Running Head: Panel Models

Panel Models in Sociological Research: Theory into Practice

Charles N. Halaby

Department of Sociology

University of Wisconsin–Madison

halaby@ssc.wisc.edu

608.263.4609

608.831.1806

December 2003

Prepared for Annual Review of Sociology 2004,

Panel Models in Sociological Research: Theory into Practice

> "[F]or some reason there is widespread, though not well articulated, opinion that in panel analysis the usual obstacles to inference and estimation are suspended for the benefit of the analyst." Otis Dudley Duncan (1972 pp.36)

I. Introduction

Panel studies are fast displacing their cross sectional counterpart at the heart of sociological research. This is the culmination of a trend that dates to the 1960s and has quickened over the last two decades. Panel studies have come to figure prominently in numerous lines of sociological research that address a broad range of subjects spread across all levels of analysis. For research aimed at variation across large-scale social units, panel data proliferate on subjects ranging from welfare spending and poverty (Huber and Stephens 2000; Moller et al. 2003) to political violence (Villareal 2002). Dramatic developments in the use of panel data also have occurred at the individual level, where panel analysis is now the preeminent form of social research on a host of educational, career (Budig and England 2001), and family (Morrison and Ritualo 2000) outcomes. In the areas of crime and deviance, which have a long history of longitudinal studies, the application of panel data to individual-level analyses continues unabated (Hagan and Foster 2001; Osgood et al. 1996). Wedged between applications to large-scale social aggregates and individuals is the growing body of panel studies of firms and organizations (Baron, Hannan and Burton 2001; Boone, Carroll and Witteloostuijn 2002).

The accelerated pace at which panel studies are emerging attests to the prevailing belief that panel data are amply suited to the analytical problems that surround the kinds of observational (i.e., non-randomized) data that are common in social research. The fundamental structure of panel data provides the analytical leverage that is indispensable for rigorously achieving the central aim that drives the bulk of quantitative research: the estimation of causal effects. When Waldfogel (1997) estimates the effect of additional children on mothers' wages; or Hagan and Foster (2001) assess the psychological and life course consequences of adolescent exposure to violence in intimate relationships; or Cherlin, Chase-Lansdale, and McCrea (1998) estimate the effect of parental divorce on the mental health of children, they are pursuing the kinds of questions that are at the forefront of panel studies and for which panel data have unique advantages.

The problem of causal inference is fundamentally one of unobservables, and unobservables are at the heart of the contribution of panel data to solving problems of causal inference. Two types of unobservables are problematic for the identification and estimation of causal parameters in nonrandomized studies: 1) time-invariant unit-specific unobservables that represent permanent properties of units (i.e., "unit effects"); and 2) time-varying unit-specific unobservables that represent transitory and idiosyncratic forces acting upon units (i.e., "disturbances").[1] Panel data offer certain advantages for dealing with such unobservables, but these advantages can only be realized through the use of statistical methods that capitalize on the structure of observations that extend across units and over time.

Methods for estimating causal parameters from panel data grow largely out of an econometric tradition that dates back three decades. Major advances in the modelling of panel data originated in the 1970s and developed further during the 1980s and 1990s. Yet sociological practice has not absorbed fully the lessons of the econometric literature. On the contrary, sociologists have been slow to capitalize on the advantages panel data offer for controlling these unobservables and mitigating the threat they pose for causal inference. Key principles that ought to routinely inform analysis are at times glossed over or ignored completely. Tests and techniques that ought

[1] As discussed below, panel data are also useful for dealing with time-varying unit-invariant unobservables that accompany the passage of time (i.e., "period effects").

to be standard practice because they capitalize on the very strengths of panel data instead are implemented in a haphazard fashion. I am not referring to arcane methods at the frontiers of applied econometrics, but rather to analytic practices that achieve the central purpose of panel data and that have been repeatedly advocated in essays appearing in sociological journals (Allison 1994), including this very journal (Hannan and Tuma 1979 ; Petersen 1993).

As a vehicle for discussing recent patterns of sociological practice and highlighting core models and methods for the analysis of panel data, I will use a compilation of selected panel studies appearing between 1990-2003 in the *American Sociological Review* and *American Journal of Sociology.* These journals arguably represent best practice, yet many of the studies appearing in them fail to exploit the opportunities that standard panel models and methods offer for dealing with unobservables. My aim is to hasten the absorption of fundamental panel analytic principles into research practice, and thereby encourage valid inference from panel data.

The substantive scope of this review falls short of both previous *ARS* articles on longitudinal analysis. My review does not approach the reach of Hannan and Tuma's (1979) wide-ranging discussion of all manner of longitudinal analysis for both metric and qualitative outcomes, or Petersen's review of duration models for event history data and static models for panel data. I concentrate on parametric methods for estimating models for metric outcomes, which covers the vast majority of panel studies in sociology. With that restriction, I cover estimation methods for static models as well as methods for dynamic models with lagged endogenous variables. I pay special attention to issues pertaining to the bias and consistency of estimators, since they have priority over matters of efficiency and the valid estimation of standard errors.[2] Hence, this review attends largely to the basics of panel data analysis, since that seems warranted by the applications appearing in the highest-profile journals in sociology.

There exist a number of very good treatments of the subjects examined below. Standard econometric texts on panel data include Hsaio (1986), Baltagi (2001), Lee (2002) and Arellano (2003), all of which are advanced. My personal preference runs toward the texts by Wooldridge (2003 Chapters 13-14; 2002b). Special mention is due Maddalla's (1987) pithy review and Allison's papers on change scores (1990) and estimating the effects of events (1994).

This article has three main sections. The first section specifies the essential advantages of panel designs and identifies basic principles that apply to the analysis of panel data. The second section discusses the key principles that accompany the formulation and estimation of static models. Dynamic models are considered in the third section.

II. Unobservables in three basic research designs

There are two fundamental observational protocols for collecting data for the purpose of causal inference (Holland 1986): different units may be exposed to different values of a causal variable and their responses compared at a single point in time; or the same units may be exposed to different values of the causal variable and their responses compared at different times. Campbell and Stanley (1963) call these the "static-group comparison" and the "one-group pretest-posttest" design, respectively. The mechanism governing exposure (or assignment) under both designs is, of course, an important consideration. By combining these two protocols, a panel design joins

[2]The technical scope of this review is mostly limited to applications in which the number of units (N) is large, the number of occasions (T) is relatively small, and the data are balanced. For extensions to incomplete panels, including discussions of selection bias, see Baltagi (2001, chapters 9 and 11.5) and Wooldridge (2003 Chapter 17), and Lee (2002 Chapter 6).

their strengths and eliminates their characteristic weaknesses because it requires less restrictive assumptions on unobservables to achieve identification and unbiased estimation of causal parameters. The main purpose of this section is to review the role of unobservables in the analysis of panel data. To that end, as well as for clarifying connections among the various research designs, I begin with over-parameterized models in anticipation of a full panel data scheme.

Static group comparison

Consider first the static-group comparison, the prototype of cross-sectional data. Assume that data are available on a response variable y_{it} for $i = 1, 2, \ldots, N$ at a single point in time ($t = 1$), and on a causal variable d_i scored $d = 1$ for the treatment group and $d = 0$ for the control group, where treatment occurs during a period τ preceding $t = 1$. One model for this design is:

$$y_{i1|d=0} = \theta_{i|d=0} + \delta_1 + \varepsilon_{i1|d=0} \tag{1}$$

for the control group; and

$$y_{i1|d=1} = \gamma + \theta_{i|d=1} + \delta_1 + \varepsilon_{i1|d=1} \tag{2}$$

for the treatment group. Here y and d are the only observed variables, γ is a parameter for the causal effect, δ_1 is a period effect common to all units, and $\varepsilon_{i1|d}$ is a transitory idiosyncratic disturbance unique to the i^{th} unit at time $t = 1$ conditional on d and θ_i. The quantity $\theta_{i|d}$ is a time-invariant unit-specific effect that captures *unobserved unit heterogeneity*, is conditional on d, but is assumed, here and for all models considered later, independent of $\varepsilon_{i1|d}$. This unit effect $\theta_{i|d}$ can be viewed as a summary of time-invariant unit-specific causes of the response variable, or as the unobserved permanent component of the ith unit's value of y_{i1}. The idea that $\theta_{i|d}$ represents time-invariant unit-specific causes implies that these causes are stable; I assume that their effect on the response variable, here normalized to unity, is also stable over time, a restriction that may be relaxed in practice. The distinction between unit effects and the disturbance is totally artificial in this design, since even $\varepsilon_{i1|d}$ is temporally invariant with the period fixed at $t = 1$. Both θ_i and ε_{i1} together comprise the composite error, $\mu_{i1|d} = \theta_{i|d} + \varepsilon_{i1}$.

The strength and weakness of this design is revealed by subtracting (1) from (2) and taking expected values:

$$E(y_{i1|d=1} - y_{i1|d=0}) = \gamma + E(\theta_{i|d=1} - \theta_{i|d=0}) + E(\varepsilon_{it|d=1} - \varepsilon_{it|d=0}) \tag{3}$$

where the term δ_1 for a period effect has dropped out. This is the key, if trivial, advantage of this design: observing both groups at the same time guarantees that confounding changes that might have otherwise accompanied the passage of time are ruled out as alternatives to d as the source of a mean difference in the response variable. Yet the causal parameter γ is still not identified without additional restrictions. One assumption is that the mean level of the disturbances is independent of the causal variable d, so that $E(\varepsilon_{i1|d=1} - \varepsilon_{i1|d=0}) = 0$. Although this *exogeneity assumption* is not unproblematic, some version of it is always required with observational data. The exogeneity assumption does not cover the term for the mean difference in the unit effects in the treatment and control groups. Even if the causal variable is exogenous and even in the absence of a causal effect (i.e. $\gamma = 0$), the model does not imply that the mean level of the response variable would be the same in both groups. The crux of the problem is that (3) expresses a between-group comparison. The treatment and control units are different, and hence possibly heterogeneous with respect to unobserved properties that may confound the attribution of effect to the causal variable. Hence, one further identifying restriction is necessary, namely, $E(\theta_{i|d=1} - \theta_{i|d=0}) = 0$, that is, unobserved heterogeneity is mean independent of the causal variable ($E(\theta_i|d) = E(\theta_i)$). This restriction on the time-invariant unit-specific unobservables is known as the “random effects” effects assumption, in which case (1) and (2) are a *random effects* model.

Under these restrictions on the unit effects and disturbances, a regression of the form

$$y_{i1} = \alpha + \gamma d_i + \mu_{i1} \tag{4}$$

yields the least-squares estimator

$$\hat{\gamma}_{ls} = (\overline{y}_{.1|d=1} - \overline{y}_{.1|d=0}) \tag{5}$$

where the right-hand-side is the observed difference in conditional sample means. Expressing (5) in terms of the causal parameter and unobservables yields:

$$\hat{\gamma}_{ls} = \gamma + (\overline{\theta}_{i|d=1} - \overline{\theta}_{i|d=0}) + (\overline{e}_{t|d=1} - \overline{e}_{t|d=0}) \tag{6}$$

which shows that the least-squares estimator captures the causal effect γ, and the sample mean differences between the treatment and control groups in the permanent and idiosyncratic components of the response variable. Under the exogeneity and random effects assumptions, the least-squares estimator is unbiased because the last two terms on the right are zero in expectation. If the restriction on the disturbances is violated, the least-squares estimator suffers from "endogeneity bias." If the random effects assumption is violated, the least-squares estimator suffers from "heterogeneity bias." It bears repeating that the distinction between endogeneity and heterogeneity bias is totally artificial for this design, since the two are indistinguishable with each unit observed only once. This distinction becomes meaningful under a panel design.

There is nothing inherent in the static group comparison that secures either the exogeneity or random effects assumption. The plausibility of these restrictions hinges on the mechanism governing exposure (or assignment) to the different values of the cause. These assumptions would be defensible under randomized assignment, or, equivalently, if the treatment and control groups were, apart form the causal effect γ, known to be randomly sampled from the same population distribution of the response variable.[3] In observational studies with arbitrary assignment to treatment, these assumptions are rarely warranted without augmenting the basic design, typically through regression adjusting $\hat{\gamma}_{ls}$ by adding measured covariates to (4). The hope is that covariates will "randomize" the conditional variation in θ_i and ε_{it}.

One group pretest-posttest

Now consider the one group pretest-posttest design. Write the observation on the dependent variable y_{it} at time $t = 0$ under cause $d = 0$ as:

$$y_{i0|d=0} = \delta_{0|d=0} + \theta_i + \varepsilon_{i0|d=0} \tag{7}$$

and the observation on the same unit at time $t = 1$ with cause $d = 1$ as

$$y_{i1|d=1} = \gamma + \delta_{1|d=1} + \theta_i + \varepsilon_{i1|d=1} \tag{8}$$

where the unit effects are the same over time. Differencing and taking expected values yields

$$E(y_{i1|d=1} - y_{i0|d=0}) = \gamma + (\delta_{1|d=1} - \delta_{0|d=0}) + E(\varepsilon_{i1|d=1} - \varepsilon_{i0|d=0}) \tag{9}$$

where the unit effects have dropped out. This equation is analogous to (3), except the differencing is across the time dimension for the same units instead of across the unit dimension at the same time. Because the units are the same at both times, unobserved properties captured by θ_i are ruled out as a source of change in the response variable. Hence, this design formally eliminates the threat of unobserved heterogeneity bias. Yet here again the causal parameter γ is not identified without additional restrictions. The last term on the right represents the association between the disturbances and the causal variable, but under the exogeneity assumption the average effects of

[3] The assumptions are also secured if a scalar mechanism determines assignment, and the mechanism is included as a variable in the regression. Such a mechanism would turn the static-group comparison into a regression discontinuity design.

transitory forces cancel out over time, so that $E(\varepsilon_{i1|d=1} - \varepsilon_{i0|d=0}) = 0$. That leaves for identification purposes the middle term in (9), which captures the over time change in the effect of time-varying unobservable forces common to all units. Hence, even in the absence of a causal effect (i.e., $\gamma = 0$), the model does not imply that the mean level of the response variable would be the same before and after exposure to treatment. That implication requires the temporal stability restriction $(\delta_{1|d=1} - \delta_{0|d=0}) = 0$, that is, there would be no change overtime in the mean of y if $\gamma = 0$. Hence, the temporal stability that was the strength of the static group comparison is lost in this design because the over time change $(\delta_{1|d=1} - \delta_{0|d=0})$ in period effects confounds the causal effect γ.[4] But under exogeneity and temporal stability, a regression of y_{it} on a constant for the common period effect $(\delta_0 = \delta_1 = \delta)$ and on d_i yields the least-squares estimator:

$$\hat{\gamma}_{ls} = (\overline{y}_{i1|d=1} - \overline{y}_{ti0|d=0}) \tag{10}$$

$$\hat{\gamma}_{ls} = \gamma + (\overline{e}_{i1|d=1} - \overline{e}_{i0|d=0}) \tag{11}$$

which is unbiased because the expectation of the last term on the right is zero by assumption.

Panel design and estimators

A panel design achieves a measure of protection against the threats of unit heterogeneity and temporal instability. It also offers relief with respect to the disturbances, since it permits identification under a weaker exogeneity assumption. To see these advantages, extend the static-group comparison backwards to time $t = 0$ when $d_i = 0$ for all units:

$$y_{i0|d=1} = \theta_{i|d=1} + \delta_0 + \varepsilon_{i0|d=1} \tag{12}$$

$$y_{i0|d=0} = \theta_{i|d=0} + \delta_0 + \varepsilon_{i0|d=1} \tag{13}$$

where $d = 1$ anticipates exposure to the cause between $t = 0$ and $t = 1$ for the treatment group and $d = 0$ for $t = 0, 1$ identifies the control group. The expected value of the difference between

[4] Observe that a dummy variable for time period would be perfectly collinear with d_i.

these equations is:

$$E(y_{i0|d=1} - y_{i0|d=0}) = E(\theta_{i|d=1} - \theta_{i|d=0}) + E(\varepsilon_{i0|d=1} - \varepsilon_{i0|d=0}) \tag{14}$$

which is the pretest analogue of the posttest expression given in (3). Again, the period effect δ_0 has been eliminated. Furthermore, the term for the difference between treatment and control in the mean of the unit effects at pretest must equal the same term in the posttest equation, since the effect of θ_i is constant over time. Hence, unlike the static-group comparison, where the random effects assumption was necessary for identification, here unobserved heterogeneity that is correlated with the causal variable may be dealt with by treating θ_i as fixed and subtracting the pretest equation from the posttest equation:

$$E(y_{i1|d=1} - y_{i1|d=0}) - E(y_{i0|d=1} - y_{i0|d=0}) = \gamma + E(\varepsilon_{i1|d=1} - \varepsilon_{i1|d=0}) - E(\varepsilon_{i0|d=1} - \varepsilon_{i0|d=0}) \tag{15}$$

On the left, the posttest difference in means has been adjusted by the pretest difference; on the right, this adjustment has eliminated the unit effects that were a source of heterogeneity bias.

Identifying the causal parameter in a panel design does not require the strong random effects restriction that unobserved sources of heterogeneity are mean independent of the causal variable; rather, the θ_i may be treated as fixed effects that are arbitrarily correlated with the causal variable, since the expression above accounts for such correlations. Equations (1) and (2) together with (12) and (13) constitute a *fixed effects* model. Nor does identification require, like the pretest-posttest design, the assumption that period effects are temporally stable. Indeed, the only identifying restriction necessary is exogeneity: the mean difference in the disturbances in the treatment and control group remains the same between pretest and posttest (i.e., $E(\varepsilon_{i1|d=1} - \varepsilon_{i1|d=0}) - E(\varepsilon_{i0|d=1} - \varepsilon_{i0|d=0}) = 0$). Here too panel data have an advantage: assuming that the *difference* in the means of the disturbances in the two groups remain stable over time is weaker than assuming, as in the static-group comparison, that the mean *levels* of the disturbances are the same at a given time.

A panel design may also be viewed from the vantage point of the pretest-posttest design. To this end, augment the pretest-postest design for the treatment group with pretest-posttest observations $y_{i1|d=0}$ and $y_{i0|d=0}$ for a control group. Then subtracting expected pretest-posttest differences for the control group from those for the treatment group yields:

$$E(y_{i1|d=1} - y_{i0|d=1}) - E(y_{i1|d=0} - y_{i0|d=0}) = \gamma + E(\varepsilon_{i1|d=1} - \varepsilon_{i0|d=1}) - E(\varepsilon_{i1|d=0} - \varepsilon_{i0|d=0}) \quad (16)$$

which is just (15) re-arranged. Note that the original pretest-posttest expression for the over time change in the mean of y in the treatment group has been adjusted by the overtime change in the mean of y for the control group. This amounts to treating the time path of the response variable in the control group as a proxy for what the time path of y in the treatment group would have been in the absence of treatment. This assumption hinges on the exogeneity restriction: the overtime mean change in unobserved time-varying causes of y is the same in both the treatment and control group, so that the last two terms on the right cancel out.

As (15) and (16) suggest, an unbiased estimator of γ may be constructed from a difference of differences in sample means without resorting to the random effects and temporal stability assumptions. The regression form of the posttest and pretest models is:

$$y_{i0} = \delta_0 + \theta_{i|d} + \epsilon_{i0} \quad (17)$$

$$y_{i1} = \delta_1 + \gamma d_i + \theta_{i|d} + \epsilon_{i1} \quad (18)$$

which upon differencing yields:

$$(y_{i1} - y_{i0}) = (\delta_1 - \delta_0) + \gamma d_i + (\epsilon_{i1} - \epsilon_{i0}). \quad (19)$$

Least-squares estimation of (19) yields the *difference-in-differences* estimator:

$$\hat{\gamma}_{dd} = (\overline{y}_{.1|d=1} - \overline{y}_{.1|d=0}) - (\overline{y}_{.0|d=1} - \overline{y}_{.0|d=0}) \quad (20)$$

of γ and $(\overline{y}_{.1|d=0} - \overline{y}_{.0|d=0})$ as an estimator of $(\delta_1 - \delta_0)$. The difference-in-differences estimator (DID) for a binary treatment variable is a member of a larger class of "first difference" (FD) estimators, based as it is on regressing the difference in the response variable between t and $t-1$ on the same difference in the causal variable. Least squares applied to (19) also yields valid standard errors and test statistics.

An alternative transformation for the purpose of estimation is to express the response and causal variables as deviations from their within-unit overtime means. Averaging (17) and (18) overtime yields:

$$\overline{y}_{i.} = (\delta_1 + \delta_0)/2) + \gamma \overline{d}_{i.} + \theta_i + \overline{\varepsilon}_{i.} \tag{21}$$

which is the "between" regression of $\overline{y}_{i.}$ on $\overline{d}_{i.}$. Time-demeaning (17) and (18) and pooling over time yields:

$$y_{it} - \overline{y}_i = (\delta_0 - \delta_1)/2 + (\delta_1 - \delta_0)p_1 + \gamma(d_{it} - \overline{d}_i) + (\epsilon_{it} - \overline{\epsilon}_i) \tag{22}$$

where p_1 is a dummy for period 1. Applying least squares to (22) yields the unbiased and consistent *fixed effects* (FE) estimator $\hat{\gamma}_{fe}$, and valid standard errors and tests statistics.[5] In the two-period case with binary causal variable, the difference-in-differences, the first difference, and the fixed effects estimators are all equivalent. By exploiting within- rather than between-group variation, these estimators achieve unbiasedness and consistency even when the random effects assumption fails because the unit effects are correlated with the explanatory variable.

The difference-in-differences estimator

Before turning to more general models, DID estimation warrants a closer look because it is in its own right a powerful approach to the analysis of two-period panel data. The DID estimator has become popular among economists as the first choice of methods for estimating the causal

[5]This supposes that the residual degrees of freedom correctly account for the estimation of N unit means.

effects of so-called "natural experiments". A natural experiment, or what Campbell and Stanley (1963) called a "quasi-experiment," is a data generating—or data *collecting*—process in which there is "a transparent exogenous source of variation in the explanatory variables that determine the treatment assignment" (Meyer 1995 p 151). Or, as described by Angrist and Krueger (2001), "[It is] a situation where the forces of nature or government policy have conspired to produce an environment somewhat akin to a randomized experiment" (p. 7). Natural experiments promise some protection against violations of the exogeneity assumption upon which the quality of DID estimation depends. To the extent that the "natural" mechanism governing assignment to values of the causal variable is not related to the time path of the outcome variable, endogeneity bias is reduced. Meyer (1995) gives a thorough review of the DID estimator in natural experiments.

There has been no shortage of applications of DID estimation to natural experiments in the last 15 years. One especially influential study is Card and Krueger's (1994) analysis of the employment consequences in the fast-food industry of an increase in the minimum wage in New Jersey. Card and Krueger surveyed restaurants in New Jersey before and after legislation raising the minimum wage, and use the change in employment over the same period in eastern Pennsylvania restaurants to "difference out" the employment shift that would have occurred in New Jersey in the absence of a new minimum wage law. Card and Krueger employ a true panel design, but the DID estimator also works for independent repeated cross-sections consisting of entirely different units at the two time periods (e.g., Meyer, Viscusi, and Durbin 1995).

The DID estimator continues to invite methodological efforts to protect against endogeneity bias. An early and influential methodological contribution is Ashenfelter's and Card's (1985) analysis of the earnings impact of training programs. They noted that the validity of the differencing procedure that leads to the difference-in-differences estimator hinges on the validity of the original model for the response varible. Ashenfelter and Card (1986) describe a variety of simple tests of this specification and the validity of the differencing procedure. The tests exploit restrictions that the model imposes on the data, and may be applied when, in addition to y_t and y_{it-1}, values of y_{it-j} for $j > 1$ are also observed for the period prior to the causal event of interest.

An especially notable line of recent research aims at locating DID estimation squarely within the potential outcome, counterfactual approach to causal effects. Abadie (2002) develops a semi-parametric DID estimator of the average effect of the treatment on the treated in cases where the exogeneity assumption fails because the time path of the response variable differs in the treatment and control groups. Athey and Imbens (2002) generalize the DID framework by introducing a nonlinear model that incorporates over time changes in the effect of time-invariant unobservables. Finally, the DID estimator has been integrated with propensity score matching methods by Heckman et al. (1997), who propose regression-adjusted difference-in-difference matching estimators of the effect of the treatment on the treated.

DID estimation extends to long panels and the analysis of the effect of an event occurring at different times for different units (e.g., Cherlin et al.). An excellent treatment of regression-adjusted DID estimation in long panels is provided by Allison (1994). A note of caution is sounded by Bertrand, Duflo, and Mullainathan (2002), who warn that in long panels serially correlated disturbances and causal variables induce severe downward bias in the DID standard errors.

II. Specification issues in static panel models

The DID, FE, and FD estimators offer researchers the capacity to dispense with the random effects assumption and still obtain unbiased and consistent estimates of parameters when unit effects are arbitrarily correlated with measured explanatory variables. This is widely regarded as the primary advantage of panel data, and why the effort to extend the benefits of fixed effect models beyond the static linear case is one of the central thrusts of econometric research on panel analysis over the last 15 years. Yet sociologists have been slow to appreciate fully the power of fixed effect models. Illustrative evidence of the disparity between sociology and economics in this respect comes from an admittedly crude search of the JSTOR archives. A search on the keywords "fixed effects" turns up 14 references in the *American Journal of Sociology* 1990-2000 and 9 in the *American Sociological Review* 1990-1997. The same search yields 68 references in the *Journal of Labor Economics* 1990-1997, 136 references in the *American Economics Review* 1990-2000, and 61 references in the *Review of Economic Statistics* 1990-1997.

Within sociology the contrast between the rate of adoption of random effects as against fixed effect models is not quite as sharp. Table 1 gives a compilation of key properties of 31 panel studies culled from *ASR* and *AJS*. Column 5 shows that 15 adopted a random effects model, 11 adopted a fixed effects model, and 5 employed both types of models. So fixed effects models have found their way into sociological research, but have not become standard practice. And the over time trend is markedly away from fixed effects and toward random effect models.[6] This despite the availability of excellent didactic discussions of the benefits of fixed effects models (Allison 1994; Firebaugh and Beck 1994).

What explains the failure of fixed effect models to more deeply penetrate sociological panel studies? In this section I review the concerns that seem to have impeded the spread of fixed effect models and led on occasion to their outright rejection. I also highlight methods that sociologists may employ to more fully capitalize on the structure of panel data, including well-known but rarely used specification tests, models for combining random and fixed effects estimators, and methods for dealing with threats to exogeneity.

Random effects, fixed effects, and unobserved heterogeneity

Generalize the model considered above to more than two periods and for regression adjustment of

[6] Before 1999, the ratio of fixed to random effects models was about 2:1; since then it is 1:3.

covariates:

$$y_{it} = \sum_k \beta_k w_{kit} + \sum_p \phi_p z_{ip} + \gamma x_{it} + \theta_i + \varepsilon_{it} \tag{23}$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$ and terms for period effects δ_t have been suppressed to simplify notation. This model includes a term θ_i for unit effects, and a zero-mean transitory disturbance ε_{it} that varies over time and units, is mean independent of θ_i and the explanatory variables in all periods, and, conditional on the explanatory variables and θ_i, has constant variance and is uncorrelated over time and across units. Hence, the explanatory variables are strictly exogenous with respect to ε_{it} ($E(x_{is}\varepsilon_{it}) = 0$ for all s, t) and the structure of the disturbance is classical. The causal variable of interest is x_{it}, which may be metric or binary, with parameter γ. Additional explanatory variables are of two types: the w_{kit} ($k = 1, \ldots, K$), which vary over time and across units, and the z_{pi} ($p = 1, \ldots, P$), which vary only between units because, like θ_i, they represent time-invariant unit characteristics.

The key issue in estimating (23) is whether the unit effects θ_i are to be treated as random or fixed. This choice hinges entirely on whether the unit effects θ_i are correlated with the explanatory variables. If the unobserved θ_i are uncorrelated with the regressors, nothing is gained by distinguishing "within" and "between" unit variation in the estimation of the parameters. In this case θ_i may be treated as a random effect, with an unbiased and consistent estimator of γ (and other parameters) obtained by applying least squares to the pooled panels of NT observations. There is, however, a gain in efficiency, as well as valid standard errors and test statistics, to be realized by taking account of the serial correlation among the composite errors that is induced by the fact that $u_{it} = \theta_i + e_{it}$ and $u_{is} = \theta_i + e_{is}, s \neq t$, contain the common θ_i. A better approach is generalized least squares (henceforth, GLS), which would yield a consistent and efficient random effects estimator $\hat{\gamma}_{re}$ as well as valid standard errors and test statistics.[7]

[7] Under the assumptions on the disturbances, the so-called random effects estimator is equivalent to GLS.

Many studies listed in Table 1 ignore the issue of unobserved unit effects altogether, or recognize such effects, but fail to assess and take steps to deal with their correlation with measured covariates. Column 6 of Table 1 summarizes the nature of the discussion of correlated unit effects in these studies. The coding of this column imposes a strict criterion by asking whether an article explicitly addresses the random effects assumption by discussing the implications of correlated unit effects for the bias and inconsistency of estimators under a random versus fixed effects model. A glance shows that the vast majority of studies do not explicitly engage this issue.[8] In most studies that employ a random effects model followed by OLS or GLS, the random effects assumption is given short shrift, with little or no discussion of the biased and inconsistent estimators that result if unit effects are correlated with the explanatory variables.

The random effects approach is especially prevalent in some areas of research. Random effects is typical of the recent cross-national panel analyses appearing in Table 1, such as Kenworthy's study of unemployment (2002), Nielsen and Alderson's studies of inequality (1997; 2002), Beckfield's study of international organizaitons (2003), and the Shofer et al. (2000) study of science and economic growth; exceptions that adopt a fixed effects approach are Pampel's (1994) astute analysis and Firebaugh and Beck's (1994) study of economic growth and welfare, which explicitly advises cross-national researchers of the benefits of fixed effect models. When T exceeds N by a fair margin, as in Huber and Stephens's study of state provision of social services (2000), Boone's et al. (2002) study of market partitioning in the Dutch newspaper industry, and Sutton's (2000) analysis of imprisonment rates, the choice of model may be moot, since GLS converges to fixed effects for fixed N as T increases.[9] Yet even in individual level analyses, where N is typically much

[8]Many more studies do observe that FE (or FD) estimation "controls" unobserved time invariant effects, although typically the implications for the quality of estimators is left implicit.

[9]Although the theory for T and N of similar magnitude is not nearly as well developed as for the fixed T and large N case that this review addresses, Wooldridge (2002 pp. 7) states that estimation methods for the latter case

larger than T, random effects models are sometimes employed with at best dim recognition of the inconsistency of the GLS estimator if unit effects are correlated with the explanatory variables (Western and Beckett 1999; Baron et al. 2001; Darnell and Sherkat 1997).

On those occasions when the choice of random or fixed effects is forthrightly addressed, the crucial distinction between unit effects that are or are not correlated with measured variables is sometimes blurred. Consider this passage from Nielsen and Alderson's (1995) cross-national study of income inequality:

> [Panel data] are amenable to the use of estimation methods that deal with potential *heterogeneity bias* [italics in original], i.e., the confounding effect of unmeasured time-invariant variables that are omitted from the regression model. ... The *fixed effects* and the *random effects* [italics in original] models are two commonly used estimation strategies designed to correct for unmeasured time-invariant factors. 1995 pp. 685).

Similarly, Western and Beckett (1999), in their study of the effect of youthful incarceration on adult employment prospects, write:

> "Random effects are also added...to adjust for unobserved heterogeneity. The random effects model...accounts for respondent-specific characteristics that are unobserved and not captured by the independent variables."

It is true that the random effects estimator recognizes heterogeneity induced by the unit effects, but only insofar as variance estimators are adjusted for the serial correlation that heterogeneity induces in the composite error. GLS estimation of a random effects model does not control or correct for the "heterogeneity problem," conventionally understood to mean *correlated* unobserved unit effects that generate bias and inconsistency in parameter estimates.

can usually be used in the former.

Absent plausible theoretical grounds or empirical evidence for the random effects assumption, bias and consistency considerations alone would lead to a fixed effects model and the same estimation methods identified earlier for the two-period case. First differencing (23) gives:

$$(y_{it} - y_{it-1}) = \sum_k \beta_k (w_{kit} - w_{kit-1}) + \gamma(x_{it} - x_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}). \tag{24}$$

Applying least squares to the pooled data yields the first-difference estimator $\hat{\gamma}_{fd}$ of γ. Alternatively, applying the fixed-effects transformation (i.e., time-demeaning the data) to (23) gives:

$$(y_{it} - \overline{y}_i) = \sum_k \beta_k (w_{kit} - \overline{w}_{ki}) + \gamma(x_{it} - \overline{x}_i) + (\varepsilon_{it} - \overline{\varepsilon}_i) \tag{25}$$

which can be estimated by least squares to yield the FE estimator $\hat{\gamma}_{fe}$ of γ. As column 7 of Table 1 shows, the fixed effects transformation is more popular than the first difference transformation among studies adopting a fixed effects model.[10]

The FE and FD estimators, both of which exploit within-unit variation as a means of purging unit heterogeneity, are unbiased and consistent under strict exogeneity, although for $T > 2$ they are not the same. The standard errors and test statistics that accompany the FE estimator $\hat{\gamma}_{fe}$ are valid if the disturbances ε_{it} are constant variance and serially uncorrelated; this holds as well for the first-difference estimator if the disturbances $(\varepsilon_{i2} - \varepsilon_{i1})$ in the *transformed* equation are constant variance and serially uncorrelated. Under these assumptions, both FE and FD estimators are efficient for a fixed effects model.[11] The efficiency of both estimators depends directly on the over

[10]The time-demeaned FE estimator is also called the "within" estimator. I will refer to both the FE and FD estimators as "within-group" estimators, or simply "within" estimators. One nice property of the time-demeaned FE estimator is that it is equivalent to using the deviation of each time-varying explanatory variable from its time mean as an instrument for the level of the variable in equation (23).

[11]If the disturbances in the original equation are serially uncorrelated, then the first difference disturbances will be correlated in adjacent periods, so that GLS applied to the first-difference equation would be optimal. Alternatively, a forward orthogonal deviations transformation could be applied to yield a difference equation in which the disturbances are serially uncorrelated (Arrelano 2003 pp. 17-18).

time variation in the explanatory variables, since one cannot get precise estimates of the effect of a change in a causal variable if not much change actually occurred.

Model choice, efficient estimation, and problematic error structures

Why are many researchers quick to reject fixed effects in favor of random effects models? Efficiency is one major criterion used to defend the choice of a random effects model and GLS estimation. Nielsen and Alderson (1997), in eschewing a fixed for random effects approach, observe:

> "[F]ixed effects [estimation] can be interpreted substantively as "throwing away" all between-[unit] variation present in the data. …[R]andom effects [estimation] is asymptotically efficient relative to fixed effects [estimation]" (1995 pp. 685-86).

This point of view, not uncommon among researchers, needs to be qualified in a number of respects. First, the efficiency advantage of GLS estimation over within group estimation strictly holds only if the disturbances are homoscedastic and serially uncorrelated. If the disturbances are either heteroscedastic or serially correlated, then the two estimators cannot be ranked in terms of efficiency (Arellano 2003 p. 41). Second, the luxury of "throwing out" between variation is the very source of the advantage that panel data provide over cross-sectional data and that within group estimators like FE and FD exploit to avoid heterogeneity bias. "Throwing out" between variation is not wasting data: it buys protection against biased and inconsistent parameter estimates. As between variation in the explanatory variables comes to dominate within variation, violations of the random effects assumption become more problematic because the GLS estimator averages between and within variation. Hence, correlated unit effects become more of an issue, not less.[12] Third, the efficiency advantage of GLS estimation of a random effects model with classical

[12] Alternatively, as T increases for fixed N, or the variance of the unit effects increases relative to the disturbance variance, the GLS estimator approaches the FE estimator, so the random effects assumption is less problematic.

disturbances hinges on the validity of the random effects assumption. This is an important point that is easily lost in discussions of the efficiency of GLS relative to within estimation. GLS depends for its optimal variance properties on consistent estimation of the model parameters. Only if the θ_i are uncorrelated with the explanatory variables is it compelling to think of the FE and FD estimators, which are unbiased and consistent regardless of the validity of the random effects assumption, as less efficient than GLS, for only then is the latter estimator consistent. To the extent that the random effects assumption is problematic, or disturbances are nonclassical, so too are claims for the efficiency advantage of GLS over within group estimators.

Correlated unit effects are a threat to more than just the efficiency if random effects estimators. The validity of variance estimators under GLS estimation also hinges on the random effects assumption. Researchers have employed a bewildering array of assumptions regarding the structure of the composite errors in an effort to secure valid standard errors when estimating random effects models. Yet the success of such efforts depends on consistent estimation of the model parameters. The GLS methods commonly used to obtain valid standard errors and test statistics for random effects models help little if unit effects are correlated with explanatory variables.

None of this is meant to suggest that within group estimation rules out problematic error structures. Yet in practice dealing with heteroscedastic or serially correlated errors is much more straightforward in a fixed effects compared to random effects context, and does not depend on the random effects assumption. Most researchers who adopt a fixed effects model are content to purge unit effects and then use the standard errors produced by applying OLS to a first-difference or time-demeaned equation (Table 1, column 9). Hence, neither Lichter et al. (1997), nor Waldfogel (1997), nor Gustafson and Johansson (1999) go beyond OLS to obtain standard errors after purging the unit effects. One reason for this is that many of the issues that ordinarily arise regarding the distribution of unobservables in random effects models are largely resolved by purging the unit effects, leaving only the disturbances as a potential source of difficulty. To be sure, problems of heteroscedastic and serially correlated disturbances that remain in the transformed equation will bias the OLS standard errors and render test statistics invalid. But there is no shortage of tests for and methods of dealing with these problems in the context of fixed effect models, although the details differ depending on the distribution of the original disturbances and whether the FE or FD transformation is used (Wooldridge, 2002, Chapter 12).

Robust variance estimators are one practical solution to problematic disturbances under FE or FD estimation. Ritualo and Morrison (2000) and McManus and Diprete (2001) use Huber-White standard errors following FE estimation to deal with heteroscedasticity. Variance estimators that are robust to general forms of heteroscedasticity and serial correlation are also available. Newey-West standard errors would be an obvious choice, since they are consistent even in the presence of heteroscedasticity and serial correlation. Even for large T and fixed N, as in some cross-national studies (Huber and Stephens 2000; Boone et al. 2002), variance estimators that are robust to cross-sectional dependence, heteroscedasticity, and serially correlated errors are available for within group estimators (Arellano 2003 p. 19). Indeed, virtually all the tests and methods for dealing with errors in a random effects context are also available for within group estimators. Verbeek (2000 p. 324) suggests that even in instances of uncorrelated unit effects, researchers may wish to employ tests and methods designed for FE estimation, since they are more straightforward than their random effects counterparts.

Model choice and time-invariant explanatory variables

Another major reason many researchers prefer random effects models and estimators is that within group estimators of fixed effect models fail to identify the parameters of observed time-invariant variables (e.g., Nielsen and Alderson 1995; Kenworthy 2002; Huber and Stephens 2000; Beckfield

2003; Baron, Hannan and Burton 2001). As a comparison of (23) and (25) (or 24) shows, one consequence of the FE (or FD) transformation is that time-invariant explanatory variables like z_p are swept away along with the unit effects, since the parameters ϕ_p cannot be separately identified from the θ_i. But this property of within estimators can hardly validate the random effects assumption, and hence does not constitute *a priori* grounds for dismissing fixed effect models out of hand. After all, if the random effects assumption is violated, then GLS parameter estimates for both time-varying and time-invariant explanatory variables are contaminated by heterogeneity bias. Nor is the failure of within estimators to identify parameters of time-invariant variables a serious cost when, as for most panel studies, research interest attaches primarily to the *changes* in a response variable brought about by *changes* in explanatory variables.

To eschew within-group estimators because they fail to identify the parameters of time-invariant variables is to underestimate the threat of heterogeneity bias and to misconstrue the principal purpose of panel data. As Wooldridge (pp. 421, 2002) notes, "In most applications, the only reason for collecting panel data is to allow for the unobserved effects [θ_i] to be correlated with the explanatory variables." Similarly, Lee (2002) observes that "the main advantage of panel data is to allow regressors to be related with the error term through [θ_i]. [T]he ability to remove time-invariant unobservable [θ_i] can be the single most important advantage of panel data" (p. 16).

Nor is it the case that a fixed effect approach sacrifices all information about the role of time-invariant explanatory variables in the process of change in a response variable.[13] Within

[13]Variables that are ostensibly time-invariant sometimes have enough time variation to identify even a baseline effect. Years of schooling completed is frequently treated as time invariant for samples of adults, but Angrist and Newey (1991) and Budig and England (2001) found in their fixed effects analyses of the NLSY that even modest changes in schooling yielded precise estimates of the wage return to schooling.

group estimators sweep away only those time invariant regressors with *time invariant* parameters. Time-invariant variables with time-varying parameters are easily handled because neither within transformation eliminates them entirely. Allison and Long's (1990) study of the effects of departmental prestige on scientific productivity and recognition is a nice example of such a specification on time-invariant aspects of academic background. For example, consider the following two-period model for scientific productivity:

$$y_{i1} = \delta_1 + \phi_1 z_i + \gamma x_{i1} + \theta_i + \varepsilon_{i1} \tag{26}$$

$$y_{i2} = \delta_2 + \phi_2 z_i + \gamma x_{i2} + \theta_i + \varepsilon_{i2} \tag{27}$$

$$\tag{28}$$

where x_{it} is prestige of current department and z_i is prestige of Ph.D. department. In this model ϕ_1 is the effect of time-invariant Ph.D. prestige on productivity at $t = 1$ and ϕ_2 is the effect at $t = 2$. The first difference equation is

$$(y_{i2} - y_{i1}) = (\delta_2 - \delta_1) + (\phi_2 - \phi_1)z_i + \gamma(x_{i2} - x_{i1}) + (\varepsilon_{i2} - \varepsilon_{i1}) \tag{29}$$

where now the coefficient of the time-invariant z_i gives the over time change in the effect of Ph.D. prestige on productivity. Although the period-specific parameters ϕ_1 and ϕ_2 are not identified, the change $(\phi_2 - \phi_1)$ is estimable. When the over time pattern of change in parameters is of interest in its own right, a fixed effect approach sacrifices little. Angrist (1995) employs a fixed effects model of wages that easily accommodates parameters for changes in the returns to schooling over time. Since time-invariant explanatory variables can be interacted with time trends or periods, fixed effect specifications also can accommodate terms that show how time-invariant explanatory variables condition the effects of time-varying explanatory variables. For example, Western's (2002) fixed effects analysis of hourly wage growth specifies race effects by fitting separate models for whites, blacks, and Hispanics.[14]

Testing for correlated unit effects

I believe that, on balance, all of the issues pertaining to the quality of estimators and the parameterization of explanatory variables favor fixed effect models and estimators when the goal is to identify and gauge the magnitude of causal parameters. I am not alone. Allison (pp. 181, 1994;) asserts that "the [fixed-effect] estimator is nearly always preferable [to the GLS random effects estimator] for estimating effects....with nonexperimental data." Nickell (p, 1418, 1981), an early proponent of the modern econometric position on the subject, writes: "[I]f one takes the view that, in any particular model, the individual effects are likely to be correlated with all the observed exogenous variables, then one is lead inexorably to the fixed effects model."

Panel data provide internal evidence bearing on the threat of heterogeneity bias, and hence guidance in the choice of model and estimators. Consider a simplified two period case of (23) in which x is the only explanatory variable. Under the exogeneity assumption, all the estimators of γ considered earlier for the two-period case, including $\hat{\gamma}_{dd}$, $\hat{\gamma}_{fd}$, $\hat{\gamma}_{fe}$, are based on within variation, and hence unbiased and consistent because they eliminate heterogeneity bias due to correlated unit effects. Denoting all these estimators as $\hat{\gamma}_w$, we have:

$$E(\hat{\gamma}_w) = \gamma \tag{30}$$

In contrast, estimators that use between-unit variation risk heterogeneity bias. Consider the "between" regression of $\overline{y}_{i.}$ on $\overline{x}_i$ that results by averaging over time. Least-squares estimation of:

$$\overline{y}_{i.} = \alpha + \gamma\overline{x}_{i.} + \theta_i + \overline{\varepsilon}_{i.} \tag{31}$$

[14] Osgood et al. (1996) provide an excellent example of an altogether different approach to accommodating time-invariant variables in a fixed effect panel analysis.

yields the "between" estimator $\hat{\gamma}_b$ of γ. This estimator has expectation:.

$$E(\hat{\gamma}_b) = \gamma + \lambda_{\theta\overline{x}_{i.}} \tag{32}$$

where the second term on the right, the parameter for the regression of θ_i on $\overline{x}_{i.}$, represents the bias in $\hat{\gamma}_b$ due to a correlation between the unit effects θ_i and the over time mean $\overline{x}_{i.}$ of the causal variable.[15] This result suggests a test of the random effects assumption: the difference $(\hat{\gamma}_b - \hat{\gamma}_w)$ gives evidence of correlated unit effects and hence heterogeneity bias.

Exactly this same principle carries over to the contrast $(\hat{\gamma}_{fe} - \hat{\gamma}_{re})$ between the GLS random effects and the FE estimators (Hausman 1978; Baltagi, 1995; Arellano 1993). This difference, which indicates the bias induced in the random effects estimator by an unaccounted for correlation between the unit effects and the explanatory variables, is the basis for a valuable specification test developed by Hausman (1978). Hausman showed how $(\hat{\gamma}_{fe} - \hat{\gamma}_{re})$ could be used to test the null hypothesis that the unit effects and the explanatory variables are uncorrelated (Hausman 1978).[16] Small values of the Hausman statistic fail to reject the null hypothesis and favor GLS estimation of a random effects model on efficiency grounds; large values favor within estimation of a fixed effects model. There is little to recommend a random effects model and GLS estimation without a Hausman test, since the power to detect heterogeneity bias is one of the main strengths of panel data. Nor is the Hausman test restricted to models with classical disturbances. Arrelano (1993) has shown that the Hausman test can be rendered robust to heteroscedasticity and serial correlation, and Metcalf (1996) has extended the test to the case where explanatory variables are correlated with the time-varying disturbance.

[15]Note that heterogeneity bias cannot be traced to within-unit variation in x_{it} around its mean $\overline{x}_{i.}$, since the unit effects θ_i are orthogonal to $(x_{it} - \overline{x}_{i.})$. Further, the correlation between $\overline{x}_{i.}$ and θ_i completely accounts for the correlation between x_{it} and θ_i; indeed, controlling for $\overline{x}_{i.}$ renders $\lambda_{\theta x_{it}} = 0$.

[16]This test should not be confused with the Breusch-Pagan test for the presence of unobserved effects in a random effects model. Breusch-Pagan tests the null hypothesis $\sigma^2_{\theta_i} = 0$ and almost always rejects it (Verbeek 2000 p. 325).

Hausman tests are routine in economics research; impressionistic evidence suggests they typically reject the random effects assumption. Yet the Hausman test is rarely used in sociological research (Table 1, column 10). Not one researcher listed in Table 1 as adopting a random effects model did so on the basis of a Hausman test that failed to reject the random effects assumption. Table 1 reveals the test was performed in only four cases, three of which indicated heterogeneity bias and favored fixed over random effects. Villareal (2002) and Budig and England (2001) adopted FE estimation after a Hausman test rejected the random effects assumption.[17] Gustafson and Johansson (1999) fit random and fixed effects models after a Hausman test failed to detect correlated unit effects. In the fourth case a random effects model was adopted despite a Hausman test that favored fixed effects (Nielsen and Alderson 1995).

Because the parameters of time-invariant variables are not identified by within estimators, it might seem that such variables are not germane to tests of correlated effects; or that such tests are not informative when research aims center on the effect of time-invariant explanatory variables. The opposite is true. Although the Hausman test compares only estimates of the parameters of time-varying explanatory variables, measured time-invariant explanatory variables figure prominently and should always be included in the random effects model. Failure to do so can have a considerable impact on the Hausman statistic. Conversely, including time-invariant control variables in the hope that they may account completely for correlated unit effects is not a valid substitute for a Hausman test and does not suffice to warrant a random effects model. In such cases a Hausman test can shed light on whether measured time-invariant variables have mitigated the threat of heterogeneity bias; and can guage the credibility of heterogeneity bias as an alternative explanation for the coefficients of such variables.

[17] Villareal (2002) used a Hausman test and FE estimation in response to a request from me in my role as deputy editor of *ASR*.

To illustrate the power and insight a Hausman test may provide, Table 2 gives the results of fitting very simple earnings equations to 1980-1987 data from the National Longitudinal Survey Youth Sample.[18] The data are annual observations for $N = 545$ full-time working males who completed their schooling by 1980. Models 1 and 2 give the FE and GLS random effects estimates of an equation containing only time-varying explanatory variables, occupational socioeconomic status and a dummy variable indicating whether the wage is set by collective bargaining (hereafter, "union"). Column 3 gives the differences between the estimates and their corresponding t-ratios, as well as the overall Hausman statistic. For both explanatory variables, the GLS estimates are significantly higher than their FE counterparts, leading to a highly significant chi-square $\chi^2 = 45.32$ and rejection of the random effects assumption. Model 3 adds a time-invariant dummy variable for race (black=1). The differences between the FE and GLS estimates are virtually unchanged, and the Hausman statistic has actually increased to $\chi^2 = 46.75$. The Hausman statistic under model 2 is not picking up heterogeneity bias due to the omission of race. Model 4 adds another time-invariant variable, years of schooling. Schooling accounts for nearly all of the previous gap between the FE and GLS estimates of the coefficient of socioeconomic status. But schooling has virtually no effect on the GLS coefficient for union, and hence the difference in the estimates remains unchanged. The overall reduction in the Hausman statistic between models 3 and 4 ($\chi^2 = 22.13$) is due exclusively to the fact that schooling has eliminated the heterogeneity bias in the GLS estimate of the effect of socioeconomic status on wages. Still, even controlling for two ostensibly important time-invariant variables, the Hausman statistic ($\chi^2 = 24.62$) for Model 4 is highly significant, so the random effects hypothesis is rejected in favor of the inference that correlated individual effects are a source of inconsistency in the GLS estimator.

[18] These data were previously analyzed by Vella and Verbeek (1998). They are discussed by Wooldridge (2000) and available for downloading at *http://ideas.uqam.ca/ideas/data/bocbocins.html.*

Model 5 of Table 2 sheds light on how the Hausman test detects departures from the random effects assumption. Model 5, which is just model 4 augmented by the time-averaged means of socioeconomic status and union, was estimated by GLS under the random effects assumption. Note first that controlling for the time means of status and union purges the previous GLS estimates of Model 4 of heterogeneity bias, leading to the very same FE estimates and the very same t-ratios as seen originally in model 1.[19] Hence, the time means of status and union are accounting for between-unit variation that is the source of heterogeneity bias. Second, the difference between the Wald chi-squares for model 4 and model 5, $\chi^2 = 24.75$, yields a test of the hypothesis that the coefficients of both time means are jointly zero, and gives the same value (rounding error aside) as the Hausman statistic for model 4. This follows because each time-mean coefficient is equal to $(\hat{\gamma}_b - \hat{\gamma}_w)$, and hence estimates the bias component $\lambda_{\theta\overline{x}_{i.}}$ given in (31) for the between estimator.[20]

The Hausman procedure is only one fairly directed test for validating a fixed effects specification. Even if the test favors a within group estimator, the unit effects model of (23) may be misspecified in other respects. Chamberlain (1982) devised a general "omnibus test" of all the restrictions implied by the fixed effects model. Angrist and Newey (1991) provide a simplified method of computing the Chamberlain test, and Arellano (1993) has shown that the Hausman test is a special case. Jakubson (1991) discusses the Chamberlain test and a variety of other tests for the validity of fixed effect models.

Random and fixed effects: a middle ground

Sociological research takes as given a strict distinction between random and fixed effects models.

[19]GLS returns the FE estimates of the time-varying explanatory variables only if the time-averaged means of *all* time-varying explanatory variables are included.

[20]Hausman (1978 p 1263) suggested that comparing the FE and between estimates yields a less powerful test than comparing FE and GLS estimates, but Hausman and Taylor (1981) show that all three paired differences of estimates yield numerically identical results.

But as model 5 of Table 2 suggests, there is a middle ground that yields the advantages of fixed effects while identifying the parameters of time-invariant regressors. This approach, which involves mixing estimators that have the desirable properties of fixed effects for time-varying explanatory variables with random effects estimators for time-invariant explanatory variables, goes to the heart of the resistance many researchers have shown to fixed effect estimation.

To fix ideas, consider a simple version of (23):

$$y_{it} = \phi z_i + \gamma x_{it} + \theta_i + \varepsilon_{it} \tag{33}$$

where z_i and x_{it} are time-invariant and time-varying covariates, respectively. Assume both regressors are strictly exogenous with respect to the disturbance ε_{it}, x_{it} is arbitrarily correlated with the unit effects θ_i, and z_i is uncorrelated with the unit effects. In other words, adopt the random effects assumption for z_i, but not x_{it}. Then this model may be estimated by a simple two-step procedure (Wooldridge 2002 p. 326). The time-demeaning transformation applied to (33) followed by least-squares yields the FE estimator $\hat{\gamma}_{fe}$, but fails to identify the parameter ϕ for the time-invariant variable. To estimate ϕ, fit by GLS the random effects model

$$\hat{\mu}_{it} = \alpha + \phi z_i + \theta_i + e_{it} \tag{34}$$

where the left-hand-side is the residual from FE estimation of (33). Since $\hat{\gamma}_{fe}$ is itself consistent, and since z_i is uncorrelated with both the unit specific and idiosyncratic errors, the random effects estimator $\hat{\phi}_{re}$ will be consistent.[21] Model 2 of Table 3 gives the random effects estimates obtained by this method for race and schooling, and the first-step FE estimates for socioeconomic status and union in the NLSY wage regression. Model 1 repeats the GLS estimates of model 4 from Table 2. The estimates for race and schooling are very close to those yielded by the full random

[21] The random effects estimator $\hat{\phi}_{re}$ will be equivalent to the between estimator $\hat{\phi}_b$ in the time-averaged version of equation 34.

effects GLS estimator in model 4.

This approach overcomes the objection that FE estimation entails the loss of information about the parameters of time-invariant measured variables, but does not circumvent the random effects assumption as it applies to time-invariant explanatory variables. The parameters of time-invariant covariates that are associated with θ_i are still not identified by a regression like (34) because ϕ cannot be distinguished from θ_i. A more powerful method that addresses this weakness while exploiting the advantages of both fixed and random effects approaches is due to Hausman and Taylor (1981). Hausman and Taylor (HT) devised an instrumental variables method that efficiently accommodates time-varying explanatory variables while relaxing the random effects assumption on time-invariant variables.[22]

A simple example gives the essence of the approach. Expand the model of (33) to:

$$y_{it} = \phi_1 z_{1i} + \phi_2 z_{2i} + \gamma_1 x_{1it} + \gamma_2 x_{2it} + \theta_i + \varepsilon_{it} \tag{35}$$

where all explanatory variables are strictly exogenous with respect to the disturbance ε_{it}. Suppose that z_{2i} and x_{2it}, but not z_{1i} and x_{1it}, are correlated with θ_i. In this setup, unbiased and consistent estimation of γ_1 and ϕ_1 is unproblematic: since x_{1it} and z_{1i} are exogenous, they act as their own instruments. Estimation of γ_2 is also unproblematic: as in FE estimation, the deviation $(x_{2it} - \overline{x}_{2i.})$ can be used as an instrument for x_{2it} because it is uncorrelated with θ_i. But an instrument is still required for the time-invariant z_2. Since the correlation of z_2 with θ_i is based on between-unit variation, a proper instrument would be one that evinces between-unit variation, that is time invariant, and that is uncorrelated with θ_i. A variable that fits this description is the mean $\overline{x}_{1i.}$ of

[22]Instrumental variable estimation is a procedure for consistently estimating the parameters of explanatory variables that are correlated with unobservables. An IV is an exogenous variable that is correlated with the explanatory variable of interest, is uncorrelated with the unobservables, and does not appear in the structural equation. Explanatory variables that are not correlated with unobservables are their own instruments.

the exogenous variable in (35).

I computed HT estimates of the NLSY wage regression under the assumption that, conditional on schooling, socioeconomic status and race are exogenous, but that union and schooling are correlated with the unit effects. This differs from the two-step procedure insofar as the random effects assumption with respect to schooling is relaxed, and the time mean of socioeconomic status is used as an instrument for schooling. The parameter estimates, given as model 3 in Table 3, are very close to those given in table 2. Indeed, the estimates for socioeconomic status and union are the very same FE estimates obtained before: the HT method returns FE estimates for all time-varying covariates in a just-identified model.[23] Baltagi (1995 pp. 118-122) gives a a detailed account of the HT method, and Arrelano and Bover (1995) provide a generalization.

Much of the power of HT models is due to the panel design itself, since the time-means of the time-varying explanatory variables may be used as instruments to identify the parameters of time-invariant regressors that are correlated with unit effects. Yet impressionsistic evidence suggests that HT models, though prominent in articles on econometric theory, are rarely applied in a practical research setting; Kim and Polachek (1994) are an exception. One reason may be that deciding which measured variables to treat as uncorrelated with the unit-specific effects introduces an element of arbitrariness. But it is important to recognize degrees of arbitrariness: the HT model is clearly a significant improvement over blindly assuming that all explanatory variables are uncorrelated with the unit effects. A second reason that HT has not been applied more widely may be the lack of a standard command in popular regression packages. This last impediment has

[23] A model is just-identified when the number of exogenous time-varying explanatory variables, say n_1, is equal to the number of time-invariant variables, say n_2, subject to correlated effects. If $n_1 < n_2$, then the model is underidentified. For $n_1 > n_2$, the model is overidentified and the estimators of the parameters of time-varying explanatory variables are more efficient than within group estimators.

recently been eliminated, since Stata 8 includes a dedicated command for HT estimation.

Endogeneity bias in static panel models

Virtually all panel analyses listed in Table 1 assume that the disturbances ε_{it} are mean independent of the explanatory variables ($E(\varepsilon_{it}|x_{it}, z_i) = 0$). Few sociological studies take steps to test for or deal with violations of the exogeneity assumption (see Table 1 column 11). Yet violations of this assumption lead to biased and inconsistent estimators whether unit effects are treated as fixed or random. This section considers tests of and remedies for such violations. Although GLS estimation of random effects models does not preclude additional precautions against violations of the exogeneity assumption (e.g., see Polachek and Kim 1994), I will focus exclusively on methods that might be undertaken following a FE or FD transformation, neither of which eliminates a correlation involving the time-varying disturbance ε_{it}. The spirit of fixed effects models does seem to invite additional effort aimed at detecting and remedying violations of exogeneity.

Correlations involving the disturbances are frequently evidence that one or more explanatory variables may be partially determined by idiosyncratic components of the response variable. For example, Budig and England (2001) hypothesize that unobserved transitory events (i.e., ε_{it}) that lead a woman to anticipate unusually high or low wages may affect the fertility decision and hence the number of children, their main explanatory variable. Or consider the Cherlin et al. (1998) study of divorce and children's mental health. If parents' decisions about their marital futures are partially determined by the disturbance ε_{it} in their children's mental health equation, then divorce will be correlated with ε_{it} and within group methods will yield biased and inconsistent estimates of the effect of divorce on mental health. Similar violations of exogeneity can occur with simple omitted variable bias.

Methods for handling endogeneity bias in panel analyses are fundamentally the same as those used with cross-sectional data. To fix ideas, suppose in (23) the time-varying explanatory variable x_{it} is correlated with the disturbance ε_{it}. Then after either FD or FE transformation, the difference form of x_{it} will be correlated with the difference form of the disturbance, and ordinary least squares will yield biased and inconsistent estimators. The conventional means of dealing with this is to identify one or more valid instrumental variables for x_{it}. Instrumental variables estimation could then be applied to the transformed equations like (24) and (25) to yield the fixed effect (FE-IV) or first difference (FD-IV) instrumental variables estimators.

To understand the issues that arise in the construction of such estimators, it pays to distinguish between strict and sequential exogeneity. Under strict exogeneity, the disturbance ε_{it} is mean independent of all past, current, and future values of x_{it} for all $t = 1, \ldots, T$, which implies that the mean of the response variable is independent of all past and future values of x_{it}, i.e., depends only on the contemporaneous value x_{it} given the unit effects. Sequential exogeneity is weaker: the disturbance ε_{it} is independent of the current and all past values of x_{is} for all $s \leq t$, in which case x_{it} is *predetermined*. This too implies that no past values of the explanatory variables affect the current mean of the response variable after controlling for current x_{it} and the unit effects θ_i.

These exogeneity restrictions highlight a difference between the FE and FD estimators that is relevant for instrumental variables estimation. For consistency, the FE estimator requires that the mean deviated form of the explanatory variables be uncorrelated with the mean-deviated form of the disturbance. Since the mean $\overline{x}_{i.}$ involves all past, current and future values of x_{it}, this condition is only met under strict exogeneity. When strict exogeneity is violated, the inconsistency in the FE estimator declines as T increases, and may be very small for large T. In contrast, the first differences of the explanatory variables and the disturbance will be mean independent under the weaker sequential exogeneity restriction. While this would seem to favor FD over FE, Wooldridge (2002 p. 302) notes that FE might be preferred because it can have less bias for large N, and because the inconsistency of the FD estimator does not decline as T increases.

The difference in the exogeneity restrictions required for consistency means that the FE estimator is somewhat weaker than the FD estimator for dealing with endogeneity bias. Assume that x_{it} with parameter γ in (23) violates strict exogeneity (but z_i and w_k are strictly exogenous). Then the strict exogeneity requirement of the fixed effect estimator rules out lagged values of the endogenous regressor, $x_{it-1}, x_{it-2}, \ldots, x_{i1}$, as valid instruments. Only strictly exogenous variables that are excluded from the equation are valid instruments. For example, if v_{it} is an excluded exogenous variable, then $(v_{it} - \overline{v}_{i.})$ would be a valid instrument for $(x_{it} - \overline{x}_{i.})$. In the event that more than two valid instruments are available for one endogenous regressor, 2SLS estimation of the mean-deviated equation (25), with all included and excluded exogenous variables treated as instruments, will yield a consistent estimator of γ. External instruments can also be used with the FD transformation, but the weaker exogeneity assumption means that lagged values of the endogenous regressor itself, $x_{it-1}, x_{t-2}, \ldots, x_{i1}$ are also available as valid instruments when $T \geq 3$. Hence, lagged levels or lagged differences of the endogenous regressor are valid instruments in the first difference equation (24). For example, $(x_{it} - x_{it-1})$ in (24) could be instrumented by x_{it-1} and x_{it-2}, or just by $(x_{it-1} - x_{it-2})$.

Instrumental variables estimation following a within-group transformation is common in economics research. Cornwell and Trumbull (1994) employ a FE-IV estimator in their study of the effectiveness of criminal justice strategies and law enforcement incentives as crime deterrents; Evans, Froeb and Werden (1993) use FE-IV in their analysis of the effect of concentration on fares in the airlines industry. FE-IV estimation is also used by Blau, Guilkey and Popkin (1996) and Kim and Polachek (1994). A study that employed the FD-IV estimator is Holt-Eakin's (1994) analysis of the effect of public sector capital accumulation on private sector productivity in American states between 1969-1986. Since a typical regressor in the first difference equation has the form $\triangle x_{it} = (x_{it} - x_{it-1})$, Holtz-Eakin uses as an instrument the twice lagged version $\triangle x_{it-2}$.

Instrumental variables estimation is a viable approach to violations of the exogeneity assumption, but is not without a price. The FE-IV and FD-IV estimators will be less efficient than the corresponding least squares estimator in a model with uncorrelated disturbances. For this reason it is useful to test for exogeneity to determine if IV estimation is necessary. Just as one would ordinarily carry out a Hausman test of the hypothesis that the unit effects θ_i are uncorrelated with the explanatory variables, so too one could employ a Hausman test of the hypothesis that an explanatory variable is uncorrelated with the time-varying disturbance in the mean-deviation or first-difference equation. Here again the basic idea is the same: if all the explanatory variables are exogenous, the FE-IV and FE-LS parameter estimates should be similar, since both would be consistent. If the two estimators yield significantly different estimates, it suggests the FE-LS estimator is inconsistent because of endogeneity bias.

Consider a test of the hypothesis that $(x_{it} - \overline{x}_{it-1})$ is uncorrelated with $(\varepsilon_{it} - \overline{\varepsilon}_{it-1})$ in (25). This is equivalent to a test of significance for the difference $(\gamma_{fels} - \gamma_{feiv})$. This test can be constructed from two least-squares regressions (Maddalla 1988, pp. 435-441; Wooldridge 2000, pp. 483-486). Let v_{it} be a valid instrument for x_{it}. Then the first step involves regressing $(x_{it} - \overline{x}_{it-1})$ on the mean deviations of v_{it} and all the exogenous variables w_{kit} in (25). Since the fitted values from this regression are a linear combination of exogenous variables, the variation in x_{it} that is a potential source of endogeneity bias must be in the residuals, say $\hat{u}_{it}$. In the second step, fit the original mean-deviated equation (25) with $\hat{u}_{it}$ included as an additional regressor:

$$(y_{it} - \overline{y}_i) = \sum_k \beta_k (w_{kit} - \overline{w}_{ki}) + \gamma (x_{it} - \overline{x}_i) + \lambda \hat{u}_{it} + e_{it} \tag{36}$$

The test of exogeneity is simply a t-test of $\lambda = 0$.[24] Rejecting this hypothesis suggests that x_{it} be treated as endogenous and the equation estimated by FE-IV; failing to reject implies that FE-LS

[24] For the case of several endogenous explanatory variables, see Maddala 2001 pp. 498-500.

will yield a consistent and more efficient estimate of γ. Applications of this test may be found in Evans, Froeb and Werden (1993) and Kim and Polacheck (1994).

III. Dynamic panel models

Most panel models appearing in sociological studies are static insofar as all explanatory variables are dated contemporaneously with the response variable. Yet dynamic models in which a lagged dependent variable appears as a regressor account for about one-third of the papers listed in Table 1 (column 12). The bulk of the dynamic regressions found in sociology take one of two basic forms. One form is

$$y_{it} = b_0 + b_1 y_{it-1} + b_2 z_i + b_3 x_{it} + v_{it} \tag{37}$$

where the endogenous variable is lagged, but other variables are dated contemporaneously with the dependent variable. This, the standard dynamic regression, was used by Pampel (1994), Podolny et al. (1996), McManus and DiPrete (2001), and Kenworthy (2002). A second form is:

$$y_{it} = b_0 + b_1 y_{it-1} + b_2 z_i + b_3 x_{it-1} + v_{it} \tag{38}$$

where now the time-varying explanatory variable is also lagged. Variations on this form were employed by Pampel (1993), Shofer, et al. (2000), Jenkins and Scanlan (2001), Hagan and Foster (2001), and Beckfield (2003). More elaborate lag structures on both the endogenous and exogenous variables can be found in practice (Western and Beckett 1999; Wilson et al. 1997).

Many researchers are notably inattentive to the issues raised by the formulation and estimation of dynamic models. Indeed, the distinction between model formulation and estimation is itself easily eclipsed in the area of dynamics. Relevant here is the contrast between models with lagged dependent variables that are generated by error dynamics in a static structural equation, and models with lagged dependent variables that represent the hypothesis of true state dependence. Models of the former type reflect ”spurious” state dependence because the coefficient of the lagged endogenous variable is due not to a direct causal effect, but rather to the persistence of unobservables that determine both lagged and contemporaneous values of the response variable. Models of true state dependence, in contrast, reflect a causal effect of past values of the response variable on current values. A methodological treatment of this distinction is given by Allison (1990), but it is rarely brought to bear in applied work in sociology. Rather, the distinction is implicit in the contrast between researchers who introduce lagged endogenous variables as a solution to estimation problems caused by unobservables in an otherwise static model, and those who posit true state dependance as a property of the structural model itself.

In the sections below, I attend briefly to dynamic regressions that reflect spurious state dependence, and then turn to the estimation of dynamic models that posit true state dependence. In neither case have sociological studies paid adequate attention to the problems of estimation posed by dynamic models with lagged endogenous variables. The focus throughout is on models with unobserved fixed effects that are correlated with the explanatory variables.

Unobservable dynamics in static models

Consider first the case in which a lagged dependent variable is included to remedy least-squares estimation problems due to unobservable error dynamics. Let the static model of interest be

$$y_{it} = \phi z_i + \gamma x_{it} + \theta_i + \epsilon_{it} \tag{39}$$

with a classical disturbance and correlated unit effects. In this model the only source of over time dynamics in the composite error $u_{it} = \theta_i + \epsilon_{it}$ is the unit effect. The dynamic form of this model is obtained by lagging (39) one period and subtracting the result from period t:

$$y_{it} = y_{it-1} + \gamma(x_{it} - x_{it-1}) + (\epsilon_{it} - \epsilon_{it-1}) \tag{40}$$

where the error dynamics have yielded a model that includes a lagged dependent variable with coefficient constrained to 1. But this is just the first-difference equation seen earlier for a static model. In other words, first-differencing yields a dynamic regression that eliminates the threat of heterogeneity bias in estimating a fundamentally static model. Yet neither (37) nor (38) have the coefficient of y_{it-1} constrained to 1 and time-varying explanatory variables appearing in first-difference form. Firebaugh and Beck (1994) made exactly this point in criticizing cross-national research based on (38); the criticism also applies to (37). Regressions like (37) and (38) do not solve the estimation problems posed by unobserved heterogeneity in a static model like (39).

An alternative approach is to introduce error dynamics through the disturbances of (39). Let

$$\epsilon_{it} = \rho\epsilon_{it-1} + u_{it} \tag{41}$$

where u_{it} is classical. Then the differencing procedure used above yields:

$$y_{it} = \rho y_{it-1} + (1-\rho)\phi z_i + \gamma x_{it} - \rho\gamma x_{it-1} + (1-\rho)\theta_i + u_{it} \tag{42}$$

where the appearance of y_{it-1} and x_{it-1} are accounted for by the serial dynamics induced by ϵ_{it-1}, leaving u_{it} serially uncorrelated. This is a dynamic regression that solves the problem of serially correlated disturbances in a static behavioral model. Yet neither of the prototype dynamic regressions found in sociological studies have this form. Nor is this model itself adequate, since it fails to account for unobserved heterogeneity. As discussed below, least squares estimation of a model with lagged endogenous variable and unobserved heterogeneity is problematic.

Unobservables in system dynamic panel models

Consider now a model of true state dependence in which earlier changes in the explanatory variables effect later distributions of the endogenous variable through the lagged endogenous variable y_{it-1}. The structural equation of the state dependence model is typically written as:

$$y_{it} = \varphi y_{it-1} + \phi z_i + \gamma x_{it} + \theta_i + \varepsilon_{it} \tag{43}$$

where ε_{it} is mean zero, constant variance, and independently distributed. This is a model of true rather than spurious state dependence because the lagged dependent variable appears not as a result of error dynamics generated by unobservables, but rather as a mechanism for transmitting to the current value of the response variable the effects of past changes in the exogenous variables. Models like (37) and (38) usually appear in sociological studies as expressions of true state dependence, although the formulation is rarely explicit on this point. Few studies rigorously develop theoretical grounds for a model of state dependence—Firebaugh and Beck (1994) make a similar point—let alone sharply distinguish it from a static model with serially correlated unobservables.[25] Studies of growth rates are one instance in which the state dependence model is explicitly adopted on theoretical grounds, as in Podolny Stuart, and Hannan's (1996) study of semiconductor sales and Sutton's (2000) cross-national study of imprisonment.

Estimating (43) is more problematic than estimating the same model without the lagged endogenous variable. Consider first the case in which the time-invariant unit effects are *uncorrelated* with the exogenous variables. In this ostensibly innocent case, the least-squares estimator is badly biased because of the correlation between the lagged endogenous variable y_{it-1} and the unit effects. This correlation is implied by the model itself: since θ_i effects y_{it}, it also effects y_{it-1}. Hence, the random effects assumption cannot hold with respect to the lagged endogenous variable; indeed, the model is formulated so that it does not hold, making tests of correlated effects moot.

The direction of the bias induced by the correlation of the lagged endogenous variable and the unit effects is well known. Least-squares is upwardly biased and inconsistent away from zero

[25] Anderson and Hsaio (1982) and Maddalla (1987) suggest a test to distinguish the spurious from true state dependence model. The test exploits the restriction imposed by the serial correlation, but not the state dependence, model: the product of the coefficient of y_{it-1} and the coefficient of x_{it} (i.e., $-\rho\gamma$ in 42) should be equal in magnitude but opposite in sign to the coefficient (i.e., $\rho\gamma$) of x_{it-1}.

for the coefficient φ of the lagged endogenous variable, with greater bias as the variance of the unobserved individual effects increases (Hsaio p. 77); and is downwardly biased toward zero for coefficients, like ϕ and γ, of the exogenous variables. There are many cases in the sociological literature where the random effects assumption is adopted and least-squares applied directly to a model like (43) (or 38) (e.g., Shofer, et al. 2000; Kenworthy 2002; Beckfield 2003).

The bias and inconsistency in the least-squares estimator of the parameters of the exogenous variables is aggravated when these variables are correlated with the unit effects. Applying the fixed effects transformation to eliminate unit effects, the time-demeaned equation is

$$(y_{it} - \overline{y}_i) = \varphi(y_{it-1} - \overline{y}_{i.-1}) + \gamma(x_{it} - \overline{x}_i) + (\varepsilon_{it} - \overline{\varepsilon}_i) \tag{44}$$

where $y_{i.-1} = (1/T)\sum_0^{t-1} y_{it-1}$. Least squares applied to this equation yields the FE estimator used by Podolny et.al. (1996). Alas, this estimator too is biased and inconsistent for all parameters. When $\varphi > 0$, the bias in the FE estimator is invariably negative and *larger* when the model includes exogenous variables, the typical case. With regard to the parameters of the exogenous variables, the FE estimator is upwardly biased for coefficients of exogenous variables that are positively related to the lagged endogenous variable; and downwardly biased for coefficients of exogenous variables that are negatively related to the lagged endogenous variable. Magnitudes of bias, which can be substantial with T small, are given by Nickell (1981 p. 1495) and Verbeek (2000 p. 328).

What is the source of bias in the fixed effects estimator? The error term $(\varepsilon_{it} - \overline{\varepsilon}_i)$ in the transformed equation is correlated with the endogenous variable term $(y_{it-1} - \overline{y}_{i-1})$ because $\overline{\varepsilon}_i$ is correlated with y_{it-1} (Nickell 1981; Baltagi, 1995 pp. 126). The bias goes to zero as T increases, but this is not much use because, as Table 1 shows, T is typically small. The first difference transformation (e.g., Sutton 2000) does not circumvent these problems because the difference term in the lagged endogenous variable will be correlated with the difference term in the disturbance.

A great deal of econometric research has gone into obtaining consistent and efficient estimators of the parameters of dynamic panel models with fixed effects. Nearly all approaches involve first transforming the original equation to eliminate the unit effects, and then applying instrumental variables estimation for the parameter of the lagged endogenous variable. An early method for obtaining consistent FE-IV estimators is due to Anderson and Hsaio (1982). Their method applies first differences to deal with the individual effects,

$$(y_{it} - y_{it-1}) = \varphi(y_{it-1} - y_{it-2}) + \gamma(x_{it} - x_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}) \tag{45}$$

and then uses y_{it-2} as an instrument for $(y_{it-1} - y_{it-2})$.The variable y_{it-2} is a valid instrument because it is correlated with $(y_{it-1} - y_{it-2})$ and is uncorrelated with $(\varepsilon_{it} - \varepsilon_{it-1})$ (if the original ε_{it} are not serially correlated). Anderson and Hsaio also suggested the twice lagged difference $(y_{it-2} - y_{it-3})$ as another valid instrument for $(y_{it-1} - y_{it-2})$, but Monte Carlo studies have shown that such an estimator is more biased and less efficient than one that uses as instruments the twice-lagged levels y_{it-2} (Arellano and Bond 1991; Kiviet 1995).

Among the studies listed in Table 1, only McManus's and DiPrete's (2001) dynamic analysis of the economic consequences of divorce and separation for men uses instrumental variable estimation. McManus and DiPrete begin with the model:

$$y_{it} = \alpha + \gamma x_{it} + \varepsilon_{it} \tag{46}$$

which they then difference to yield:

$$(y_{it} - y_{it-1}) = \gamma(x_{it} - x_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}) \tag{47}$$

where y_{it} is a measure of economic well-being. They then conjecture that the change in economic welfare from $t-1$ to t depends on the level of economic welfare at $t-1$. To accommodate this idea, they add y_{it-1} to the difference equation (47) to give:

$$(y_{it} - y_{it-1}) = \varphi y_{it-1} + \gamma(x_{it} - x_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}) \tag{48}$$

which they then fit using lagged values y_{it-2} as instruments for y_{it-1}. An alternative approach would have been to include the lagged endogenous variable in the original model of (46), yielding:

$$y_{it} = \varphi y_{it-1} + \gamma x_{it} + \theta_i + \varepsilon_{it} \tag{49}$$

where I have added unit effects. This, the standard dynamic panel model, also expresses the idea that the change in economic well being from $t-1$ to t depends on the level of economic well-being at $t-1$.[26] Applying the first difference operator to (49) yields:

$$(y_{it} - y_{it-1}) = \varphi(y_{it-1} - y_{it-2}) + \gamma(x_{it} - x_{it-1}) + (e_{it} - e_{it-1}) \tag{50}$$

which is the same as (45). This is not the model estimated by McManus and DiPrete, but it too may be fitted by the Anderson and Hsaio method, with the twice lagged y_{it-2} used as an instrument for $(y_{it-1} - y_{it-2})$.

Other approaches to instrumental variable estimation of dynamic panel models have developed as alternatives to Anderson-Hsaio. Most notable is the generalized method of moments (GMM) approach pioneered by Arrellano and Bond (1991) and later extended by Arrelano and Bover (1995) in the context of HT models. In an effort to improve efficiency, Arrelano and Bond (1991) developed two estimators that include among potential instruments not just lagged levels of the endogenous variable, like Anderson and Hsaio (1982) and Holtz-Eakin, Newey, and Rosen (1988), but also lagged levels and differences of predetermined exogenous variables and strictly exogenous variables, respectively. One estimator, GMM1, assumes that the idiosyncratic errors are homoscedastic and serially uncorrelated, while the second, GMM2, is robust. The two estimators are asymptotically

[26]Rewrite 49 as $(y_{it} - y_{it-1}) = \vartheta y_{it-1} + \gamma x_{it} + \theta_i + \varepsilon_{it}$, where $\vartheta = (\varphi - 1)$.

equivalent, but Monte Carlo studies have shown that GMM1 tends to outperform GMM2 (Kiviet 1995; Judson and Owen 1999). Baltaggi (1995 Chapter 8) gives a detailed exposition of the work of Arrelano and Bond (1991) and Arrelano and Bover (993).

Should researchers use the Anderson-Hsaio or Arrelano-Bond estimator? Arellano and Bond (1991) report Monte Carlo analyses that show that GMM1 is much more efficient than Anderson-Hsaio for the coefficient of the lagged endogenous variable when the true parameter is large, say, $\varphi \geq .80$. Yet for other values of φ, inspection of their findings reveals that the bias and standard errors of the Anderson-Hsaio estimator for the coefficients of both the lagged dependent variable and the exogenous variables are very similar to those of GMM1. Indeed, Kiviet (1995 p. 70), using a different Monte Carlo protocol, reports that the Anderson-Hsaio estimator "shows smaller bias than GMM1 and its efficiency compares quite favorably." Judson and Owen (1999) also report that the Anderson-Hsaio estimator compares favorably to the Arellano-Bond GMM1, especially for large T. One similarity worth noting is that both estimators use lagged endogenous variables as instruments, and hence rely for consistency on the original disturbances being serially uncorrelated; or, equivalently, the absence of second order serial correlation of the disturbances in the difference equation. Arellano and Bond (1991) derived tests for this important condition.

The Anderson-Hsaio and Arellano-Bond IV methods are now standard approaches to estimating the parameters of dynamic panel models with fixed effects. Nonetheless, it is good to keep in mind Kiviet's (1995 pp. 72) observation:

> [F]or dynamic panel data models the use of instrumental variables estimation methods may lead to poor finite sample efficiency. ...As yet, no technique is available that has shown uniform superiority in finite samples over a wide range of relevant situations as far as the true parameter values and the further properties of the data generating mechanism is concerned.

For this reason, it makes sense to experiment with both estimators when dealing with dynamic panel models.[27] As a practical matter this is possible because dedicated commands for both estimators are available in some statistical software, including Stata 7 and 8.[28]

IV. Conclusion

Appearances not withstanding, I have found that the studies identified in Table 1 yield a more impressive portrait of the state of panel data research in sociology than I had expected when first embarking upon this exercise. Yet these studies also indicate that sociologists have only just begun to exploit the power of panel data. The primary goal of this review has been to encourage sociologists to capitalize on the opportunities that panel data offer for securing the validity of causal inferences. To this end, I have described the central problems posed by unobservables and presented a range of simple methods for tackling them. Although I have pressed the advantages of a fixed effects approach, this has been motivated as much be the desire to restore balance as by the very real strengths such an approach has for ruling out alternative effects that may confound causal inference. But Duncan's caution issued more three decades ago is well worth repeating: "[T]he making of a causal inference is not a simple affair that can be reduced to a formula applied mechanically to a set of panel data on two or more variables" (1972 pp. 36). Or perhaps it is more accurate to observe that causal inference cannot be reduced to any *one* formula applied to data. Because causal inference from observational data is by its nature precarious, it pays to experiment with the host of basic techniques that panel data make available and that this article has surveyed.

---

[27]Kiviet (1995) introduced some new estimators that, judging from his Monte Carlo results and those of Judson and Owens (1999), have good properties.

[28]The Anderson-Hsaio FD-IV estimator is implemented in Stata's *xtivreg* command, and the Arellano-Bond estimators GMM1 and GMM2 can be obtained with the *xtabond* command. The latter command also offers the specification tests Arellano and Bond (1991) proposed for serial correlation in the disturbances.

References

Abadie A. 2002. Semiparametric difference-in-differences estimators. Cambridge, Mass.: Dep. Econ., Harv. Univ.

Alderson AS, Nielsen F. 2002. Globalization and the great U-turn: income inequality trends in 16 OECD countries. *Am. J. Sociol.* 107:1244-99

Allison PD. 1990. Change scores as dependent variables in regression analysis. *Sociol. Meth.* 20:93-114

Allison P. 1994. Using panel data to estimate the effects of events. *Sociol. Meth. Res.* 23:174-99

Allison P, Long JS. 1990. Departmental effects on scientific productivity. *Am. Sociol. Rev.* 55:469-78

Angrist JD, Krueger AB. 2001. Instrumental variables and the search for identification: from supply and demand to natural experiments. Cambridge, Mass.: Nat. Bur. Econ Res.

Angrist JD, Newey WK. 1991. Over-identification tests in earnings functions with fixed effects. *J. Bus. Stat.* 9:317-24

Angrist JD. 1995. The economic returns to schooling in the West Bank and Gaza Strip. *Am. Econ. Rev.* 85:1065-87

Arellano M. 2003. *Panel Data Econometrics.* Oxford: Oxford University Press. 231 pp.

Arellano M, Bover O. 1995. Another look at the instrumental variable estimation of error-componenets models. *J. Econmet.* 68:29-51

Arellano M. 1993. On the testing of correlated effects in panel data." *J. Economet.* 59: 87-97

Arellano M, Bond S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employmnet equations. *Rev. Econ. Stud.* 58:277-97

Ashenfelter O, Card D. 1985. Using the longitudinal structure of earnings to estimate the effects of training programs. *Rev. Econ Stat.* 67:648-60

Athey S, Imbens G. 2002. Identification and inference in nonlinear difference-in-differences models. Palo Alto, Ca.: Stanford Univ.

Baltagi BH. 2001. *Econometric Analysis of Panel Data.* Chichester:John Wiley and Sons. 257 pp.

Baron JN, Hannan MT, Burton MD. 2001. Labor pains: change in organizaitonal models and employee turnover in young, high-tech firms. *Am. J. of Sociol.* 106:960-1012

Beckfield J. 2003. Inequality in the world polity: the structure of international organization. *Am. Sociol. Rev.* 68:401-24

Bertrand M, Duflo E, Mullainathan S. 2002. How much should we trust difference-in-differences estimates? Cambridge, Mass.: Nat. Bur. Econ Res.

Blau DM, Guilkey DK, Popkin BM. 1996. Infant health and the labor supply of mothers. *J. Hum. Res.* 31:90-139

Boone C, Carroll GR, van Witteloostuijn A. 2002. Resource distributions and market partitioning: Dutch daily newspapers, 1968-1994. *Am. Sociol. Rev.* 67:408-31

Budig MJ, England P. 2001. The wage penalty for motherhood. *Am. Sociol. Rev.* 66: 204-25

Campbell DT, Stanley JC. 1963. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally. 84 pp.

Card D, Krueger AB. 1994. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *Amer. Econ. Rev.* 84:772-93

Chamberlain G. Multivariate regression models for panel data. *J. Economet.* 22:5-46

Cherlin AJ, Lindsay Chase-Lansdale P, McRae C. 1998. Effects of parental divorce on mental health throughout the life course. *Am. Sociol. Rev.* 63:239-49

Cornwell C, Trumbull WN. 1994. Estimating the economic model of crime with panel data. *Rev. Econ. Stat.* 76:360-66

Darnell A, Sherkat DE. 1997. The impact of Protestant fundamentalism on educational attainment. *Am. Sociol. Rev.* 67:306-15

Duncan OD. 1972. Unmeasured variables in linear panel models. *Sociol. Method.* 4:36-82

Evans WN, Froeb LM, Werden GJ. 1993. Endogeneity in the concentration-price relationship: causes, consequences, and cures. *J. Indust. Econ.* 41:431-38

Firebaugh G, Beck FD. 1994. Does economic growth benefit the masses? growth, dependence, and welfare in the third world. *Am. Sociol. Rev.* 59:631-53

Gustafsson B, Johansson M. 1999. In search of smoking guns: what makes income inequality vary over time in different countries. *Am. Sociol. Rev.* 64:585-605

Hagan J, Foster H. 2001. Youth violence and the end of adolescence. *Am. Sociol. Rev.* 66:874-99

Hannan MT, Tuma N. 1979. Methods of temporal analysis. *Ann. Rev. Sociol.* 1979:303-28

Hausman JA. 1978. Specification tests in econometrics. *Economet.* 46:1251-72

Hausman JA, Taylor WE. 1981. Panel data and unobservable individual effects. *Economet.* 49:1377-98

Holland P. 1986. Statistics and causal inference. *J. Am. Stat. Assoc.* 81:945-60

Heckman JJ, Ichimura H, Todd PE. 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. *Rev. Econ Stud.* 64:605-54

Holtz-Eakin D. 1994. Public-sector capital and the productivity puzzle. *Rev. Econ. Stat.* 76:12-21

Hsaio C. 1986. *Analysis of Panel Data.* New York: Cambridge University Press. 246 pp.

Huber E, Ragin C, Stephens JD. 1993. Social democracy, Christian democracy, constitutional structure, and the welfare state. *Am. J. of Sociol.* 99:711-749

Huber E, Stephens JD. 2000. Partisan governance, women's employment, and the social democratic service state. *Am. Sociol. Rev.* 65:323-42

Jakubson G. 1991. Estimation and testing of union wage effect using panel data. *Rev. Econ. Stud.* 58:971-91

Jenkins JC, Scanlan SJ. 2001. Food security in less developed countries, 1970-1990. *Am. Sociol. Rev.*66:718-44

Judson RA, Owen AL. 1999. Estimating dynamic panel data models: a guide for macroeconomists. *Econ. Lett.* 65:9-15

Kenworthy L. 2002. Corporatism and unemployment in the 1980s and 1990s. *Am. Sociol. Rev.* 67:367-88

Kilbourne BS, Farkas G, Beron K, England P. 1994. Returns to skill, compensating differentials, and gender bias: effects of occupational characteristics on the wages of white women and men. *Am. J. Sociol.* 100:689-19

Kim MK, Polachek SW. 1994. Panel estimates of male and female earnings functions. *J. Hum. Res.* 29:406-28

Lee M. 2002. *Panel Data Econometrics: Methods-of-Moments and Limited Dependent Variables.* San Diego: Academic Press. 195 pp.

Lichter DT, McLaughlin DK, Ribar DC. 1997. Welfare and the rise in female-headed families *Am. J. Sociol.* 103:112-43

Maddala, GS. 1987. Limited dependent variable models using panel data. *J. Hum. Res.* 22:307-38

Maddala, GS. 2001. *Introduction to Econometrics.* New York: Macmillan. 636 pp.

McManus, PA, DiPrete TA. 2001. Losers and winners: the financial consequences of separation and divorce for men. *Am. Sociol. Rev.*66:246-68

Metcalf GE. 1996. Specification testing in panel data with instrumental variables. *J. Economet.* 71:291-307

Meyer BD. 1995. Natural and quasi-experiemtns in Economics. J. Bus. Econ. Stat 13:151-61

Meyer BD, Viscusi WK, Durbin DL. 1995. Workers' compensation and injury duration: evidence from a natural experiment. *Amer. Econ Rev.* 85:322-340

Moller S, Bradley D, Huber E, Nielsen F, Stephens JD. 2003. Determinants of relative poverty in advanced capitalist democracies. *Am. Sociol. Rev.* 68:22-51

Morrison DR, Ritualo A. 2000. Routes to children's recovery after divorce: are cohabitation and remarriage equivalent? *Am. Sociol. Rev.* 65:560-80

Mundlak, Y. 1978. On the pooling of time series and cross section Data. *Economet.* 46:69-85

Nickell, S. 1981. Biases in dynamic models with fixed effects. *Economet.* 49: 1417-26

Nielsen F, Alderson AS. 1995. Income inequality, development, and dualism: results from an unbalanced cross-national panel. *Am. Sociol. Rev.* 60:674-701

Osgood DW, Wilson JK, O'Malley PM, Bachman JG, Johnston LD. 1996. Routine activities and individual deviant behavior. *Am. Sociol. Rev.* 61:635-55

Pampel FC. 1994. Population aging, class context, and age inequality in public spending. *Am. J. Sociol.* 100:153-95

Podolny JM, Stuart TE, Hannan MT. 1996. Networks, knowledge, and niches: competition in the worldwide semiconductor industry, 1984-1991. *Am. J. Sociol.* 102:659-89

Schofer E, Ramirez FO, Meyer JW. 2000. The effects of science on national economic development, 1970-1990. *Am. Sociol. Rev.* 65:866-87

Sutton JR. Imprisonment and social classification in five common-law democracies, 1955-1985. *Am. J. Sociol.* 106:350-86

Waldfogel J. 1997. The effects of children on women's wages." *Am. Sociol. Rev.* 62:209-17

Western B, Beckett K. 1999. How unregulated is the u.s. labor market? the penal system as a labor market institution. *Am. J. Sociol.* 104:1030-60

Western, B. The impact of incarceration on wage mobility and inequality. *Am. Sociol. Rev.* 67:526-546

Villareal A. 2002. Political competition and violence in Mexico: hierarchical social control in local patronage structures. *Am. Sociol. Rev.* 67:477-98

Wilson J, Musick M. 1997. Who cares? Toward an integrated theory of volunteer work. *Am. Sociol. Rev.* 62:694-713

Wooldridge JM. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Mass.: The MIT Press. 752 pp.

Wooldridge JM. 2003. *I*ntroductory Econometrics: A Modern Approach. South-Western College Publishing. 863 pp.

Zhou X. 2000. Economic transformation and income inequality in urban China: evidence from panel data. *Am. J. Sociol.* 105:1135-74

Table 1. Characteristics of Selected Panel Studies from the *American Sociological Review* and the *American Journal of Sociology*, 1990-2003

author(s)	*journal (1)*	*year (2)*	*N (3)*	*T (4)*	*model (5)*	*correlated unit effects addressed? (6)*	*fe/fd (7)*	*estimator (8)*	*st. errors (9)*	*Hausman test (10)*	*exogeneity assumption addressed? (11)*	*lagged endogenous? (12)*	*lagged exogenous? (13)*
Allison & Long	*ASR*	1990	173	2	fe	no	fe/fd	ml	ml	no	no	no	no
Pampel	*AJS*	1994	18	38	re/fe	no	fe	gls/ fe	gls/ols	no	no	yes	yes
Firebaugh & Beck	*ASR*	1994	62	23	fe	no	fd	ols	ols	no	yes	no	yes
Kilbourne et al.	*AJS*	1994	5000	10	fe	no	fe	ols	ols	no	no	no	no
Neilsen & Alderson	*ASR*	1995	88	3	re	no	n/a	gls	gls	yes	no	no	no
Osgood et al.	*ASR*	1996	1800	5	fe	no	fe	ols	ols	no	no	no	no
Podolny et al.	*AJS*	1996	113	6	fe	no	fe	ols	ols	no	no	yes	no
Waldfogel	*ASR*	1997	2000	15	fe	no	fe/fd	ols	ols	no	no	no	no
Darnell & Sherkat	*ASR*	1997	1400	3	re	no	n/a	ml	ml	no	no	no	no
Lichter et al.	*AJS*	1997	3053	2	fe	no	fe/fd	ols	ols	no	no	no	no
Wilson & Musick	*ASR*	1997	2900	2	re	no	n/a	ml	ml	no	no	yes	no
Cherlin et al.	*ASR*	1998	12759	5	re/fe	no	fe	ml/ols	ml/ols	no	no	no	no
Western & Beckett	*AJS*	1999	2917	10	re	no	n/a	gls	gls	no	no	yes	yes
Gustafson & Johansson	*ASR*	1999	16	various	re/fe	yes	fe	gls/ols	ols	yes	no	no	no

Table 1. Characteristics of Selected Panel Studies from the *American Sociological Review* and the *American Journal of Sociology*, 1990-2003

author(s)	*journal* (1)	*year* (2)	*N* (3)	*T* (4)	*model* (5)	*correlated unit effects addressed?* (6)	*fe/fd* (7)	*estimator* (8)	*st. errors* (9)	*Hausman test* (10)	*exogeneity assumption addressed?* (11)	*lagged endogenous?* (12)	*lagged exogenous?* (13)
Zhou	*AJS*	2000	4730	11	re	no	n/a	ml	ml	no	no	no	no
Huber & Stephens	*ASR*	2000	16	25	re	no	n/a	gls	pcse	no	no	no	yes
Morrison & Ritualo	*ASR*	2000	1377	14	fe	no	fe	ols	robust	no	no	no	no
Sutton	*AJS*	2000	5	31	fe	no	fd	ols	pcse	no	no	yes	yes
Schofer et al.	*ASR*	2000	80-112	2	re	no	n/a	ols	ols	no	no	yes	yes
Baron et al.	*AJS*	2001	101	4	re	no	n/a	gee	robust	no	no	no	no
Budig & England	*ASR*	2001	5287	8	fe	yes	fe	ols	robust	yes	no	no	no
Mcmanus & Diprete	*ASR*	2001	1983	14	fe	no	fd	iv	robust	no	no	yes	no
Jenkins & Scanlan	*ASR*	2001	88	2	re	no	n/a	ols	ols	no	no	yes	yes
Hagan & Forster	*ASR*	2001	5000	2	re	no	n/a	ols	ols	no	no	yes	yes
Alderson & Nielsen	*AJS*	2002	16	~12	re	no	n/a	gee	robust	no	no	no	no
Kenworthy	*ASR*	2002	16	18	re	no	n/a	ols	pcse	no	no	yes	no
Boone et al.	*ASR*	2002	11	27	re	no	n/a	gls	gls	no	no	no	no
Westem	*ASR*	2002	5400	16	re/fe	no	fe	ols	ols	no	no	no	no

Table 1. Characteristics of Selected Panel Studies from the *American Sociological Review* and the *American Journal of Sociology*, 1990-2003

author(s)	*journal (1)*	*year (2)*	*N (3)*	*T (4)*	*model (5)*	*correlated unit effects addressed? (6)*	*fe/fd (7)*	*estimator (8)*	*st. errors (9)*	*Hausman test (10)*	*exogeneity assumption addressed? (11)*	*lagged endogenous? (12)*	*lagged exogenous? (13)*
Villareal	*ASR*	2002	1811	3	re/fe	no	fe	gls/ols	gls/ols	yes	no	no	no
Moller et al.	*ASR*	2003	14	~4	re	no	n/a	ols	robust	no	no	no	no
Beckfield	*ASR*	2003	90	2 & 6	re	no	n/a	gls	gls	no	no	yes	yes

Notes: In columns (5) and (7), "re" stand for random effects; "fe" for fixed effects, "fd" for first difference, and "n/a" is not applicable. In columns (8) and (9), "ols" is ordinary least squares, "gls" is generalized least squares, "ml" is maximum likelihood, "iv" is instrumental variables, and "pcse" is panel-corrected standard errors

Table 2. GLS random effects and OLS fixed effects parameter estimates and Hausman test statistics for short and long versions of wage equations, full-time employed males, NLSY 1980-1987 (N=544; T=8).

	model 1	*model 2*		*model 3*		*model 4*		*model 5*
Independent Variables	*fixed-effect estimates*	*GLS random effect estimates*	*gls-fe difference*	*GLS random effect estimates*	*gls-fe difference*	*GLS random effect estimates*	*gls-fe difference*	*GLS random effect estimates*
occupational status (SEI/10)	.037 (6.40)	.046 (8.67)	.009 (4.06)	.046 (8.56)	.009 (4.06)	.039 (7.11)	.002 (0.79)	.037 (6.40)
union (=1)	.083 (3.93)	.121 (6.22)	.038 (4.95)	.124 (6.39)	.041 (4.95)	.124 (6.41)	.041 (4.72)	.083 (3.93)
schooling (years)	--	--	--	--	--	.064 (7.30)	--	.056 (5.62)
black (=1)	--	--	--	-.140 (2.89)	--	-.130 (2.77)	--	-.150 (3.17)
SEI time mean								.029 (1.71)
union time mean								.256 (4.87)
Wald chi-square for GLS models		104.37 (.0000)		113.54 (.0000)		171.21 (.0000)		195.96 (.0000)
Hausman chi-square			45.32 (.0000)		46.75 (.0000)		24.62 (.0000)	

Note: Appearing in parentheses below the coefficients are the absolute values of the t-ratios; below the differences in GLS and FE estimates are Wald chi-squares.

Table 3 Two-step fixed effect/GLS estimates and full Hausman/Taylor estimates of NLSY wage regression, full-time employed males, 1980-1987 (N=544; T=8).

	model 1	*model 2*	*model 3*
Independent Variables	*GLS random effect estimates*	*Two-step FE & GLS random effect estimates*	*Hausman-Taylor estimates*
socioeconomic status (SEI/10)	.039 (7.11)	.037 (6.40)	.037 (6.40)
union (=1)	.124 (6.41)	083 (3.93)	083 (3.93)
schooling (years)	.064 (7.30)	.064 (7.35)	.068 (3.44)
black (=1)	-.130 (2.77)	-.125 (2.61)	-.124 (2.58)
Wald chi-square for GLS models	171.21 (.0000)		98.9 (.0000)