRACT

A large and growing line of research has used longitudinal data to eliminate unobservable individual effects that may bias cross-section parameter estimates. The resulting estimates, though unbiased, are generally quite imprecise. This study shows that the imprecision can arise from the measurement error that commonly exists in the data used to represent the dependent variable in these studies. The example of economists' salaries, which are administrative data free of measurement error, demonstrates that estimates based on changes in longitudinal data can be precise. The results indicate the importance of improving the measurement of the variables to which the increasingly high-powered techniques designed to analyze panel data are applied. The estimates also indicate that the payoff to citations to scholarly work is not an artifact of unmeasured individual effects that could be biasing previous estimates of the determinants of academic salaries.

Daniel S. Hamermesh
Department of Economics
Michigan State University
East Lansing, MI 48824
(517) 355-5238

I. Introduction

Many studies of wage determination have estimated models involving fixed individual effects using longitudinal data. Time-series methods and "first-differencing" circumvent the unmeasurability of the fixed effects and produce estimates of the structural parameters that accord with prior expectations. However, the explanatory power of the equations is often remarkably low relative to that of the corresponding cross-section equations. For example, the $\bar{R}^2$ in Mincer (1983) are around .4 in cross-section equations, but less than .02 in equations that are effectively ten-year differences. Equally great differences are present in Mellow (1981) using the same (Panel Study of Income Dynamics) data but taking only one-year differences. A similar difference is produced in Stafford-Duncan (1980) using another household survey.

The discrepancies between the $\bar{R}^2$ in equations based on changes and those based on levels are not quite so huge in other studies. Nonetheless, substantial differences also exist in work by Lazear (1976), Brown (1980) and Duncan-Holmlund (1983). Presumably, similar differences exist outside the realm of labor economics (though with the exception of Holtz-Eakin et al, 1985, the techniques have not been widely applied in other areas). The discrepancies indicate no bias in the parameter estimates. They do, though, coincide with an imprecision in the estimates that prevents one from making strong statements about the true values of the parameters.

A number of authors (Griliches-Hausman, 1986; Freeman, 1984; Chowdhury-Nickell, 1985; and Jakubson, 1986) have recently discussed how measurement error in an **independent variable** produces biases when differencing methods are used on longitudinal data. In Section II I show how measurment error in the dependent variable can cause the imprecise, but unbiased estimates noted in the empirical studies discussed above. I then estimate a model involving fixed individual effects in a new set of data collected for this purpose. Using these data (covering the salaries of a group of academics), the estimated $\bar{R}^2$ and the standard errors of the estimated parameters differ little between cross-section and longitudinal models.

II. Fixed-Effects Models with Noisy Dependent Variables

Define the model as:

$$y^*_{it} = \beta X_{it} + \phi_i + \epsilon_{it}, \tag{1}$$

where β is the vector of parameters, and X is a vector of variables observed at time t, which could include separate constant terms for each t. y^* is the true value of the dependent variable, ϕ is a vector of unmeasured person-specific effects, and ϵ is an i.i.d. disturbance. Let y_{it} be defined as:

$$y_{it} = y^*_{it} + \theta_{it}, \tag{2}$$

where $E(\theta_{it}) = 0$; $E(\theta^2_{it}) = \sigma^2_\theta$, and $E(\theta_{it}\theta_{it-1}) = \rho\sigma^2_\theta$. The formulation of y_{it} in (2) specifies the wage as being measured with an autocorrelated error.

Using (1) and (2) the variance of the error in (1) estimated over levels is:

$$\sigma^2_\phi + \sigma^2_\epsilon + \sigma^2_\theta ,$$

while that estimated over changes is:

$$2\{\sigma_{\epsilon}^2 + \sigma_{\theta}^2[1-\rho]\} \ .$$

Let the variance of y_{it} be σ_y^2, and let the autocorrelation of y_{it} be defined as ρ', so that $E(y_{it}y_{it-1}) = \rho'\sigma_y^2$. Then when (1) is estimated over levels:

$$R^2 = 1 - [\sigma_{\phi}^2 + \sigma_{\epsilon}^2 + \sigma_{\theta}^2]/\sigma_y^2.$$

When (1) is estimated over changes:

$$R^2 = 1 - 2\{\sigma_{\epsilon}^2 + \sigma_{\theta}^2[1-\rho]\}/\{\sigma_y^2[1-\rho']\} \ .$$

The difference between these coefficients of determination is:

$$\text{(3)} \qquad D = \{\sigma_{\theta}^2[\rho' - \rho] + \rho'\sigma_{\epsilon}^2 - \sigma_{\phi}^2[1-\rho']\}/\{\sigma_y^2[1-\rho']\} \ .$$

If autocorrelation in the y exceeds that in the measurement error, greater measurement errors increase D. We have no way of estimating either ρ or ρ' in general. However, in their validation study of the Panel Study of Income Dynamics Duncan-Hill (1985) find the highest one-year autocorrelation of the measurement error in earnings to be .43. Calculations on their data show that the one-year autocorrelation of the logarithm of measured earnings is .76.[1] They also show that the variance ratio of measurement error to the true value of the logarithm of annual earnings is around .5. Taken together, these findings and observations suggest that it is unsurprising that the literature is replete with studies in which D is very large: Wage and earnings data in the PSID and the other large household surveys that underlie the studies discussed in Section I are characterized by substantial measurement error that is not very highly autocorrelated.

III. Longitudinal Estimates of the Determinants of Academic Salaries

To demonstrate that the estimation of models to account for fixed effects by taking deviations around means need not lead to imprecise estimates, consider the example of academic salaries. The general model specifies the dependent variable, the logarithm of the real compensation of the i'th faculty member at time t, as:

(1') $$y_{it} = \beta X_{it} + \phi_i + \epsilon_{it} .$$

I estimate (1') in a variety of ways:

1. OLS estimates based on cross-sections of data at two points in time.

2. OLS estimates based on deviations of the variables from their means over time (the "within" estimator), the same estimator used in the studies referenced in Section I.

3. GLS estimates (Judge et al, 1985, pp. 522, passim.), essentially a weighted average of the data used to produce "within" and "between" OLS estimates of (1'). This technique has not been widely used in the literature on fixed effects in the determination of wages. It can produce more efficient parameter estimates.

In Hamermesh et al (1982) we examined the determinants of the 1979-80 academic-year salaries of 148 full professors of economics in 7 large public universities. Variables included in X were the elapsed time since the individual obtained the Ph.D., EXP; the average number of citations by others in the previous five years (from the Social Science Citation Index), CIT; and a dummy variable indicating whether the individual had been or currently was an administrator at or above the level of department chair, AD. Additional data were obtained from 6 of the 7 schools on the additional administrative experience and the salary in 1985-86 of the 100 full professors still on their faculties. With additional data on citations these formed the basis for a 1985-86 cross section. It and the appropriate subset of the original data set form a longitudinal data file on these 100 individuals.

The data on salaries were transformed to yield a measure of compensation in real terms (that removes cross-section differences in living costs).[2] The means and standard deviations of the means of the logarithm of this measure and of CIT are shown in Table 1 for the 6 schools separately and for the pooled sample of 100 observations.[3] The average rate of citations increased in all 6 schools over the six-year period. The standard deviations of means of the citations data are huge, partly because the distribution of citations across observations is highly skewed. Examining the autocorrelations of each variable (the $r_{79,85}$), we see that there is very strong persistence over time in rates of citation, and somewhat less persistence in salaries.

The first two panels of Table 2 present estimates of (1') based on the two cross sections of data, both for each of the 6 schools separately and for the pooled set of data. Comparing the results across the two years there is no overall structural change (other than a shift in the constant term).[4] Tests for the inclusion of school-specific dummy variables in the pooled data also yielded test statistics that were not significant (1.00 and 1.78 for 1979-80 and 1985-86, each distributed $F(5, 91)$), and the parameter estimates differed little from those presented in the Table.

Each additional year of experience raises pay by between ½ and 1 percent even in this relatively homogeneous sample. Administrative experience raises base salaries by over 10 percent. Whether this is a compensating differential for time lost from research that

Table 1

Descriptive Data, Salaries and Citations [a]

Log Real Compensation	Pooled	School Number 1	2	3	4	5	6
1979	10.67 (.17)	10.68 (.22)	10.76 (.22)	10.63 (.13)	10.66 (.07)	10.69 (.14)	10.64 (.14)
1985	11.13 (.19)	11.09 (.26)	11.22 (.22)	11.11 (.14)	11.23 (.12)	11.13 (.17)	11.10 (.15)
$r_{79,85}$	.76	.93	.63	.49	.74	.85	.75
Citations							
1975-79	20.56 (24.58)	31.28 (34.29)	30.94 (28.90)	18.74 (18.06)	28.98 (33.64)	12.21 (10.10)	7.27 (7.88)
1981-85	28.42 (40.40)	33.50 (32.86)	53.00 (70.97)	27.38 (27.92)	41.42 (61.80)	16.18 (16.04)	11.62 (14.63)
$r_{79,85}$	.86	.87	.93	.88	.97	.82	.57
N =	100	20	13	19	11	17	20

[a] Standard deviations of the means in parentheses.

Table 2

Cross-Sections and Changes, Equation (1) [a]

		School Number					
1979-80	Pooled	1	2	3	4	5	6
EXP	.0109	.0090	.0148	.0100	.0049	.0101	.0088
	(5.56)	(1.61)	(3.37)	(2.53)	(.90)	(2.80)	(1.72)
CIT	.0032	.0045	.0045	.0039	.0012	.0086	.0036
	(6.36)	(4.96)	(3.18)	(3.12)	(1.94)	(3.07)	(.83)
AD	.103	.253	.202	.229	---	.074	-.054
	(2.99)	(4.96)	(1.82)	(2.33)		(1.02)	(-.48)
$\bar{R}^2$	.449	.681	.613	.445	.209	.444	.110
1985-86							
EXP	.0037	.0078	.0055	-.0036	-.0035	.0136	-.001
	(1.52)	(1.15)	(.72)	(-.63)	(-.44)	(2.64)	(-.15)
CIT	.0025	.0057	.0018	.0023	.0013	.0088	.0058
	(5.94)	(5.06)	(1.83)	(2.07)	(2.66)	(3.96)	(2.58)
AD	.091	.313	.109	.107	---	.154	-.033
	(2.35)	(4.04)	(.63)	(1.31)		(2.10)	(-.48)
$\bar{R}^2$	.256	.616	.034	.248	.354	.481	.187
Changes							
CIT	.0022	.0014	.0019	.0047	.0013	.0065	.0009
	(4.47)	(1.13)	(1.74)	(2.47)	(1.68)	(4.99)	(.46)
AD	.125	.148	---	.088	---	.093	.163
	(3.07)	(2.23)		(.99)		(2.46)	(2.17)
$\bar{R}^2$	.212	.171	.144	.219	.154	.748	.127

---

[a] t-statistics in parentheses here and in Table 3.

would raise other productivity-enhancing characteristics, or whether it stems from rewards being paid to administrative activities per se, is not knowable from this analysis. Scholarly achievement, as indicated by citations of one's work by others, has a substantial impact on salaries: Each extra citation per annum raises salary by about .3 percent. The coefficient implies a difference in salary of 6.3 percent for a change equal to the interquartile range of CIT.

The $\bar{R}^2$ in these samples are fairly large in most cases, and in the pooled sample (whose size begins to approach those of the smaller data sets used in other studies of wages) it is quite typical. However, as a comparison to the third panel of Table 2 shows, the $\bar{R}^2$ are not substantially larger than those that are produced when (1') is estimated over changes between the two cross sections.[5] In this panel the equations fit nearly as well as in the estimates over levels of the variables. The coefficient estimates are in general signficantly nonzero, and they do not differ greatly from the cross-section coefficients. In particular, the similarity of the estimated effect of CIT in the data on changes demonstrates the substantive point that previous estimates of its importance in academic salary determination reflect more than its correlation with unobserved unchanging individual characteristics.

The sample does not include the 37 professors who were in the universities in 1979-80 but not in 1985-86. It is possible that these people systematically self-selected out of their old jobs in a way that makes the estimates of (1') inconsistent (Solon, 1986). I cannot rule out this possibility; however, the only significant predictor of whether a professor dropped out of the sample was EXP;

and of the 37 drop-outs, 25 either retired or died. Moreover, the hypothesis that the structure of (1') for these people differs from that characterizing the 100 people who remained in the sample cannot be rejected at the 95-percent level of confidence once a separate constant for the drop-outs is included in (1').[6]

The relatively high values of $\bar{R}^2$ obtained in the data on changes do not arise because there are no individual effects. The test-statistic proposed by Breusch-Pagan (1980), which is distributed $\chi^2(1)$, attains values of 7.06, 15.20 7.45, 13.00, 10.49, 6.23, 6.07, for the pooled sample and the six school samples respectively. Each of these is significantly different from zero at least at the 98 percent level of confidence. This suggests that the error component ϕ has nonzero variance.

Table 3 presents GLS estimates of (1'). These are essentially weighted averages of the data used to produce "within" and "between" estimators of β, with a weight, α, for the "within" estimator equalling one minus the square root of the ratio of the estimated variances from the "within" and unrestricted equations. In the estimates for three of the schools the weighting coefficient was negative, rendering the technique inapplicable. In the pooled sample and in two other schools the coefficient $\hat{\alpha}$ listed in Table 3 is small. Only in School 5 are the weights on the "within" and "between" estimates roughly the same.

The GLS estimates differ from the cross-section estimates in that they add a dummy variable, D1, that allows for a separate constant term for each cross section. Aside from this, though, they are very similar to the estimates listed in Table 2, both to

Table 3

GLS Estimates, Equation (1) [a]

	Pooled	School Number 1	5	6
EXP	.0080	.0082	.0207	.0041
	(4.79)	(1.80)	(2.88)	(1.13)
CIT	.0029	.0048	.0132	.0048
	(8.60)	(6.67)	(4.19)	(2.48)
AD	.122	.267	.120	-.026
	(4.57)	(5.36)	(1.27)	(-.51)
D1	-.384	-.332	-.250	-.410
	(-18.47)	(-7.62)	(-3.79)	(-9.17)
$\bar{R}^2$	.801	.846	.791	.787
$\hat{\alpha}$	.155	.279	.570	.129

[a] $\hat{\alpha} < 0$ for schools 2, 3 and 4.

the cross sections and the changes. The similar results produced by all the changes and GLS estimation procedures suggest that the simple cross-section estimates produced here and in Hamermesh et al (1982) are not merely artifacts of a failure to use additional information provided by panel data.

Unlike the numerous other studies that have generated fairly precise estimates of cross-section wage equations and very imprecise estimates of the same equations estimated over changes, this study produces estimates whose precision differs little when data on changes are used. **Why are these results different from other results?** In this study measurement error is small or absent: The data are from administrative records obtained directly from the individuals who set the salaries of the people included in the sample. That being so, it is not surprising that the differences in the adjusted R^2 between the levels and changes equations are much smaller here than in other studies.[7] That the R^2 are somewhat higher in the levels equations than in those estimated over (six-year) first differences is due partly to the high autocorrelation in the (true) earnings measures used here: The correlation of 1979-80 log real compensation with its value in 1985-86 is .76.

IV. Conclusions

Under reasonable assumptions about underlying correlations and error variances, measurement error in the dependent variable is likely to lead to relatively imprecise parameter estimates in equations estimated by differencing out the fixed effects. In the equations estimated here no such imprecision exists because measurement error is absent from the dependent variable. These considerations suggest we face a difficult choice: Either we obtain data that are more appropriate for use in fixed-effects models; or we recognize that the nature of the data underlying most studies makes the application of standard techniques designed to handle fixed individual effects questionable.

The application demonstrates a clear relationship between a direct measure of academic productivity --- citations by other academics --- and the economic rewards to academics. Previous work has indicated that scholars whose work is more widely cited are better paid. This study shows that the higher pay does not stem from unmeasured characteristics that are correlated with greater attention to a scholar's research. Instead, increases in a scholar's professional recognition produce increases in his or her compensation. In at least one segment of academe pay reflects performance, as measured by the impact of one's ideas.

REFERENCES

Trevor Breusch and Adrian Pagan, "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," Review of Economic Studies, 47 (April 1980), 239-253.

Charles Brown, "Equalizing Differences in the Labor Market," Quarterly Journal of Economics, 94 (February 1980), 113-134.

Gopa Chowdhury and Stephen Nickell, "Hourly Earnings in the United States," Journal of Labor Economics, 3 (1985), 38-69.

Greg Duncan and Daniel Hill, "An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data," Journal of Labor Economics, 3 (October 1985), 508-532.

----------- and Bertil Holmlund, "Was Adam Smith Right After All? Another Test of the Theory of Compensating Wage Differentials," Journal of Labor Economics, 1 (October 1983), 366-379.

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

Zvi Griliches and Jerry Hausman, "Errors in Variables in Panel Data," Journal of Econometrics, 31 (1986), 93-118.

Daniel Hamermesh, George Johnson and Burton Weisbrod, "Scholarship, Citations and Salaries: Economic Rewards in Economics," Southern Economic Journal, 49 (October 1982), 472-481.

Douglas Holtz-Eakin, Whitney Newey and Harvey Rosen, "Implementing Causality Tests with Panel Data, with an Example from Local Public Finance," National Bureau of Economic Research, Technical Working Paper No. 48, June 1985.

George Jakubson, "Measurement Error in Binary Explanatory Variables in Panel Data Models," Princeton University, Industrial Relations Section, Working Paper No. 209, August 1986.

George Judge, W. E. Griffiths, R. Carter Hill, Helmut Lutkepohl and Tsoung-Chao Lee, The Theory and Practice of Economics. New York: John Wiley and Sons, 1985.

Edward Lazear, "Age, Experience and Wage Growth," American Economic Review, 66 (September 1976), 548-558.

Wesley Mellow, "Unionism and Wages: A Longitudinal Analysis," Review of Economics and Statistics, 63 (February 1981), 43-52.

Jacob Mincer, ""Union Effects: Wages, Turnover and Job Training," Research in Labor Economics, Supplement 2 (1983).

Gary Solon, "Bias in Longitudinal Estimation of Wage Gaps," National Bureau of Economic Research, Technical Working Paper No. 58, June 1986.

Frank Stafford and Greg Duncan, "Do Union Members Receive Compensating Wage Differentials?" American Economic Review, 70 (June 1980), 355-371.

W. Kip Viscusi and Charles O'Connor, "Adaptive Responses to Chemical Labelling," American Economic Review, 74 (December 1984), 942-956.

</ref_section>

FOOTNOTES

1. The additional calculations were kindly provided by Greg Duncan.

2. If the campus is in a major metropolitan area, each observation was dividied by COL[1 - FR], where COL is the ratio of the medium-budget cost of living in the area and FR is the ratio of fringe benefits to salary. The former is from Bureau of Labor Statistics, Handbook, Bulletin 2070, Table 155; the latter are from annual reports on the economic status of the profession, from the AAUP Bulletin for 1979-80, and from Academe, for 1985-86. The range of the fringe/salary ratio is from .15 to .24 in 1979-80, and from .11 to .25 in 1985-86.

3. The schools are numbered exactly as in Hamermesh et al (1982).

4. F-tests for structural change (other than shifts of the intercepts) yielded the following test statistics for the pooled samples and the 6 schools separately: 2.09, F(3, 192); .27, F(3,26); 1.45, F(3,30); .30, F(3,32); .58, F(3,32); .83, F(3,18); and .23, F(3,14). None of these is significant at usually applied critical levels. However, test-statistics for changes in specific parameters were significant for the coefficient on EXP in the pooled sample and in the sample covering School 3. That the effect of EXP on compensation is smaller in the second period is consistent with the observation that all the members of the sample are six years older than in the first period.

5. The EXP variable obviously drops out of the changes estimates because each person aged 6 years between the two dates. In Schools 2 and 4 no additional sample members acquired administrative experience between these dates.

6. The test statistic is F(3, 129) = 2.73.

7. Viscusi-O'Connor (1984) present equations estimated over both levels and changes in which the estimated R^2 are fairly close. Like the estimates here, the wage data underlying their equations are obtained from administrative records rather than from recall by workers.